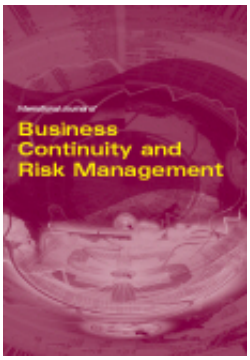


International Journal of Business Continuity and Risk Management

ISSN online: 1758-2172 - ISSN print: 1758-2164
https://www.inderscience.com/ijbcrm

DDoS analysis using machine learning: survey, issues, and future directions

Lalmohan Pattnaik, Suneeta Satpathy, Bijay Kumar Paikaray, Pratik Kumar Swain

DOI: 10.1504/IJBCRM.2024.10060876

Article History:

Received:	15 May 2023
Last revised:	05 July 2023
Accepted:	07 July 2023
Published online:	06 March 2024

DDoS analysis using machine learning: survey, issues, and future directions

Lalmohan Pattnaik

Faculty of Emerging Technologies,
Sri Sri University,
Cuttack, India
Email: lalmohan.p@srisriuniversity.edu.in

Suneeta Satpathy

Center for AI & ML,
SOA University,
Odisha, India
Email: suneeta1912@gmail.com

Bijay Kumar Paikaray*

Center for Data Science,
SOA University,
Odisha, India
Email: bijaypaikaray87@gmail.com
*Corresponding author

Pratik Kumar Swain

Faculty of Emerging Technologies,
Sri Sri University,
Cuttack, India
Email: pratikkumarswain.official@gmail.com

Abstract: Technology has evolved as humanity's new religion in this generation. With everyone switching to online services for their work during the COVID-19 pandemic, digitisation increased more sharply afterwards. The distributed denial of service (DDoS) assault is one of many online dangers that needs to be taken seriously by companies or customers offering cloud services or in need of services respectively. Such threats make the customers deprived of cloud services by overburdening the network with the number of packets causing the shutdown of cloud services. In order to trick current detection systems, attackers are also evolving with the technologies and modifying their attack strategies. Every day, enormous amounts of data are produced, processed, and stored, with typical detection technologies unable to identify new and sophisticated DDoS attacks. This research study thoroughly examines the previous work on DDoS threat analysis using machine learning, as well as its difficulties and potential future applications.

Keywords: denial of service; DoS; distributed denial of service; DDoS; machine learning; cloud service.

Reference to this paper should be made as follows: Pattnaik, L., Satpathy, S., Paikaray, B.K. and Swain, P.K. (2024) 'DDoS analysis using machine learning: survey, issues, and future directions', *Int. J. Business Continuity and Risk Management*, Vol. 14, No. 1, pp.57–76.

Biographical notes: Lalmohan Pattnaik is currently pursuing his PhD at Sri Sri University, Cuttack, India. He received his MTech in Computer Science Engineering from BPUT, Odisha, India in 2010. He is presently working as an Assistant Professor in Faculty of Emerging Technologies in Sri Sri University, Cuttack, Odisha. He has an overall 13 yrs. of teaching experience. His research areas include security, cloud, and ML.

Suneeta Satpathy holds the position of Associate Professor at Sri Sri University Cuttack Odisha. She earned her MCA degree from OUAT BBSR Bhubaneswar, India, and PhD from Utkal University Bhubaneswar, Odisha. Her areas of research include computer forensics, digital forensics, cyber security, data fusion, data mining big data analysis, decision mining, and machine learning.

Bijay Kumar Paikaray is currently working as an Associate Professor in the Department Center for Data Science, SOA University, Odisha, India. His areas of research include high-performance computing, information security, machine learning and IoT.

Pratik Kumar Swain is currently pursuing his BTech in Computer Science specialisation in cyber security at Sri Sri University, Cuttack, India. His research areas include cyber security, cloud computing, digital forensics and machine learning.

1 Introduction

In the last five years, there has been a considerable increase in the use of cloud services, which has resulted in a surge in cloud threats that target both users and cloud service providers. Users can access computational resources remotely thanks to cloud services, which eliminate the need for local servers or personal computers. Infrastructure as a service (IaaS), platform as a service (PaaS), and software as a service (SaaS) are the three main categories under which cloud services fall. IaaS gives customers control over their operating systems and applications by making computer resources available on a pay-per-use or free basis. Web-based application deployment and development environments are provided by PaaS, allowing clients to concentrate on developing their apps without worrying about the supporting infrastructure (Arshi et al., 2020). Through web browsers, SaaS provides full access to software programs over the internet. Numerous advantages, including scalability, flexibility, cost-effectiveness, and accessibility, are offered by such services. Threats to cloud security, which might jeopardise the confidentiality, integrity, and availability (CIA) of cloud resources, rank as one of the major problems with cloud services. Cloud security procedures are used to protect the CIA's online assets (Mahjabin et al., 2017; Wankhede and Kshirsagar, 2018).

A variety of techniques and procedures are included in cloud security to safeguard the data and resources used by cloud computing. This covers issues with access control, data protection, and compliance in addition to guaranteeing the CIA of data (Nassif et al., 2021; Butt et al., 2020).

The article presents an overview of the challenges encountered by cloud services in Section 2. Subsequently, in Section 3 discussion of the risks associated with cloud security is provided. Section 4 narrates a brief explanation of DoS and DDoS attacks in cloud services. Further, Section 5 details the machine learning algorithm and its necessity for DDoS analysis with a systematic literature review of all articles that fall in the said domain. Analysis and Interpretation of the literature reviewed are presented in Section 6 followed by the concluding remarks and future of the study in Section 7.

2 Challenges encountered in cloud services

A variety of technical and non-technical issues can affect the dependability, security, and performance of cloud services.

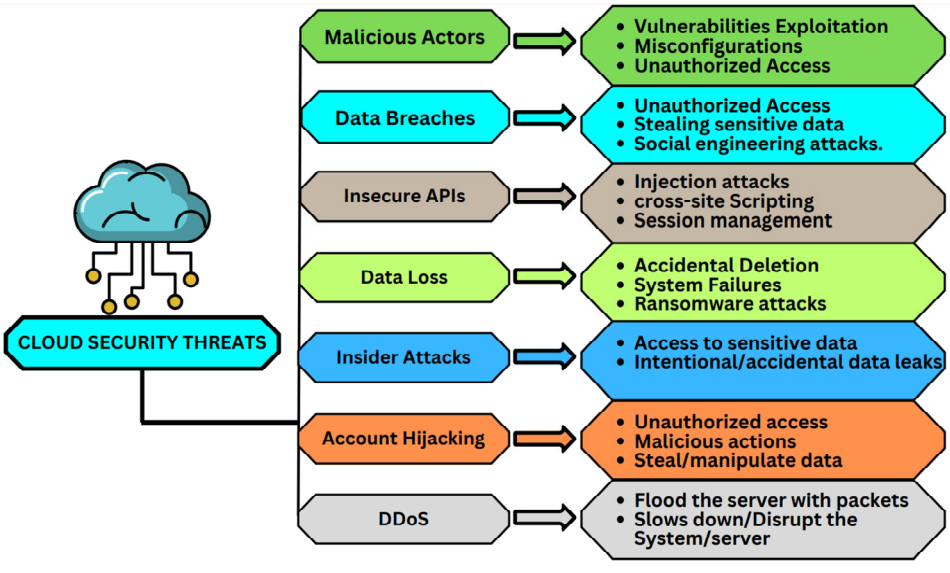
- *Security threats*: These are among the most frequent issues that cloud services encounter (Yang et al., 2020; Oginga and Masese, 2022). Cloud services must guarantee the security and legal compliance of the data and applications hosted on their platforms. They must also take precautions against data breaches, cyberattacks, and other security risks.
- *Scalability*: To satisfy changing customer needs, cloud services must be able to scale their resources up or down fast. In order to make sure that resources are deployed effectively, this calls for a high degree of automation and flexibility.
- *Network performance*: Cloud services rely on quick and dependable network connections to provide data and apps to consumers. Any interruptions or latency problems might have a negative impact on user experience and performance.
- *Vendor lock-in*: Since cloud services frequently call for the usage of proprietary tools and technology, switching providers or integrating with other systems may be challenging for consumers (Kumar and Kumar, 2022).
- *Cost control*: Depending on the available resources, cloud services may be expensive, and costs may rise quickly as more clients use more resources (Potluri et al., 2020). Cloud service providers must provide transparent pricing and use policies as well as cost-effective solutions.
- *Data portability and migration*: Users must be able to move their data and applications across cloud service providers or between on-premises and cloud environments (Nayak et al., 2022). Standards for data portability and compatibility between various cloud systems are necessary for this.
- *Service level agreements (SLAs)*: Reliable SLAs that ensure uptime, performance, and data accessibility must be provided by cloud service providers (Achilleos et al., 2019). Additionally, they must make it simple for users to track and report on SLA compliance, as well as give transparency around SLA data.

From the aforementioned issues, security risks provide the biggest problem for cloud services, providers of cloud services and their clients.

3 Cloud security threats

Cloud security threats encompass risks that have the potential to impact the CIA of data, resources, and computing environments (Paikaray et al., 2020). Malicious individuals can exploit new attack vectors and vulnerabilities that emerge with the adoption of cloud services. These risks encompass data breaches, unauthorised access, data loss, unreliable APIs, insider threats, account takeovers, denial of service (DoS), distributed denial of service (DDoS) attacks, as well as the exploitation of vulnerabilities in cloud applications and infrastructure, as illustrated in Figure 1.

Figure 1 Threats to cloud services (see online version for colours)



4 DoS and DDoS attack

In a DoS attack, a single computer system is employed to inundate a specific server or system with an excessive amount of traffic, leading to its gradual degradation and eventual inability to function (Paikaray et al., 2020; Moreno-Vozmediano et al., 2019). There exist various methods to execute DoS attacks aimed at compromising the availability of the targeted server or system. Figures 2 and 3 provide an illustration of the common taxonomy of DoS attacks.

Figure 2 DoS attack (see online version for colours)

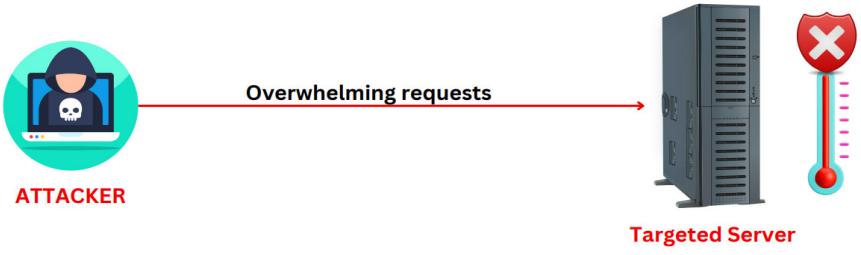


Figure 3 Taxonomy of DoS attacks (see online version for colours)

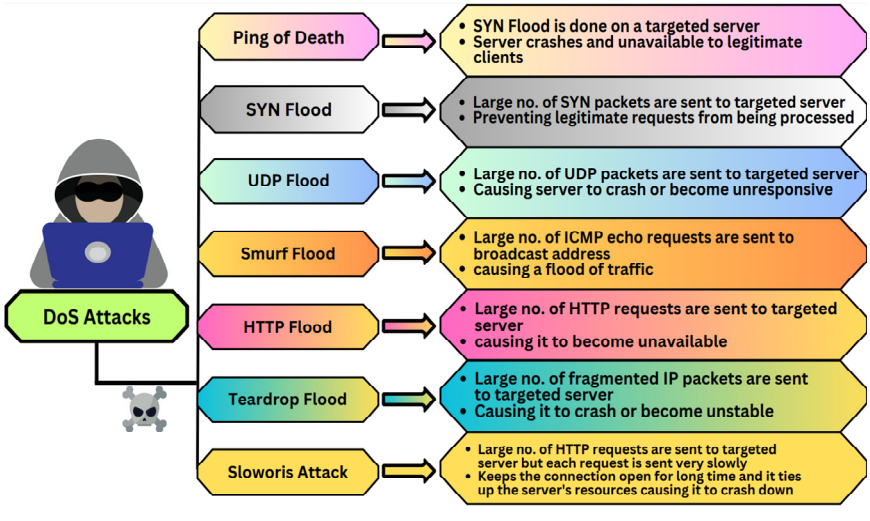
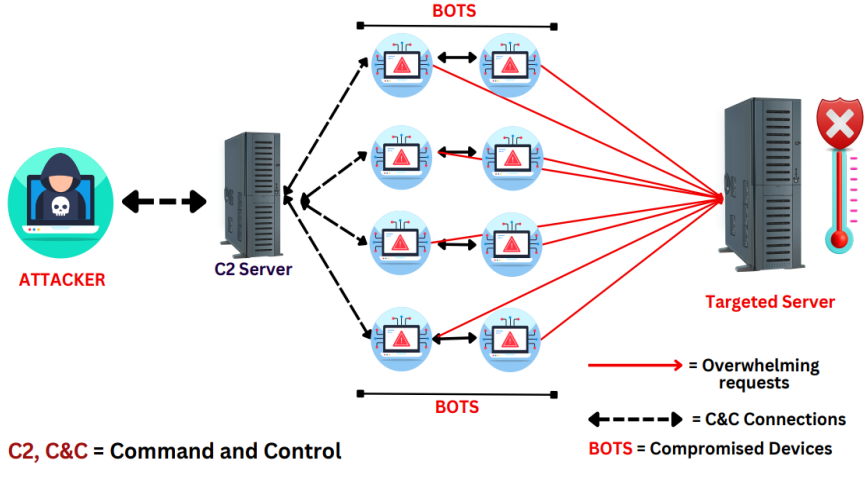


Figure 4 DDoS attack (see online version for colours)



In assaults like DDoS, a botnet or computer network is used by the attacker in order to mount a coordinated attack against the target system. A no. of compromised devices that ends up forming the botnet may be under the attacker's control. All of the botnet's devices simultaneously flood the target system with requests, exhausting its resources and making it unavailable.

Figures 4 and 5 give a brief taxonomy of DDoS threats that comes with different flavours of assaults.

4.1 Composition of DoS and DDoS attack

While DDoS assaults are more complex and involve more sources of attack than DoS attacks, they have a similar anatomy (Tabrizchi and Rafsanjani, 2020).

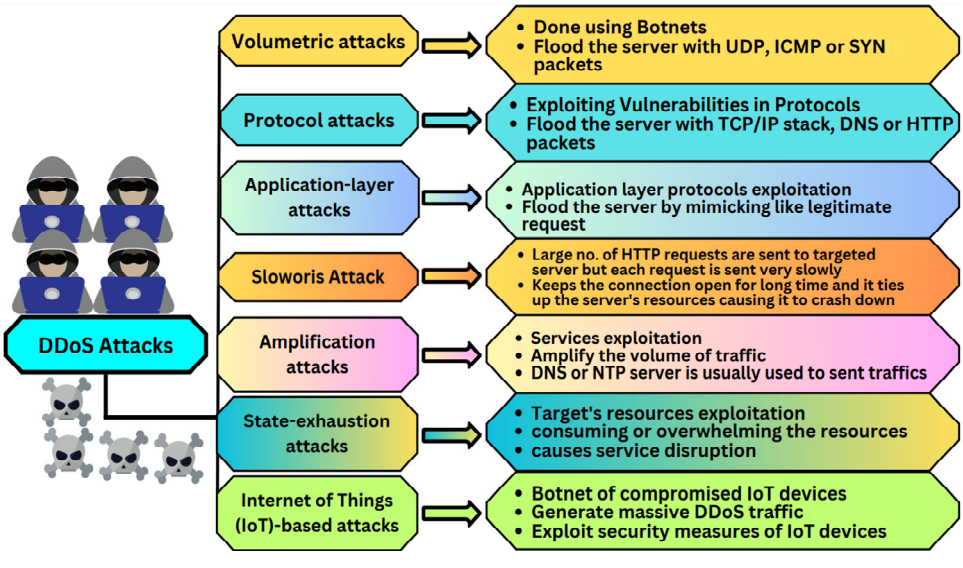
The following steps are commonly included in the anatomy of a DoS attack:

1 *Inspection*: The assailant locates the target system's weaknesses.

2 *Exploitation*: The attacker floods the target system with traffic, using up all of its resources and making it inoperable.

3 *DoS*: This type of threat occurs when genuine users are unable to access the target system.

Figure 5 Types of DDoS attacks (see online version for colours)



The anatomy is similar to a DDoS attack, however, there are additional procedures to take:

1 *Creating a botnet*: The hacker infects a huge number of computers with malware that allows for remote control.

2 *Command and control*: To plan the attack and select the victim, the attacker transmits commands to the botnet.

3 *Attack launch*: The victim is subjected to a coordinated attack by the botnet, which inundates it with an overwhelming volume of traffic and consumes its resources.

4 *Amplification*: The attacker uses amplification techniques like DNS amplification, to magnify the traffic volume sent to the target.

There exist various defence strategies that can be employed to safeguard against DoS and DDoS attacks (Sureshkumar and Baranidharan, 2021). Presented below are a few instances of prevalent defence mechanisms:

The target system bandwidth can be increased to defend against assaults like DoS and DDoS. To achieve this, the architecture can be augmented with features like extra servers, content delivery networks, load balancers, etc.

To mitigate the impact of DoS and DDoS attacks, network firewalls serve as an effective measure by filtering network traffic. By employing network firewalls, the passage of malicious packets to the targeted machine can be obstructed (Ranjan et al., 2015). Additionally, it is possible to restrict communications based on relevant factors like IP addresses of source and destination, port numbers, etc.

Network traffic can be monitored by using intrusion detection and prevention systems (IDPS) and activity patterns or artefacts can be identified as indicators of compromise for DoS and DDoS attacks. Following that, measures can be taken to stop the traffic or inform the system admin.

Content distribution networks (CDNs) are used for distributing content among several servers in various regions. Such type of network can make it more difficult for attackers to focus on a single point of failure.

DoS and DDoS assaults can be avoided by restricting the number of joining points that can be made to the target system. Use of rate limitation or connection throttling can be used to accomplish this. DDoS attacks can be defended against by using anti-DDoS services, which are specialised services. To detect and stop malicious activity, they often combine network monitoring, traffic filtering, and other methods.

These are only a few illustrations of defence strategies that can be employed to fend off DoS and DDoS attacks. To ensure the maximum level of protection, it is crucial to have a comprehensive strategy that incorporates numerous levels of defence. Machine learning algorithms for DDoS attack detection:

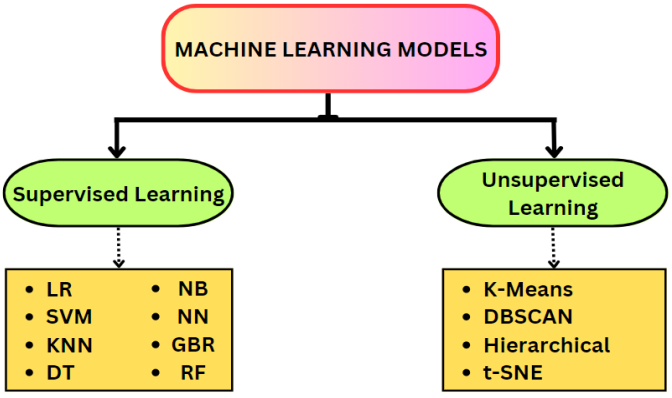
Machine learning allows computers to acquire knowledge from data and predict and give judgements without having to be explicitly programmed. It involves using datasets to teach computer systems to find patterns and relationships, resulting in increased performance and accuracy. ML which includes supervised, unsupervised, semi-supervised, and reinforcement learning, find application in diverse domains such as fraud detection, natural language processing, autonomous vehicles, image recognition and recommendation systems.

5 ML algorithms

Computers can acquire knowledge from data and generate hypotheses or draw conclusions about that data by utilising mathematical models, also known as machine learning algorithms. These algorithms, which form the basis of machine learning, enable computers to develop and learn over time shown in Figure 6.

Numerous machine learning algorithms exist, each with its own advantages and drawbacks. Below is a list of a few of the DDoS detection algorithms.

Figure 6 Classification of ML models (see online version for colours)



Using one or more input variables, a continuous output value can be predicted using the linear regression method. Machine learning methods identify the most suitable line of best fit that effectively captures the relationship between the input and output variables.

- *Linear regression* is a statistical modelling technique used to understand the relationship between a dependent variable and one or more independent variables. It assumes a linear relationship between the variables and finds the best-fit line that minimises the sum of squared residuals. The line can be used to make predictions or infer the impact of changes in the independent variables on the dependent variable.
- *Decision trees* method builds a model of decisions and their outcomes that resembles a tree. It is frequently applied to client segmentation and data mining.
- *Random forests* order to increase accuracy and decrease overfitting, the random forest method builds numerous decision trees and aggregates their output.
- *Support vector machines* (*SVM*) are used for classification and regression tasks. It creates a hyperplane in a high-dimensional space to separate data points into different classes, maximising the margin between them. SVM is effective for handling both linearly separable and nonlinearly separable datasets.
- *The naive Bayes* algorithm, which is based on the Bayes theorem, is used for probabilistic classification. Spam filtering and natural language processing both frequently employ it.
- *K-nearest neighbours* classification technique groups data according to the consensus of its k-nearest neighbours. It is frequently employed in anomaly detection and recommender systems.
- *Neural networks* are to replicate the composition and functionality of the human brain. Natural language processing, speech recognition, and image recognition are just a few of the many applications it is used in.

Data points are grouped together using the clustering method based on how similar they are. It is frequently employed in consumer profiling and market segmentation. These are just a few examples of frequently used machine learning methods for DDoS attack classification or detection.

5.1 Systematic literature review protocol

A systematic literature review methodology is a detailed procedure for conducting an exhaustive and methodical review of pertinent papers on a particular research issue or topic. The protocol of a systematic literature review plays a crucial role in ensuring that the review process is open, reproducible, and objective.

The protocol for the systematic literature review in this paper is as follows.

The major objective of doing a systematic review is to identify and formulate the research questions that will guide the review process. The examination of the information acquired from the final selection of applicable research papers will next be used to answer these questions. The research inquiries that were considered for this review are listed below:

1 What are the latest cutting-edge machine learning algorithms for analysing DDoS attacks, along with their limitations and constraints?

2 What kind of DDoS attacks may be analysed using machine learning techniques?

3 What kinds of features can be derived from network traffic data to be used in machine learning DDoS analysis?

4 What additional features from network traffic data may machine learning be utilised to extract for DDoS analysis?

5 What are the most effective techniques and potential applications of machine learning for DDoS analysis of network traffic data?

6 What privacy and ethical considerations should be made and how may they be addressed when using machine learning to analyse DDoS?

7 How might machine learning be used with other approaches, such as rule-based techniques and anomaly detection, to improve DDoS analysis?

8 What potential academic and industrial applications of machine learning-based DDoS analysis exist, and what implications do they have for security?

5.2 Search strategy

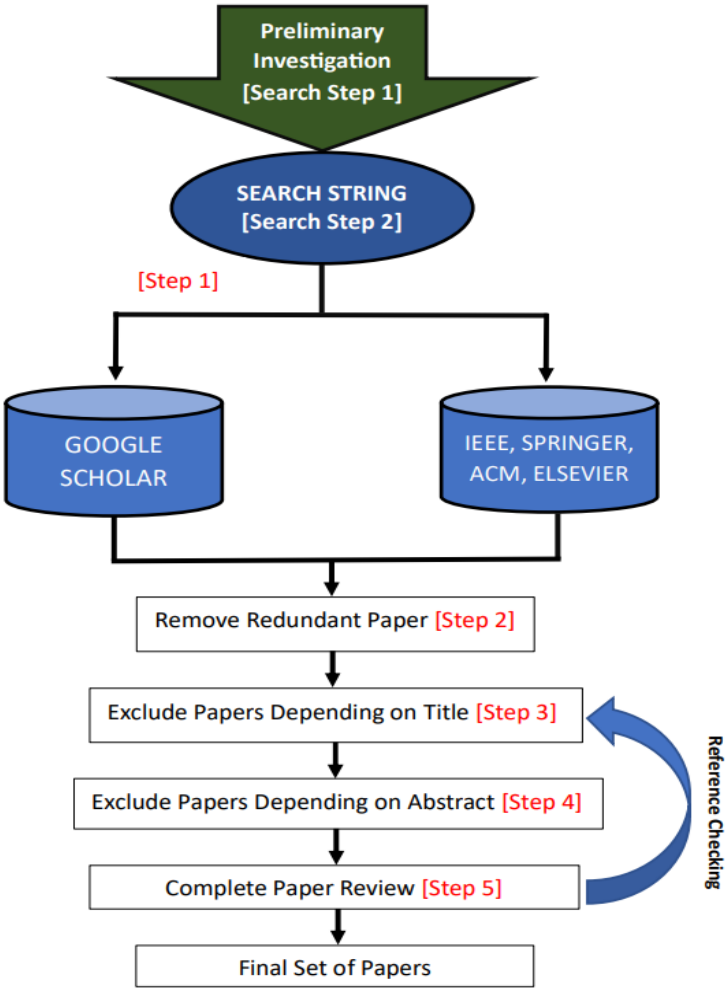
Starting a thorough search procedure will allow for the completion of a systematic survey. For a systematic survey to be successful, a good search strategy must be developed, which requires selecting a number of relevant databases from which to draw relevant material. Between 2018 and 2023, a two-step search process was used for this investigation, with Step 1 employing the digital libraries of the ACM, IEEE, Springer, and ScienceDirect. The academic search engine Google Scholar is used in step two of the search to make sure no relevant literature was overlooked. In order to further hone the search term, the researchers selected ten highly cited and relevant articles. Several digital libraries utilised search queries such as ‘Detection of DDoS attacks using machine

learning' or 'Machine learning approach for DDoS detection', with slight variations. The search outcomes from the selected digital libraries were further improved using filtering methods. Figure 7 depicts how several steps were carried out during the survey.

5.3 Study selection process

The primary objective of study selection in this research was to filter out irrelevant material that did not align with the specified research objectives. Included in the selection were studies that built upon previous pertinent research. In search step 1, a total of 4,506 entries were formed by combining the first 1,354 entries from search step 2 with the 3,152 entries collected in the initial search phase. Following the removal of 214 duplicate items from the previous stage, articles were further eliminated based on their titles (3251), abstracts (758), and full texts (283). Ultimately, after the fifth round of screening, a total of 46 research publications were chosen for inclusion in the study.

Figure 7 SLR for DDoS analysis (see online version for colours)



Inclusion criteria

- The selection process only included papers that were relevant.
- Research findings outlining a machine learning technique for recognising DDoS attacks.
- Research findings addressing the study's objectives.
- Research findings building on earlier research in the field, and investigations that were closely related but differed in some important ways are considered independent primary investigations.
- Items released between 2018 and 2021.

Exclusion criteria

- Redundant research studies and articles with insufficient information.
- Other languages except the English language are used for writing papers.
- Study subjects, critiques, editorial pieces, discussions, articles presenting data, brief communications, publications related to software, encyclopedic entries, posters, abstracts, tutorials, ongoing research, keynote presentations, and invited speeches unrelated to the study topic.

5.4 Reference checking

After reviewing the entire texts of the 32 papers that were selected for the study, the references were evaluated to make sure no relevant material had been missed. 76 more publications were discovered throughout the evaluation, and they were evaluated for conformity with the inclusion and exclusion criteria using their titles, abstracts, and full texts. As a result of this approach, 12 articles were excluded based on full texts, 51 articles were eliminated based on abstracts, and 11 articles were discarded based on titles. After these filtering steps, a total of 71 articles were removed, leaving only two articles for reference checking and final selection.

5.5 Data extraction

In order to address the study's objectives, each article was meticulously read to gather the necessary data. Relevant information was then extracted and recorded in a pre-designed form that includes various fields, such as article title, methodology, datasets utilised, number of features, identification of attack and legitimate classes, preprocessing strategy, experiment setup/model performance optimisation, performance metrics, strengths and weaknesses, along with a summary of the article. The utilisation of these fields facilitated the investigation of research questions and the critical evaluation of the compiled articles. For a detailed description of the data extraction fields, please refer to Table 1.

Table 1 Review of machine learning application for DDoS analysis

Year	*Algorithm used*	*Dataset*	*Results*
Amrish et al. (2022)	ANN, KNN, decision tree, RF	CICDDoS-2019	The most accurate model was the artificial neural network, which scored 99.95% accuracy.
Liu et al. (2022)	SVM and LR	CICDDoS-2019	Compared to logistic regression, the SVM model performs better.
Islam et al. (2022)	Random forest, KNN, and SVM	Own simulated dataset	SVM performed best with an accuracy of 99.5%.
Sudar et al. (2021)	Decision tree and SVM	KDD-99	Greater accuracy and detection rates were achieved by both DT and SVM.
Lucky et al. (2020)	Decision tree (light weight)	CICDDoS-2019 CICDDoS-2017	99.9% accuracy detected.
Pande et al. (2021)	Random forest	NSL-KDD	A 99.76% accuracy rate is achieved.
Priya et al. (2020)	Random forest, KNN, and NB	Own simulated dataset	A 98.5% accuracy rate is attained.
Nadeem et al. (2021)	Random forest, KNN, SVM, decision tree, NB	NSL-KDD	The best accuracy rate was achieved by RF, which was 99.97%.
Saghezchi et al. (2022)	All supervised and unsupervised algorithms	Own simulated dataset	With a 99.9% accuracy rate, DT did the best.
Wani et al. (2019)	Random forest, SVM and NB	KDD-99	SVM performed the best, with an accuracy rate of 99.7%.
Bindra and Sood (2019)	Random forest, KNN, GNB and SVM	CICIDC-2017	RF (96% accuracy) had the highest accuracy.
Aysa et al. (2020)	Random Forest, NN, SVM and decision tree	Own simulated dataset	High accuracy in detection was achieved using RF and DT in combination.
Sambangi and Gondi (2020)	Multiple LR	CICIDC-2017	97.86% accuracy is found.
Khuphiran et al. (2018)	DFF and SVM	DARPA-2009	SVM became the best performer with an accuracy of 93.01%
Yungaicela-Naula et al. (2021)	RF, SVM, KNN, MLP, LSTM, GRU, CNN	CICDDoS-2017 CICDDoS-2019	GRU accomplished better with 99.94% accuracy
Polat et al. (2020)	ANN, NB, SVM, KNN	Own simulated dataset	KNN accomplished an accuracy of 98.3% which was better in comparison to other
Amjad et al. (2019)	NB and random forest	KDD-99	NB outperformed NB in terms of detection.
Radivilova et al. (2019)	Machine learning algorithms	Own simulated dataset	RF showed a better performance.

Table 1 Review of machine learning application for DDoS analysis (continued)

Year	*Algorithm used*	*Dataset*	*Results*
Sarraf (2020)	LSVM and decision tree	CICIDS-2017	DT's achievement was 100% accuracy.
Mishra et al. (2021)	NB, KNN and random forest	Own Simulated Dataset	99.76% accuracy gained.
Aytaç et al. (2020)	Random forest, ANN, KNN, NB, KNN, SVM and decision tree	CICDDoS-2019	SVM outperformed with a 99.7% accuracy rate.
Peneti and Hemalatha (2021)	Multilayer perceptron, random forests, XGBoost and AdaBoost	CICIDS-2017	RF performance was best.
Gaur and Kumar (2022)	Random forest, decision tree, KNN and XGBoost	CICDDoS-2019	The accuracy of XGBoost (ANOVA), which performed better, was 98.34%.
Mohmand et al. (2022)	Random forest and XGBoost	UNSW-NB 15	XGBoot did better, with an accuracy of 90%.
Manikumar and Maheswari (2020)	Random forest, decision tree, and KNN	Own simulated dataset	RF fared better, with a 95% accuracy rate.
Nakka and Devi (2023)	Random forest, NB and SVM	UNSW-NB15 UNB ISCX 12 NSL-KDD	RF gave better results with 99% accuracy
Gurulakshmi and Nesarani (2018)	Support vector machine	Own simulated dataset	SVM did a better job of foreseeing anomalous activity.
Yusof et al. (2018)	SVM merged with PTA (Packet threshold algorithm)	Own simulated dataset	PTA-SVM was more effective at spotting DDoS attacks.
Aslam et al. (2022)	NB, SVM, random Forest, LR and KNN merged into a framework as AMLSDM-EV), Static AMLSDM	Own simulated dataset	AMLSDM-EV maintained a better performance than Static AMLSDM-(SVM/NB/RF/KNN/LR)
Alzahrani and Alzahrani (2021)	Random forest, LR, KNN, NB, Decision tree and SVM,	CICDDoS-2019	DT & RF are acquired 99% of the time, however, DT is superior because of reduced computational time.
Machaka et al. (2022)	SVR, LGR, ANN and K-means	DARPA IDS	94% accuracy achieved.
Abbas and Almhanna (2020)	Random forest and NB	CICDDoS-2019	RF achieved an accuracy of 99.9% and was the better performer
Sambangi Sambangi (2020)	Multiple linear regression	CICIDS-2017	MLR achieved a precision of 97.86%.

Table 1 Review of machine learning application for DDoS analysis (continued)

Year	*Algorithm used*	*Dataset*	*Results*
Doshi et al. (2018)	Neural network, KNN, decision tree, random forest, LSVM, decision tree	Own simulated dataset	DT fared best, achieving 99% accuracy.
Saini et al. (2020)	WEKA, J48, random forest and NB	Own simulated dataset	RF & NB were outperformed by J48 in terms of outcomes.
Shaaban et al. (2019)	Neural networks, convolutional neural networks, KNN, SVM	Own simulated dataset NSL-KDD	CNN did the best, with 99% accuracy.
Rimal and Rajapraveen (2020)	Support vector machine and NB	CICIDS-2017	SVM fared the best, with a 99.68% accuracy rate.
Pei et al. (2019)	SVM and random forest	Own simulated dataset	RF got the results with an accuracy of 99%
Idhammad et al. (2018)	Information gain ratio, network entropy estimation, extra trees algorithm and co-clustering	NSL-KDD UNB-ISCX12 UNSW-NB15	The NSL-KDD, UNB ISCX 12, and UNSW-NB15 datasets have accuracy levels of 98.23%, 99.88%, and 93.71%, respectively.
Tuan et al. (2020)	Support vector machine, ANN, NB, decision tree, unsupervised algorithms	UNBS-NB15 KDD-99	USML outperformed other models with accuracy ratings of 94.78% in the UNBS-NB-15 and 98.08% in the KDD99.
Nalayini and Katiravan (2022)	SVM, LR, KNN, RF, DT and NB	CICIDS-2017	99.88% of the performance of RF was accurately measured.

6 Analysis and interpretation

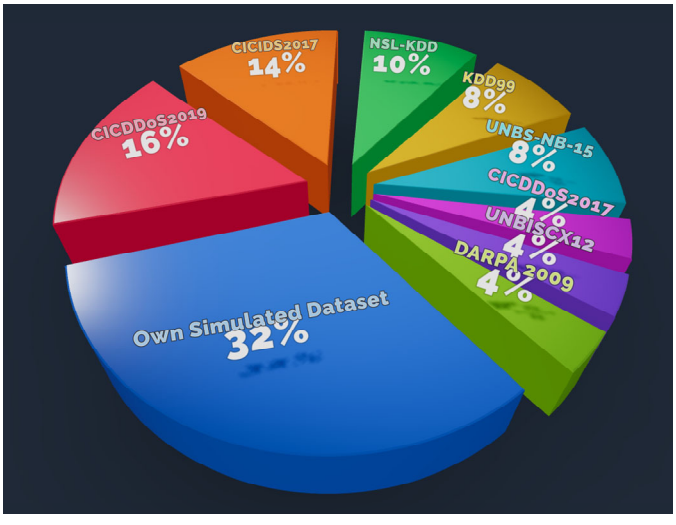
This section looks at works that looked into supervised and unsupervised algorithms for detecting DDoS assaults in the context of ML in cloud security. We used a search approach strategy that entails eliminating plenty of irrelevant papers in order to reduce the number of publications that exactly meet our study objective. We also employed quality evaluation criteria to ensure that the selected papers provide results that have been synthesised. Detecting DDoS attacks can be difficult due to different variants and attack patterns, making it challenging to differentiate them from normal traffic. To address this issue, researchers have developed multiple machine learning and deep learning techniques for DDoS attack detection over time. However, these approaches have notable limitations since adversaries continuously evolve their strategies. Leveraging the insights from this analysis, Based on ML/DL methodologies, we assessed and ranked the most recent cutting-edge DDoS attack detection systems.

Table 1 presents a compilation of relevant research in DDoS attack detection, based on the suggested classification for ML/DL techniques. The corresponding findings of each study are included. Several studies have reported accuracy rates exceeding 99%.

However, it is important to note that the majority of these evaluations were conducted using offline data analysis. It should be recognised that performance metrics may vary when applied to real-time case studies. Furthermore, it should be highlighted that different assessment methods and datasets were employed in these investigations, making it difficult to compare the outcomes. Figure 8 illustrates the datasets that were most commonly utilised in the literature.

In comparison, self-formed simulated datasets are used as 32% of total studies, CICDDoS2019 dataset is used as 16% of the total studies, CICIDS2017 dataset used as 14% of the total studies, NSL-KDD dataset used as 10% of total studies, KDD99 dataset is used as 8% of the total studies, UNBS-NB-15 dataset is used as 8% of the total studies, CICDDoS2017 dataset used as 4% of the total studies, UNBISCX12 dataset used as 4% of the total studies and DARPA 2009 dataset is used a 4% of the total studies. SVM, Decision Tree, ANN, Random Forest, USML, GRU and CNN achieved 99% accuracy.

Figure 8 Usage of datasets (see online version for colours)



The results of our investigation on performance metrics also pointed to the following. While 22 researchers each utilised the measures for recall, precision, and F1-score, only six studies each employed the FPR and AUC metrics. 36 of the publications that were assessed employed accuracy evaluations to evaluate their approaches. It can be shown that most research did not include information on how long their methods took to test or train, despite the fact that this information is essential for deploying the system in real-world or production environments.

Based on our review of ML/DL methods for DDoS attack identification in this study, the results highlight the following areas that require further investigation in future work:

- *Failure to apply in the real world*: The majority of research has focused on dissecting these models without evaluating how well they work in actual DDoS attack scenarios. As a result, there is a pressing need for ML/DL models that have been proven reliable in practical applications.

- *Models based on machine learning and deep learning that get dynamically updated*: Having models that can be regularly updated is crucial to effectively identify new forms of attacks, as attack patterns are constantly changing and evolving. But as of right now, there are no DL models that meet this condition in the literature.
- *Compact ML/DL model design*: Networks like the internet of things, MANETS, and wireless sensor networks have a finite amount of computing power and memory. Therefore, it is essential to have compact models that are capable of identifying security issues. Future work will focus on creating adaptable and effective DL models for these situations.
- *The occurrence of appropriate datasets*: The current datasets lack sufficient diversity in terms of attack types and data quality, leading to biased detection systems that are unable to recognise all types of attacks. As a result, having a variety of datasets is essential to ensuring reliable and efficient detection models.

To achieve substantial advancements in the field of cyber threat analysis and its subsequent identification and prevention, it is imperative to overcome the above-mentioned hurdles.

7 Conclusions and future direction

Attacks by DDoS, which can hurt internet users in a number of ways, will continue to be a severe danger to many big and small enterprises. Few regions must be given top attention because of the dearth of publicly accessible data, the length of calculation times, and the potential for human operators to detect DDoS attacks. Finding sophisticated attacks with unexpected patterns is one of the hardest challenges on the Internet since conventional threat detection technologies cannot accomplish it. In this work, we did a thorough analysis of the literature with a focus on distributed and DDoS assaults. To recognise DDoS attacks, several machine learning methods are utilised. Given how simple and inexpensive it is to launch one, such assaults have already beyond a critical threshold, and their frequency is only predicted to rise. Based on current trends in their cost, performance, and availability, it is anticipated that the frequency of DDoS attacks will continue to rise over time. In light of this, the current objective is to develop a dynamic and compact machine learning model that can effectively recognise and evaluate DDoS attacks.

References

Abbas, S.A. and Almhanna, M.S. (2020) 'Distributed denial of service attacks detection system by machine learning based on dimensionality reduction', *ICMAICT 2020*.

Achilleos, A.P., Kritikos, K., Rossini, A., Kapitsaki, G.M., Domaschka, J., Orzechowski, M., Seybold, D., Griesinger, F., Nikolov, N., Romero, D. and Papadopoulos, G.A. (2019) 'The cloud application modelling and execution language', *J. Cloud Comput*., December, Vol. 8, No. 1, p.20, DOI: 10.1186/s13677-019-0138-7.

Alzahrani, R.J. and Alzahrani, A. (2021) 'Security analysis of DDoS attacks using machine learning algorithms in networks traffic', *Electronics*, Vol. 10, No. 23, p.2919, https://doi.org/10.3390/electronics10232919.

Amjad, A., Alyas, T., Farooq, U. and Arslan Tariq, M. (2019) 'Detection and mitigation of DDoS attack in cloud computing using machine learning algorithm', *EAI Endorsed Transactions on Scalable Information Systems*, Vol. 6, No. 23, p.e7, https://doi.org/10.4108/eai.29-7-2019.159834.

Amrish, R., Bavapriyan, K., Gopinaath, V., Jawahar, A. and Kumar, C.V. (2022) 'DDoS detection using machine learning techniques', *Journal of IoT in Social, Mobile, Analytics, and Cloud*, Vol. 4, No. 1, pp.24–32, DOI: 10.36548/jismac.2022.1.003.

Arshi, M., Nasreen, M.D. and Madhavi, K. (2020) 'A survey of DDOS attacks using machine learning techniques', *E3S Web Conf.*, Vol. 184, p.01052, DOI: 10.1051/e3sconf/202018401052.

Aslam, M., Ye, D., Tariq, A., Asad, M., Hanif, M., Ndzi, D., Chelloug, S.A. et al. (2022) 'Adaptive machine learning based distributed denial-of-services attacks detection and mitigation system for SDN-enabled IoT', *Sensors*, Vol. 22, No. 7, p.2697, MDPI AG, http://dx.doi.org/10.3390/s22072697.

Aysa, M.H., Ibrahim, A.A. and Mohammed, A.H. (2020) 'IoT Ddos attack detection using machine learning', *2020 4th International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT)*, Istanbul, Turkey, pp.1–7, DOI: 10.1109/ISMSIT50672.2020.9254703.

Aytaç, T., Aydın, M.A. and Zaim, A.H. (2020) 'Detection DDOS attacks using machine learning methods', *Electrica*, Vol. 20, No. 2, pp.159–167.

Bindra, N. and Sood, M. (2019) 'Detecting DDoS attacks using machine learning techniques and contemporary intrusion detection dataset', *Aut. Control Comp. Sci*., Vol. 53, pp.419–428, https://doi.org/10.3103/S0146411619050043.

Butt, U.A., Mehmood, M., Shah, S.B.H., Amin, R., Shaukat, M.W., Raza, S.M., Suh, D.Y. et al. (2020) 'A review of machine learning algorithms for cloud computing security', *Electronics*, Vol. 9, No. 9, p.1379, MDPI AG, http://dx.doi.org/10.3390/electronics9091379.

Doshi, R., Apthorpe, N. and Feamster, N. (2018) 'Machine learning DDoS detection for consumer internet of things devices', *2018 IEEE Security and Privacy Workshops (SPW)*, San Francisco, CA, USA, pp.29–35, DOI: 10.1109/SPW.2018.00013.

Gaur, V. and Kumar, R. (2022) 'Analysis of machine learning classifiers for early detection of DDoS attacks on IoT devices', *Arab J. Sci. Eng*., Vol. 47, pp.1353–1374, https://doi.org/10.1007/s13369-021-05947-3.

Gurulakshmi, K. and Nesarani, A. (2018) 'Analysis of IoT bots against DDOS attack using machine learning algorithm', *2018 2nd International Conference on Trends in Electronics and Informatics (ICOEI)*, Tirunelveli, India, pp.1052–1057, DOI: 10.1109/ICOEI.2018.8553896.

Idhammad, M., Karim, A. and Belouch, M. (2018) 'Semi-supervised machine learning approach for DDoS detection', *Appl. Intell*., https://doi.org/10.1007/s10489-018-1141-2.

Islam, U. et al. (2022) 'Detection of distributed denial of service (DDoS) attacks in IoT based monitoring system of banking sector using machine learning models', *Sustainability*, July, Vol. 14, No. 14, p.8374, DOI: 10.3390/su14148374.

Khuphiran, P., Leelaprute, P., Uthayopas, P., Ichikawa, K. and Watanakeesuntorn, W. (2018) 'Performance comparison of machine learning models for DDoS attacks detection', *2018 22nd International Computer Science and Engineering Conference (ICSEC)*, Chiang Mai, Thailand, pp.1–4, DOI: 10.1109/ICSEC.2018.8712757.

Kumar, M.A. and Kumar, K.A. (2022) 'A survey on cloud computing security threats, attacks and countermeasures: a review', *International Journal of Human Computations & Intelligence*, Vol. 1, No. 3, pp.13–18.

Liu, Z., Qian, L. and Tang, S. (2022) 'The prediction of DDoS attack by machine learning', in *Third International Conference on Electronics and Communication; Network and Computer Technology (ECNCT 2021)*, March, pp. 681–686, doi: 10.1117/12.2628658.

Lucky, G., Jjunju, F. and Marshall, A. (2020) 'A lightweight decision-tree algorithm for detecting DDoS flooding attacks', *2020 IEEE 20th International Conference on Software Quality, Reliability and Security Companion (QRS-C)*, Macau, China, pp.382–389, DOI: 10.1109/QRS-C51114.2020.00072.

Machaka, P., Ajayi, O., Kahenga, F., Bagula, A. and Kyamakya, K. (2022) 'Modelling DDoS attacks in IoT networks using machine learning', in *International Conference on Emerging Technologies for Developing Countries*, Springer Nature, Cham, Switzerland, December, Vol. 9, pp.161–175.

Mahjabin, T., Xiao, Y., Sun, G. and Jiang, W. (2017) 'A survey of distributed denial-of-service attack, prevention, and mitigation techniques', *International Journal of Distributed Sensor Networks*, Vol. 13, No. 12, DOI: 10.1177/1550147717741463.

Manikumar, D.V.V.S. and Maheswari, B.U. (2020) 'Blockchain based DDoS mitigation using machine learning techniques', *2020 Second International Conference on Inventive Research in Computing Applications (ICIRCA)*, Coimbatore, India, pp.794–800, DOI: 10.1109/ICIRCA48905.2020.9183092.

Mishra, A., Gupta, B.B., Peraković, D., Peñalvo, F.J.G. and Hsu, C-H. (2021) 'Classification based machine learning for detection of DDoS attack in cloud computing', *2021 IEEE International Conference on Consumer Electronics (ICCE)*, Las Vegas, NV, USA, pp.1–4, DOI: 10.1109/ICCE50685.2021.9427665.

Mohmand, M.I., Hussain, H., Khan, A.A., Ullah, U., Zakarya, M., Ahmed, A. and Haleem, M. (2022) 'A machine learning-based classification and prediction technique for DDoS attacks', *IEEE Access*, Vol. 10, pp.21443–21454, DOI: 10.1109/ACCESS.2022.3152577.

Moreno-Vozmediano, R., Montero, R.S., Huedo, E. and Llorente, I.M. (2019) 'Efficient resource provisioning for elastic cloud services based on machine learning techniques', *J. Cloud Comput*., December, Vol. 8, No. 1, p.5, DOI: 10.1186/s13677-019-0128-9.

Nadeem, W., Hock, G., Goh, Ponnusamy, V. and Aun, Y. (2021) 'DDoS detection in SDN using machine learning techniques', *Computers, Materials and Continua*, Vol. 71, pp.771–789, DOI: 10.32604/cmc.2022.021669.

Nakka, R.D.S.N. and Devi, D. (2023) 'Semi-supervised machine learning approach for DDoS detection', *International Journal of Communication and Computer Technologies*, Vol. 11, No. 2, pp.120–126, DOI: 10.31838/ijccts/11.02.15.

Nalayini, C.M. and Katiravan, J. (2022) *Detection of DDoS Attack Using Machine Learning Algorithms*, 26 July, Vol. 9, No. 7, ISSN-2349-5162, JETIR, SSRN [online] https://ssrn.com/abstract=4173187, http://www.jetir.org.

Nassif, A., Abu Talib, M., Nasir, Q., Albadani, H. and Albab, F. (2021) 'Machine learning for cloud security: a systematic review', *IEEE Access*, p.1, DOI: 10.1109/ACCESS.2021.3054129.

Nayak, S.K., Swain, S.K., Mohanta, B.K. and Paikaray, B.K. (2022) 'Secure framework for data leakage detection and prevention in IoT application', in *2022 IEEE 2nd International Symposium on Sustainable Energy, Signal Processing and Cyber Security (iSSSC)*, IEEE, December, pp.1–6.

Oginga, R. and Masese, N. (2022) 'Evaluating causes of policy violation in cloud environment', *International Journal of Engineering Research & Technology (IJERT)*, January, Vol. 11, No. 1, pp.1–15.

Paikaray, B.K., Dewangan, P., Swain, D. and Chakravarty, S. (2020) 'An extensive study on medical image security with ROI preservation: techniques, evaluations, and future directions', in *Machine Learning and Information Processing: Proceedings of ICMLIP 2019*, Springer, Singapore, pp.465–476.

Pande, S., Khamparia, A., Gupta, D. and Thanh, D.N.H. (2021) 'DDOS detection using machine learning technique', in Khanna, A., Singh, A.K. and Swaroop, A. (Eds.): *Recent Studies on Computational Intelligence. Studies in Computational Intelligence*, Vol. 921, Springer, Singapore, https://doi.org/10.1007/978-981-15-8469-5_5.

Pei, J., Chen, Y. and Ji, W. (2019) 'A DDoS attack detection method based on machine learning', *Journal of Physics: Conference Series*, Vol. 1237, p.032040, DOI: 10.1088/1742-6596/1237/3/032040.

Peneti, S. and Hemalatha, E. (2021) 'DDOS attack identification using machine learning techniques', *2021 International Conference on Computer Communication and Informatics (ICCCI)*, Coimbatore, India, pp.1–5, DOI: 10.1109/ICCCI50826.2021.9402441.

Polat, H., Polat, O. and Cetin, A. (2020) 'Detecting DDoS attacks in software-defined networks through feature selection methods and machine learning models', *Sustainability*, Vol. 12, No. 3, p.1035, https://doi.org/10.3390/su12031035.

Potluri, S., Mangla, M., Satpathy, S. and Mohanty, S.N. (2020) 'Detection and prevention mechanisms for DDoS attack in cloud computing environment', *2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT)*, Kharagpur, India, pp.1–6, DOI: 10.1109/ICCCNT49239.2020.9225396.

Priya, S.S., Sivaram, M., Yuvaraj, D. and Jayanthiladevi, A. (2020) 'Machine learning based DDOS detection', *2020 International Conference on Emerging Smart Computing and Informatics (ESCI)*, Pune, India, pp.234–237, DOI: 10.1109/ESCI48226.2020.9167642.

Radivilova, T., Kirichenko, L., Ageiev, D. and Bulakh, V. (2019) 'Classification methods of machine learning to detect DDoS attacks', *2019 10th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS)*, Metz, France, pp.207–210, Doi: 10.1109/IDAACS.2019.8924406.

Ranjan, R., Swain, D. and Paikaray, B. (2015) 'Efficient key management and cipher text generation using BCD coded parity bits', *Procedia Computer Science*, Vol. 57, pp.703–709.

Rimal, A. and Rajapraveen, K.N. (2020) 'DDOS attack detection using machine learning', *International Journal of Emerging Technologies and Innovative Research*, Vol. 7, No. 6, pp.185–188, ISSN: 2349-5162 [online] http://www.jetir.org.

Saghezchi, F.B., Mantas, G., Violas, M.A., de Oliveira Duarte, A.M. and Rodriguez, J. (2022) 'Machine learning for DDoS attack detection in Industry 4.0 CPPSs', *Electronics*, Vol. 11, No. 4, p.602, MDPI AG, http://dx.doi.org/10.3390/electronics11040602.

Saini, P., Behal, S. and Bhatia, S. (2020) 'Detection of DDoS attacks using machine learning algorithms', DOI: 10.23919/INDIACom49435.2020.9083716.

Sambangi, S. and Gondi, L. (2020) 'A machine learning approach for DDoS (distributed denial of service) attack detection using multiple linear regression', *The 14th International Conference on Interdisciplinarity in Engineering – INTER-ENG 2020*, https://doi.org/10.3390/proceedings2020063051.

Sambangi, S. and Gondi, L. (2020) 'A machine learning approach for DDoS (distributed denial of service) attack detection using multiple linear regression', *The 14th International Conference on Interdisciplinarity in Engineering – INTER-ENG 2020*, MDPI, http://dx.doi.org/10.3390/proceedings2020063051.

Sarraf, S. (2020) 'Analysis and detection of DDoS attacks using machine learning techniques', *American Scientific Research Journal for Engineering, Technology, and Sciences*, Vol. 66, No. 1, pp.95–104.

Shaaban, A., Abd-Elwanis, E. and Hussein, M. (2019) 'DDoS attack detection and classification via convolutional neural network (CNN)', pp.233–238, DOI: 10.1109/ICICIS46948.2019.9014826.

Sudar, K.M., Beulah, M., Deepalakshmi, P., Nagaraj, P. and Chinnasamy, P. (2021) 'Detection of distributed denial of service attacks in SDN using machine learning techniques', in *IEEE Int. Conf. Comput. Commun. Informat. (ICCCI)*, 27 January, pp.1–5, DOI: 10.1109/ICCCI50826.2021.9402517.

Sureshkumar, V. and Baranidharan, B. (2021) 'A study of the cloud security attacks and threats', in *Journal of Physics: Conference Series*, July, Vol. 1964, No. 4, p.042061, IOP Publishing.

Tabrizchi, H. and Rafsanjani, M.K. (2020) 'A survey on security challenges in cloud computing: Issues, threats, and solutions', *J. Supercomput*., December, Vol. 76, No. 12, pp.9493–9532, DOI: 10.1007/s11227-020- 03213-1.

Tuan, T., Long, H., Son, L., Priyadarshini, I., Kumar, R. and Son, N. (2020) 'Performance evaluation of Botnet DDoS attack detection using machine learning', *Evolutionary Intelligence*, Vol. 13, p.3, DOI: 10.1007/s12065-019-00310-w.

Wani, A.R., Rana, Q.P., Saxena, U. and Pandey, N. (2019) 'Analysis and detection of DDoS attacks on cloud computing environment using machine learning techniques', *2019 Amity International Conference on Artificial Intelligence (AICAI)*, Dubai, United Arab Emirates, pp.870–875, DOI: 10.1109/AICAI.2019.8701238.

Wankhede, S. and Kshirsagar, D. (2018) 'DoS attack detection using machine learning and neural network', *2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA)*, Pune, India, pp.1–5, DOI: 10.1109/ICCUBEA.2018.8697702.

Yang, P., Xiong, N. and Ren, J. (2020) 'Data security and privacy protection for cloud storage: a survey', *IEEE Access*, p., DOI: 10.1109/ACCESS.2020.3009876.

Yungaicela-Naula, N.M., Vargas-Rosales, C. and Perez-Diaz, J.A. (2021) 'SDN-based architecture for transport and application layer DDoS attack detection by using machine and deep learning', in *IEEE Access*, Vol. 9, pp.108495–108512, DOI: 10.1109/ACCESS.2021.3101650.

Yusof, M.A.M., Ali, F.H.M. and Darus, M.Y. (2018) 'Detection and defense algorithms of different types of DDoS attacks using machine learning', in Alfred, R., Iida, H., Ag. Ibrahim, A. and Lim, Y. (Eds.): *Computational Science and Technology. ICCST 2017. Lecture Notes in Electrical Engineering*, Vol. 488, Springer, Singapore, https://doi.org/10.1007/978-981-10-8276-4_35.