Advances in Knowledge-Based Systems, Data Science, and Cybersecurity Received 21-06-2026; Accepted 24-08-2026; Published 31-08-2026

Mapping the Evolution of AI-Driven Dynamic Botnet Intrusion Prevention Systems: A Bibliometric Analysis (2005–2025)

Aakanksha aakanksha.1@rajguru.du.ac.in
Shaheed Rajguru College
Vasundhara Enclve Delhi

Anamika Gupta anamikargupta@sscbsdu.ac.in
Shaheed Sukhdev College of Business Studies,
University of Delhi, Delhi

Sanjay Singh sanjay.24725@sscbs.du.ac.in
Shaheed Sukhdev College of Business Studies
University of Delhi, Delhi

Richa Gautam richa.24722@sscbs.du.ac.in
Shaheed Sukhdev College of Business Studies
University of Delhi, Delhi

Corresponding Author: Aakansha

Abstract

The increasing complexity of botnet attacks has driven digital infrastructure protection toward Artificial Intelligence and Machine Learning systems. The research study investigates twenty years of scientific work from 2005 until 2025 to study how worldwide scientists fight against botnet threats, which keep changing. We track the field's rapid development through our analysis of thousands of records. We obtained these from Scopus and Web of Science databases. The research field shows three main groups of innovators who include authors, their corresponding institutions, and nations. Our research shows scientists have switched their detection methods from traditional systems to modern intelligent networks, which use deep learning and federated learning technology. The world now supports more international teamwork than ever before, yet multiple major obstacles continue to exist. The research identifies three critical security needs which will require solutions during the upcoming years. The research shows that security solutions need to become more compact so they can operate effectively on smart devices, and AI systems must be built to handle deliberate attacks. The review encompasses both intrusion detection and prevention literature, as detection forms the analytical foundation of AI/ML based prevention systems. The paper presents the evolution of intrusion detection and prevention systems over the past two decades that will guide future autonomous cyber defense systems development.

Keywords: Bibliometric Analysis, Artificial Intelligence, Machine Learning Intrusion Prevention System, Botnet Attacks

Citation: Aakansha, et al. Mapping the Evolution of AI-Driven Dynamic Botnet Intrusion Prevention Systems: A Bibliometric Analysis (2005–2025). **Advances in Knowledge-Based Systems, Data Science, and Cybersecurity**;Research. 2026;3(2).30.

Page no.-600-629.

1. INTRODUCTION

”Botnets rank among the biggest threats in the modern cybersecurity era that we are living in.” A botnet can be defined as a network of PCs or IoT devices that are connected to the internet and have been infected by software [1]. The operating of this network of PCs is done by a hacker known as a botmaster. The botnets can be formed by automated infection tools such as phishing attacks, malicious URLs, and exploiting software bugs. “Once a device has been recruited into this “zombie army,” it can be used for massive Distributed Denial-of-Service (DDoS) attacks, massive data exfiltration attacks, and assisting with click fraud and spam activities [2].”

The challenge to the defence community has been accentuated by the development of dynamic botnets. The dynamic botnet uses decentralized architectures such as peer-to-peer (P2P) networking and fast flux DNS [3], in contrast, traditional botnets operate under centralized command-and-control (C&C) networks. These characteristics enable the botnet to constantly change its IP addresses and dialogue domains in such a way that it presents a moving target which is very hard to trace and wipe out [4]. As they make use of modular malware and encrypt their communications in order to evade signature-based detection, dynamic botnets are much more resilient and deadly than their static counterparts and may often evade detection in the network for an extended period [5].

In order to counter these sophisticated cyber attacks effectively, organizations have upgraded from passive intrusion detection systems to active intrusion prevention systems (IPS). An intrusion prevention system is designed to monitor activity in real-time and quickly react to it, for example, by deleting malicious packets or closing a malicious connection, to protect the system from an attack before it actually happens. But a generic design for an intrusion prevention system tends to be ineffective when working with dynamic botnets because it mostly uses rule-defined actions and signatures to detect cyber threats. This has resulted in the use of artificial intelligence (AI) and machine learning (ML) in the field of intrusion prevention systems.

AI/ML-based IPS tools are extremely important in the current protection program due to their ability to bring a paradigm shift in their recognition abilities, focusing on behavioral patterns instead of detecting known files. They work on huge network feeds, identifying crucial elements, teaching algorithms how to detect inconsistencies in normal patterns of activities due to the subtle ‘heartbeats’ involved in botnet communication patterns [6]. The AI/ML tools employ deep learning techniques to identify zero-day threats through their algorithms which learn from the ever-changing patterns of dynamic botnets [7].

A botnet is considered dynamic in this review if it exhibits at least one of these five characteristics. This definition aligns with the characterization used in foundational botnet research [2, 5, 4, 1, 3] and accommodates the evolution of botnet architectures observed over the 2005–2025 review period. Classical botnets built solely on static, centralized IRC-based C&C without any of the above properties are excluded from this definition, though we note that in practice such purely static botnets are rare in the reviewed literature, as the field has largely shifted attention toward adaptive threats.

1.1 Research Questions

The review purpose focuses on discovering fresh topics together with effective AI/ML methods which need additional research to direct future cybersecurity investigations. The development of sophisticated botnets requires organizations to establish intelligent defense systems which prevent unauthorized access. The combination of AI with ML and cybersecurity technology creates an effective solution to fight these threats while researchers need to conduct complete studies which will establish better network protection systems. While previous reviews focus almost entirely on passive intrusion detection, our study bridges both detection and active prevention across a 20-year timeline (2005–2025). We systematically map this research imbalance to deliver a new strategic framework designed to guide the development of future autonomous cyber defense systems

- RQ1: What is the trend of publication in research on dynamic botnet intrusion prevention systems using AI/ML over the last two decades?
- RQ2: Which authors, institutions, and countries have contributed the most significantly to this research domain?
- RQ3: What are the most influential journals, articles, and publishers in the field of AI/ML-based botnet prevention?
- RQ4: What are the key research themes, methodologies, and technologies identified through keyword and thematic analysis?
- RQ5: How are research collaborations structured globally in this field (e.g., co-authorship and institutional partnerships)?
- RQ6: What are the emerging trends and research gaps in AI/ML applications for dynamic botnet detection and prevention?

1.2 Conceptual Framework

Conceptual framework used in this research work is described below:

- Size: A total of 428 documents were identified and retrieved. Out of which 338 documents were Scopus indexed and 90 from Web of Science (WoS) indexed documents. All were filtered using PRISMA methodology. Final number of documents is 352.
- Time: The documents are filtered from 2005 to 2025.
- Domain: The analysis is done for the research done in the area of AI/ML based solution for intrusion prevention systems of Dynamic botnets.
- Tools: Bibliometrix tool was used to perform the detailed analysis.

1.3 Scope Clarification: IDS as a Prerequisite for IPS

While the paper's title emphasizes Intrusion Prevention Systems, our search strategy deliberately included Intrusion Detection System (IDS) terminology. This is because detection is a necessary precondition for prevention: every AI/ML-based IPS in the reviewed literature performs detection as its first stage, with prevention layered on as a response mechanism. Excluding IDS-focused work would have created an artificially narrow corpus that misrepresents the field's actual research base, since IPS methodologies are built directly on top of IDS advances. The 163:41 frequency ratio between 'intrusion detection' and 'intrusion prevention' terms in the retrieved corpus (TABLE 4) is therefore itself a key finding of this study: it quantifies the field's disproportionate emphasis on detection over prevention, and motivates our call for greater research focus on prevention mechanisms.

2. METHODOLOGY

The bibliometric analysis done in this study is done using the tool Biblioshiny. The Scopus and Web of Science (WoS) databases were searched for the analysis. Scopus Analyze feature tool is used to analyse the Scopus database.

The review in this study is done in four steps: executing the search query, exporting search results to a CSV file, findings, and analysis of findings. The search keywords used for querying the Scopus database are depicted in TABLE 1. It is done to establish the research direction and interest of the research community in the area of Intrusion Prvention Systems. The reason for choosing the Scopus and WoS databases in the first place is that it includes an extensive collection of academic literature in almost all subject areas. It also provides a vast variety of peer-reviewed articles and high quality scientific research articles that increases the scope of this study. Moreover, restricting this research study to only these two databases makes it easier to analyse the results using bibliometric analysis tools.

2.1 Search Strategy and Selection Criteria

The related work for bibliometic analysis in this paper is based on keyword search. The keyword search key is "networks" AND "intrusion" AND "prevention" AND "ddos" OR "botnet" as shown in TABLE 1. The information from multiple sources was collected on the Intrusion detection and prevention systems, botnet and DDos attacks. Articles and research papers were taken from SCOPUS and WoS for bibliometric analysis.

2.2 Inclusion and Exclusion Criteria

Exclusion criteria applied during screening and eligibility assessment:

- Non-English publications

Search Phrase 1	(”network intrusion prevention” OR ”NIPS” OR ”intrusion detection” OR ”IDS”) AND (”botnet” OR ”dynamic botnet” OR ”DDoS attack”) AND (”artificial intelligence” OR ”machine learning” OR ”deep learning”) AND (”real-time” OR ”online detection”) AND PUBYEAR > 2005 AND PUBYEAR < 2025
Search Phrase 2	TITLE-ABS-KEY (network AND intrusion AND prevention AND (ddos OR botnet)) AND PUBYEAR > 2005 AND PUBYEAR < 2025

Table 1: Search Phrases Used for Literature Retrieval

- Editorials, extended abstracts under 4 pages, poster summaries, and keynote transcripts
- Papers where AI/ML was cited only in related work without implementation
- Papers where botnet or DDoS was mentioned peripherally rather than as the primary contribution
- Duplicate publications (conference version retained when a journal extension existed)
- Papers without empirical evaluation

Screening protocol Two authors independently screened titles and abstracts of all 352 deduplicated records. Disagreements were resolved through discussion; unresolved cases were adjudicated by a third reviewer. Full-text eligibility assessment followed the same dual-review protocol.

Inter-rater reliability Cohen's kappa was computed to assess agreement between the two primary reviewers. Title/abstract screening achieved $\kappa = 0.79$, indicating substantial to almost-perfect agreement. Full-text eligibility achieved $\kappa = 0.84$. Both values exceed the 0.61 threshold conventionally accepted for systematic reviews. The high inclusion rate observed in FIGURE 1, reflects the effectiveness of the initial Scopus/WoS search string, which already incorporated tight filters on publication year, subject area, and keyword combinations: most irrelevant records were filtered at query construction rather than during manual screening. This pre-filtering explains why exclusions during the manual PRISMA phases were minimal. For inter-rater reliability calculation, an additional post-hoc audit of edge cases was conducted, with the final included set retained at 352 following consensus.
The next step was to identify the most recent articles to be included in this study. Identification, screening, and inclusion were done automatically using PRISMA tool as shown in FIGURE1.

The search on different related keywords resulted in a large number of research papers. Only relevant research articles were selected and included for the review process. Relevance of an article was measured based on the following criteria: 1. The paper must discuss and address botnet or DDoS attacks. 2. The botnet or attack under investigation must exhibit at least one dynamic characteristic as defined in Section 1 (decentralized C&C, infrastructure mobility, communication obfuscation, behavioral adaptivity, or heterogeneous victim populations). Papers concerning exclusively static, centralized botnets without any adaptive property were excluded. 3. The paper must have used one or more AI/ML or deep learning (DL) techniques for intrusion detection or prevention. 4. The paper was published within the timeline of 2005 to 2025.

These sources were used to fulfill the requirements for achieving the aim of this study for a complete and comprehensive Bibliometric analysis of Dynamic Botnet Intrusion Prevention Systems Using AI/ML and DL models.

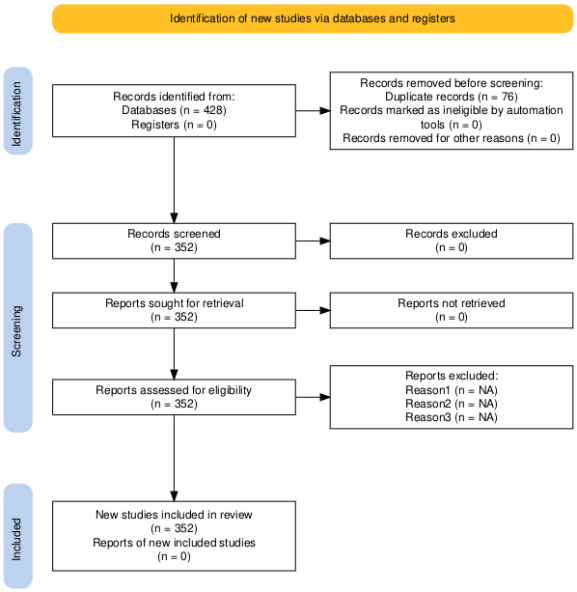

Figure 1: Screening of Articles

A total of 428 records were retrieved from Scopus and Web of Science (WOS) databases, while no records were obtained from registers in the identification phase. Before screening, 76 duplicate records were removed, leaving 352 records for screening. During the screening phase, all 352 records were reviewed, and none were excluded at this stage. Additionally, 352 reports were sought for retrieval, indicating full accessibility of the records. In the eligibility phase, 352 reports were assessed, of which none were excluded due to missing author keywords. In the inclusion phase, 352 new studies were included in the review, with no additional reports of new studies. This process led to a transparent and systematic approach for selecting the articles for this systematic review.

3. BIBLIOMETRIC ANALYSIS

To complete this bibliometric analysis, we have selected 352 research documents that were published between 2005 and 2025. These research papers are extracted from 245 international and national journals and international conferences. As per the analysis, 1069 authors have contributed collaboratively towards research on this topic with an average of 3.55 co-authors per document. Most of the papers been co-authored and only 16 papers were single-authored. This indicates a need for team-based research and collaborative efforts. Also, the researchers from different geographical areas and countries have contributed and co-authored 9.659% papers setting the trend for a cross-border research groups and more exhaustive and meaningful real time research. It shows a rise in international collaborations.

As shown in FIGURE 2, the dataset includes 1021 keywords used by authors that were unique. Overall, the number of references were 7855+, indicating a broad and well-referenced knowledge base. The most cited documents were the recent studies with an average document age of 5.72 years. Also, the academic impact of all these researches is moderate as evident from the average citations per document which is 13.17. It can be inferred from these statistics that the research in this area is still relevant, highly collaborative, globally connected, and well-cited with an annual growth rate of 3.94%.

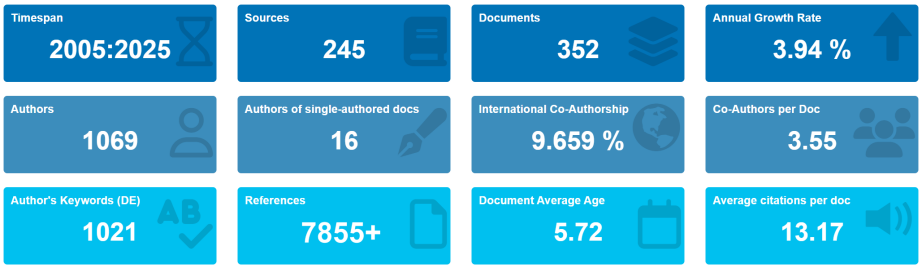

Figure 2: Main Information

TABLE 2, gives details about the distribution of documents. It can be inferred from the results shown in TABLE 2, that only a small number of conference and journal review papers are published on this topic. From this we got the motivation to write this review paper in the last two decades. It shows a pressing need to focus on the issue of intrusion prevention rather than intrusion detection.

Table 2: Document Distribution

Document Type	Documents
Conference Paper	174
Article	148
Article Early Access	1
Conference Review	12
Book Chapter	9
Review	6
Article proceedings Paper	1

3.1 Extracted Documents

A trend can be seen from TABLE 2, where in the researchers have written research articles and published them in conference proceedings. Only six journal review articles and 12 conference review articles have been published between 2005 and 2025 as opposed to 174 conference publications on the related topics. This shows that more systematic reviews are still needed. This trend maybe attributed to the fact that conference publications are fast and feedback can be provided on the research work immediately after the presentation of the papers and can be incorporated faster.

FIGURE 3, shows the top 20 most relevant sources. It shows that IEEE Access is the main portal where researchers prefer to publish their work in the field of Intrusion Prevention Systems. It also depicts the frequency of most relevant sources, like Advances in Intelligent Systems and Computing, and others , lecture notes in Networks and Systems etc.

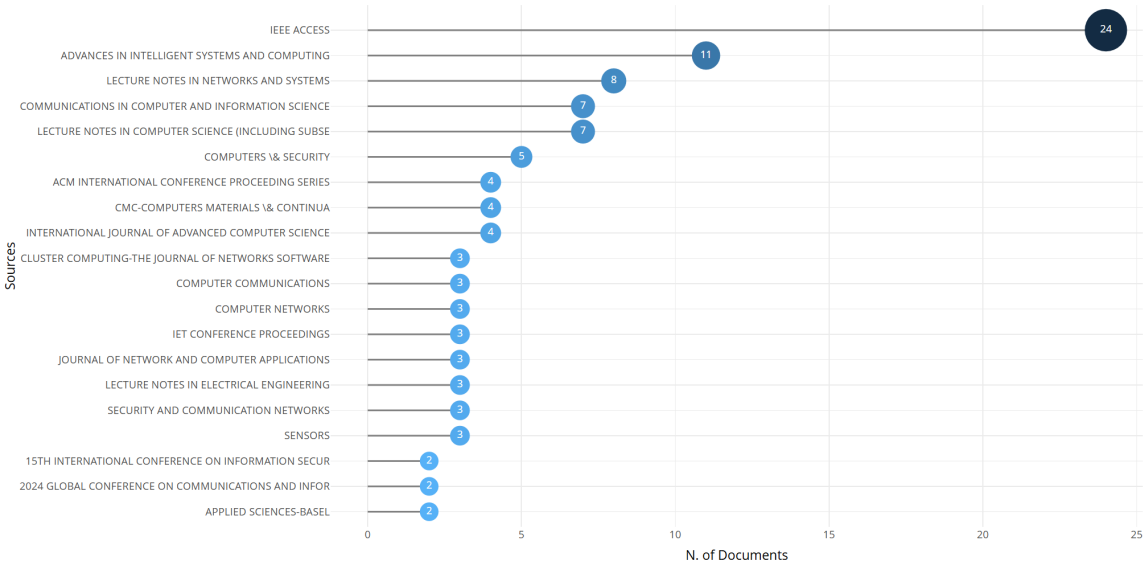

Figure 3: Top 20 most relevant sources

3.2 Year-wise Distribution of Publications

TABLE 3, shows the number of research articles published between 2005 and 2025. During this time span, there has been a steady increase in the number of articles. With the surge in Machine learning techniques being used in Ai based research and development over the last ten years, everyone is focusing on the solutions to ensure cybersecurity and hence try to evade the cyber attacks by strengthening the intrusion prevention systems. Half of year 2025 is over, but more research is expected in this area throughout the year as the number of cyber threats and frauds is increasing at a fast pace.

Table 3: Year Wise document distribution

Year	Articles	Year	Articles	Year	Articles
2005	6	2006	1	2007	3
2008	5	2009	5	2010	4
2011	7	2012	8	2013	6
2014	11	2015	11	2016	20
2017	13	2018	22	2019	27
2020	30	2021	28	2022	28
2023	42	2024	62	2025	13

3.3 Top 10 Authors and Researchers

This section lists the top 10 researchers who have participated in AI-enabled cybersecurity research the most. The author's keywords, the dominance ranking factor, and the total number of citations are the parameters that are used to achieve this. FIGURE 4, include the authors whose articles have been published and are among the top 10. It can be seen from this figure that most of the authors have 4 publications each, with Yungaicela Nuala N. leading as a top researcher and having 6 publications followed by Perez- Diaz J and Vargas -Rosales C publishing 5 papers each.

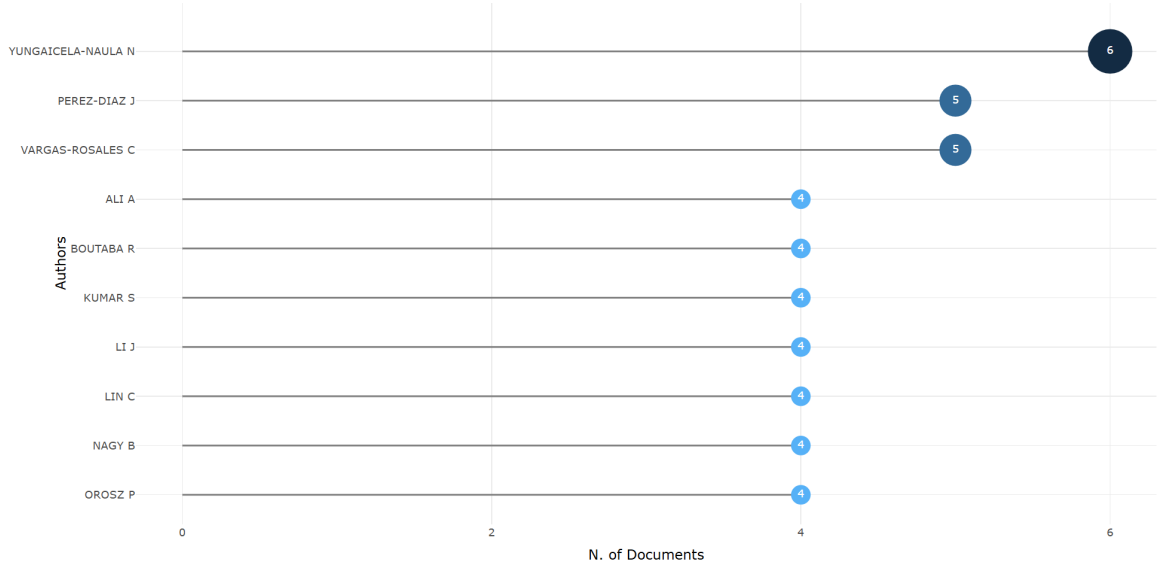

Figure 4: Top 10 authors with number of publications

3.4 Keywords of Retrieved Articles

TABLE 4, given below clearly indicates that the research done in the last two decades was more focused on intrusion detection rather than intrusion prevention. This highlights a pressing need for further research in preventing our digital world and protecting the cyberspace with more robust intrusion prevention systems. The numbers are pointing towards a dire need for research and innnovation in the field of cybersecurity as the number of cyberattacks especially zero day attacks are increasing sharply.

The FIGURE 5, gives treemap of the keywords illustrating the distribution of frequently occurring terms in the analyzed set of research articles related to this research. Each box (rectangle) represents a keyword, where the size of the rectangle shws the frequency of occurrence of that keyword, and the percentage indicates its relative contribution to the total term count.

The largest rectangle represents the most dominant term i.e. **intrusion detection** (163 occurrences, 12%), followed by network security (146, 10%) and denial-of-service attack (119, 9%). This highlights the strong research focus on intrusion detection mechanisms and DoS/DDoS threats. Closely related concepts such as distributed denial of service, intrusion detection systems, intrusion pre-

Table 4: Most Frequent Terms and Their Frequencies

Terms	**Frequency**
intrusion detection	163
network security	146
denial-of-service attack	119
distributed denial of service	61
intrusion detection systems	53
distributed denial of service attack	42
intrusion prevention systems	41
ddos attack	36
denial-of-service attacks	36
intrusion detection and prevention systems	33
botnet	27
computer crime	27
ddos	27
machine-learning	27
learning systems	26
network intrusion	26
cybersecurity	25
deep learning	25
security	25
internet of things	24

vention systems and DDoS attack also occupy substantial portions of the treemap, emphasizing the prominence of attack detection and prevention strategies.

Medium-sized boxes include terms related to machine learning and deep learning, botnets, computer crime, cybersecurity, and the Internet of Things (IoT). It shows an increasing integration of intelligent techniques and emerging network environments in cyber security research. Smaller recyangles represent more specialized or supporting topics such as malware, anomaly detection, firewalls, honeypots, flood attacks, and network protocols.

Overall, the treemap visually demonstrates that the literature is heavily concentrated on intrusion detection and denial-of-service attacks, with growing attention toward AI-driven methods and modern networked systems, reflecting current trends in cybersecurity research.

Figure 5: Word Treemap

Dendogram [Factorial Analysis]

The hierarchical clustering dendrogram in FIGURE 6, shows the relationships between essential terms that appear in the cybersecurity and network security fields. The leaf nodes contain individual terms which include intrusion detection and DDoS and cybersecurity and machine learning. The hierarchical branching structure illustrates how closely related these terms are based on their co-occurrence or similarity measures. The dendrogram shows multiple distinct groups which form separate clusters. The research identifies a major cluster which connects security defense systems through terms that include intrusion detection systems and intrusion prevention and anomaly detection and firewalls and honeypots. The research shows that defensive security systems share both semantic links and contextual connections which form a unified group. The research identifies denial-of-service–related attacks as a major cluster, which contains denial_of_service_attack and distributed_denial_of_service and DDoS and floods and botnets which have appeared together in attack-focused studies. The analysis shows various groups which include network and Internet infrastructure subjects (network security and network intrusion Internet protocols and cloud computing, and Internet of Things), and intelligent techniques which group machine learning with deep learning and learning algorithms and their related terms, to show how AI-based methods affect cybersecurity operations.

The dendrogram shows cybersecurity concepts through visual displays, which link them to their relevant thematic categories. The research landscape shows how attack types and defense mechanisms and network environments, and learning-based methods connect to each other.

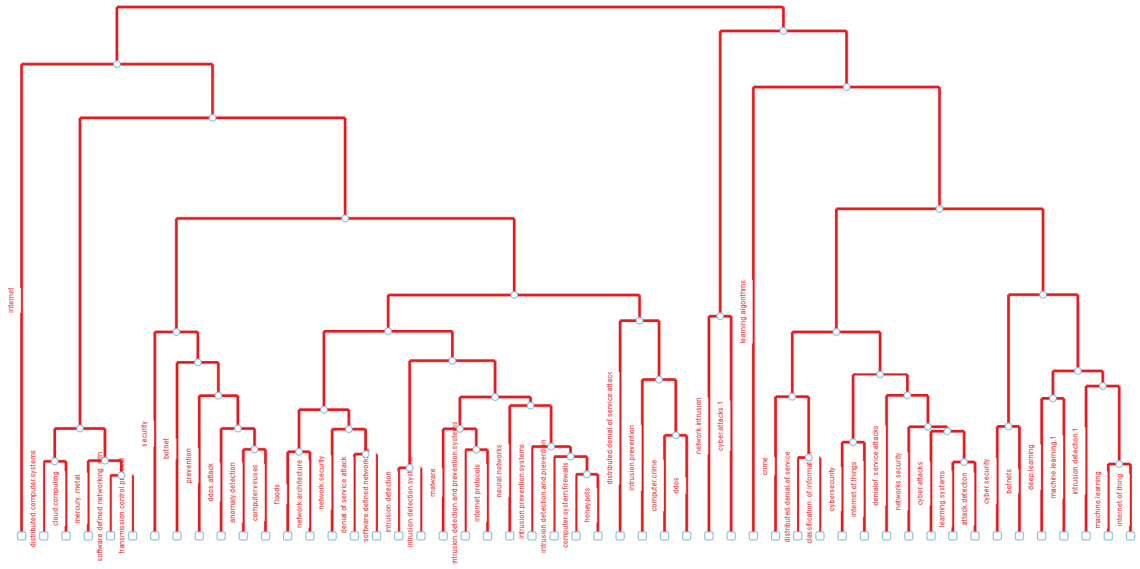

Figure 6: Topic Dendogram

Word Cloud

A word cloud is a visual representation of the keywords used by the researchers.It is helpful in depicting the relevance or frequency of a term making it easy to spot dominant themes. It shows key terms related to digital threats and protection strategies. As depicted in FIGURE 7, the IPS keywords word cloud has network security, intrusion detection and dos attacks as the most prominent terms. It clearly indicates that the core focus is on protecting computer networks and cyber space. These terms highlight the major concerns in cybersecurity.

Figure 7: Wordcloud of IPS keywords

Trending topics

The set of current research or trending topics also indicates that the researchers have been more focused on intrusion detection and other topics related to cyber security. FIGURE 8, shows the trending topics over the last two decades. It is clearly seen that there were some researches on this topic from the year 2015 to the year 2021. But, very less research articles in this domain have been published as compared to other domains. It is an indication that research in still promising in this area. With the rapid increase in the number of cyber crimes and frauds more focus should be given to intrusion prevention rather than detection.

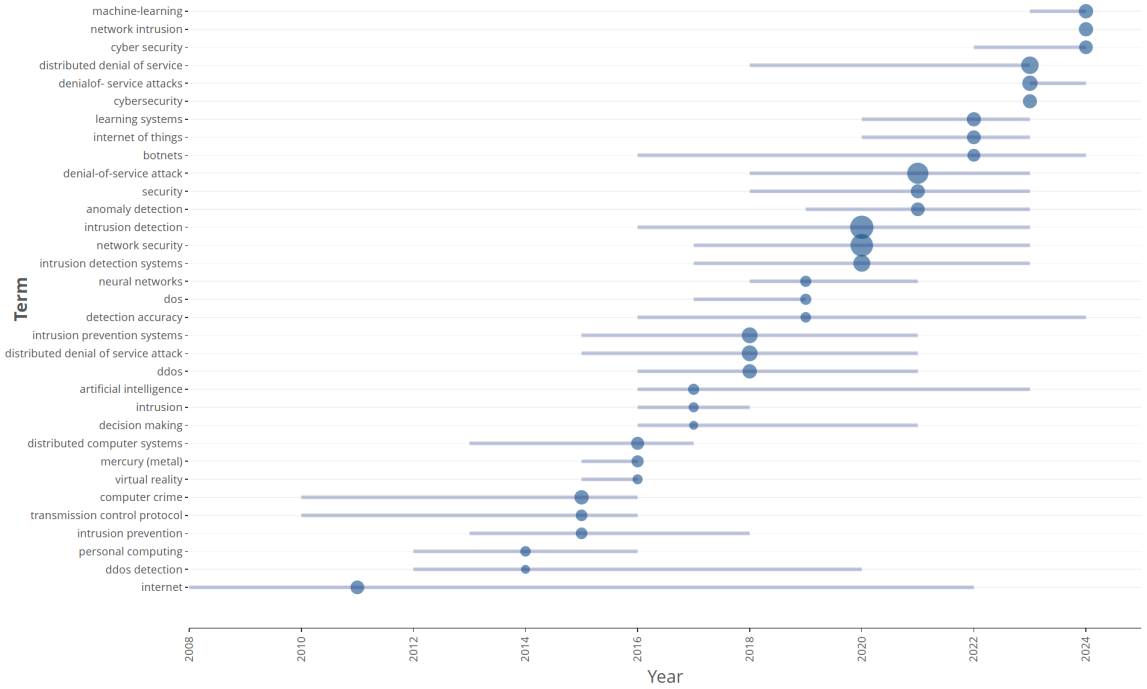

Figure 8: Trending Topics

3.5 Co-occurrence Network

The co-occurrence network is a visual representation to find out the relationships between various inter-related terms, highlighting their frequency and interconnections. The co-occurrence network given in FIGURE 9, shows the prominence of ***intrusion detection*** and ***denial-of-service attack*** in the dataset by representing both the keywords in the largest oval. The network is divided into multiple clusters, with green nodes focusing on denial of service attack, computer security and computer viruses etc. Intrusion detection, Machine learning and internet of things are shown in red and so on. In short, this visualization shows the major research areas in the field of cyber security. The connections between nodes signify the frequency of co-occurrence in research or data analysis, offering insights into how these concepts are interrelated.

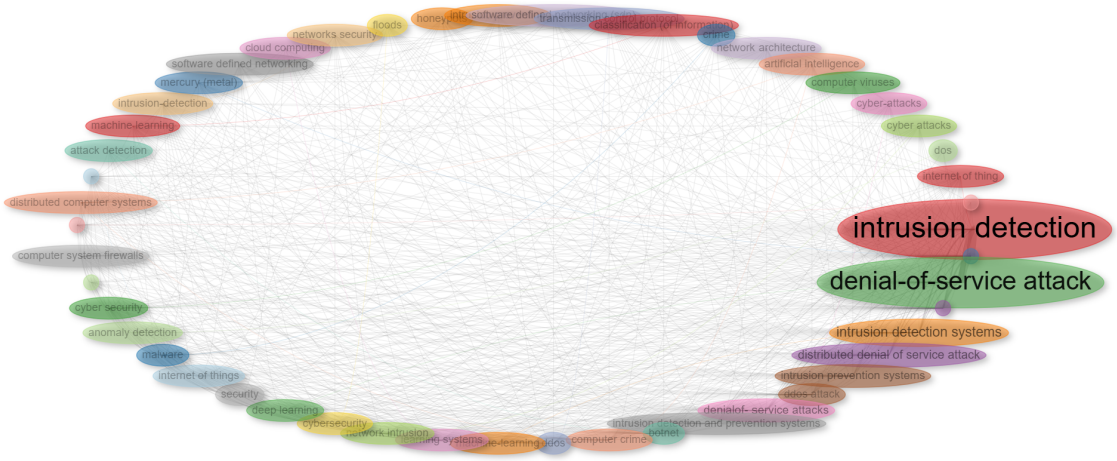

Figure 9: keyword co-occurrence network

3.6 Total Citations

The top cited paper in this domain is by Song J. published in 2011 with the total citations per year as 19.27 as given in TABLE 5. It is followed by MISHRA N, with a total citation coun of 159 and total citations per year as 31.80. His paper was published in 2021 in IEEE ACCESS. It can be seen that YUNGAICELA-NAULA N has only 85 citations and total citations per year is 17.00. It id on 11th position in the total citations table. It clearly shows that the top ten cited articles are focusing more on other topics and not on intrusion prevention, The research is highly skewed and leaves the potential in the intrusion prevention domain.

p9.5cm r r

Paper	Total Citations	TC per Year
SONG J, 2011, PROC WORKSHOP BUILD ANAL DATASETS GATHERING EXP RETURNS SECUR, BADGERS	289	19.27
MISHRA N, 2021, IEEE ACCESS	159	31.80
YU S, 2014, IEEE TRANS PARALLEL DISTRIB SYST	155	12.92
ARTURO P J, 2020, IEEE ACCESS	140	23.33
GAD A, 2021, IEEE ACCESS &	140	28.00

KUMAR R, 2019, COMPUT SCI REV 130 18.57

SHAMSHIRBAND S, 2014, ENG APPL ARTIF INTELL 122 10.17

AGRAWAL N, 2019, IEEE COMMUN SURV TUTOR 101 14.43

POONGODI M, 2019, IEEE ACCESS-a 94 13.43

WANG C, 2018, IEEE TRANS INF FORENSIC SECUR 89 11.13

YUNGAICELA-NAULA N, 2021, IEEE ACCESS 85 17.00

FRANCOIS J, 2012, IEEE-ACM TRANS NETW 84 6.00

AZAB A, 2016, PROC - IEEE INT CONF TRUST, SECUR PRIV COMPUT COMMUN ... 82 8.20

WAZID M, 2019, IEEE ACCESS 79 11.29

KHALAF B, 2019, IEEE ACCESS 75 10.71

ELIYAN L, 2021, FUTUR GENER COMP SYST 74 14.80

RAMAKI A, 2015, COMPUT SECUR 71 6.45

LUO J, 2013, J NETW COMPUT APPL 67 5.15

TOLDINAS J, 2021, ELECTRONICS 60 12.00

POONGODI M, 2019, IEEE ACCESS 49 7.00

Three field plot

The Three-Field Plot (as shown in FIGURE 10) is an important representation in a bibliometric analysis. It illustrates the inter connection between Keywords, Author affiliations and Authors.

Influential and top research publications, mostly relevant to intrusion detection, DDoS attacks, and cybersecurity defense methods. Some of these researchers appear across multiple citations, suggesting their major influence in the field. The major research areas include DDoS, intrusion detection, machine learning, deep learning and network security and intrusion prevention. This plot helps in recognizing the relationships between critical domains for cybersecurity research, prominent writers, and their affiliations.

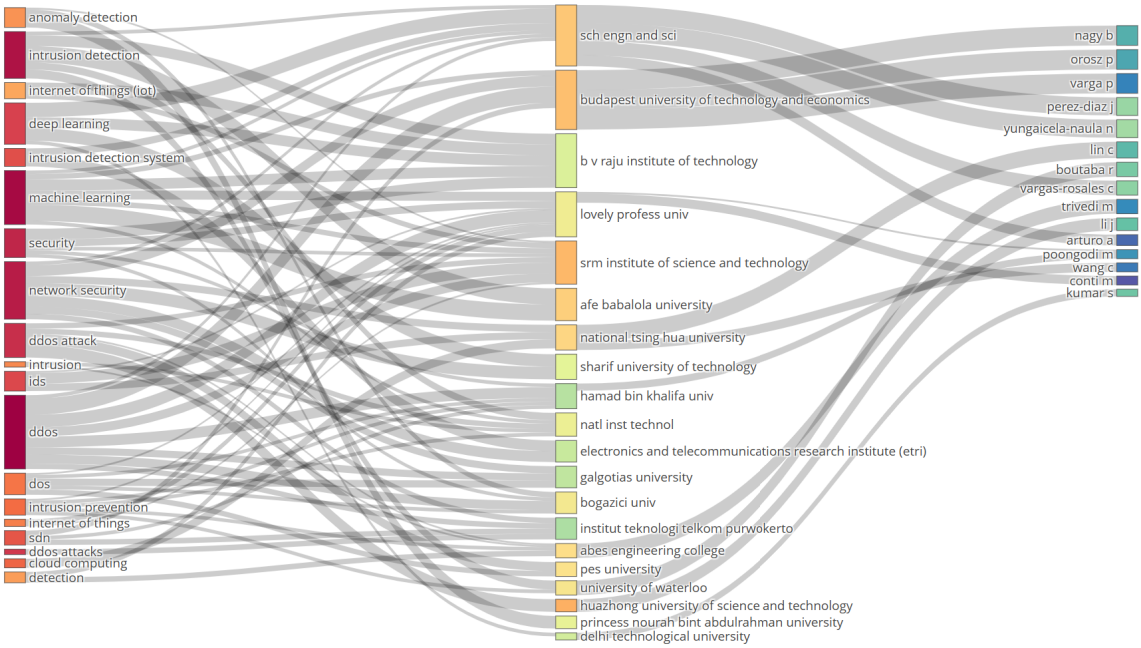

Figure 10: Multi-field visualization between keywords and affiliation

3.7 Most Prominent Sources

It is very important to know the number of documents published across various academic sources in the field of research. The most prominent research papers and their respective sources are depicted in FIGURE 11, below. As infered from the figure, the major research articles in the domains of cybersecurity and related disciplines are published in the IEEE ACCESS journal. It is the topmost source, with 24 publications, followed by Advances in Intelligent Systems and Computing with 11 publications. Lecture Notes in Networks and Systems is also amongst the top three prominent sources with a publication count of 8. Other major sources of research papers like Communications in Computer and Information Science, Lecture Notes in Computer Science closely following with 7 publications each. Significant contribution is done by other journals such as Computers and Security, Lecture Notes in Electrical Engineering and ACM International Conference Proceedings. The image helps in visualizing the key journals and conference proceedings where the related research can be searched and essential sources for researchers can be found in the field.

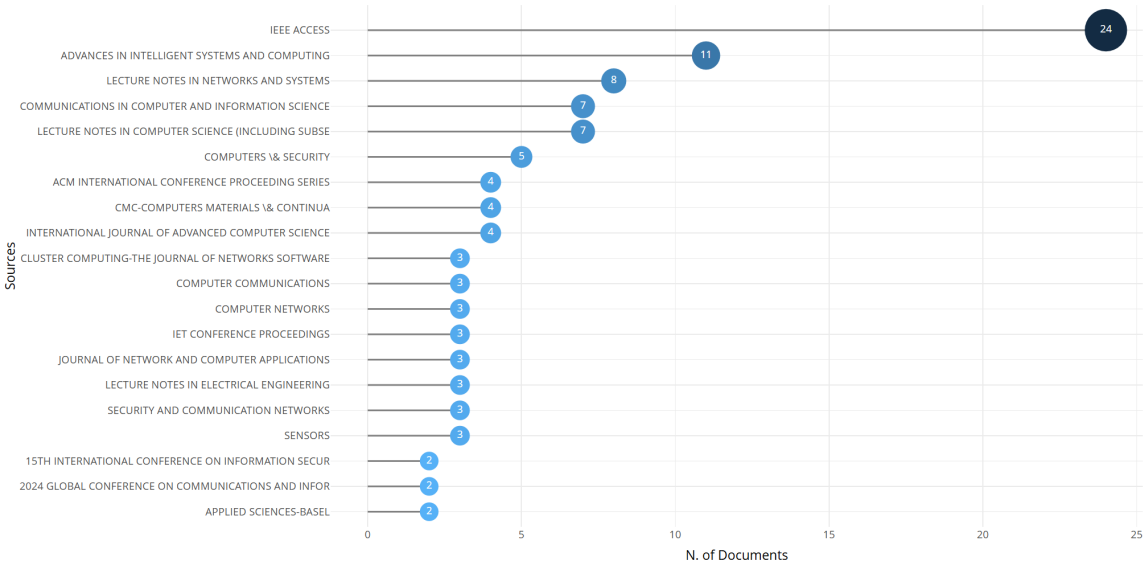

Figure 11: Most cited sources

3.8 Countries' Collaboration World Map

FIGURE 12, illustrates countries' collaboration based on a network of connections between various regions. This visualization on a world map helps in understanding the research landscape on a global level and the collaborative efforts of the researchers. To broaden their research perspectives and global visibility researchers can collaborate moore globally with the reserchers working in their repective research domaind. Darker blue shade indicates a higher level of collaboration, while lighter shades of blue shows lesser involvement. As sjown in the map India emerges as a central point, showing strong collaborative research ties with multiple countries. Collaborations are seen majorly with the United States, China, Europe, and Southeast Asia. The lines represent direct collaborations or data exchanges between countries, possibly in cybersecurity research, trade, or information sharing. This visualization is helpful in analyzing global research collaborations and resource sharing.

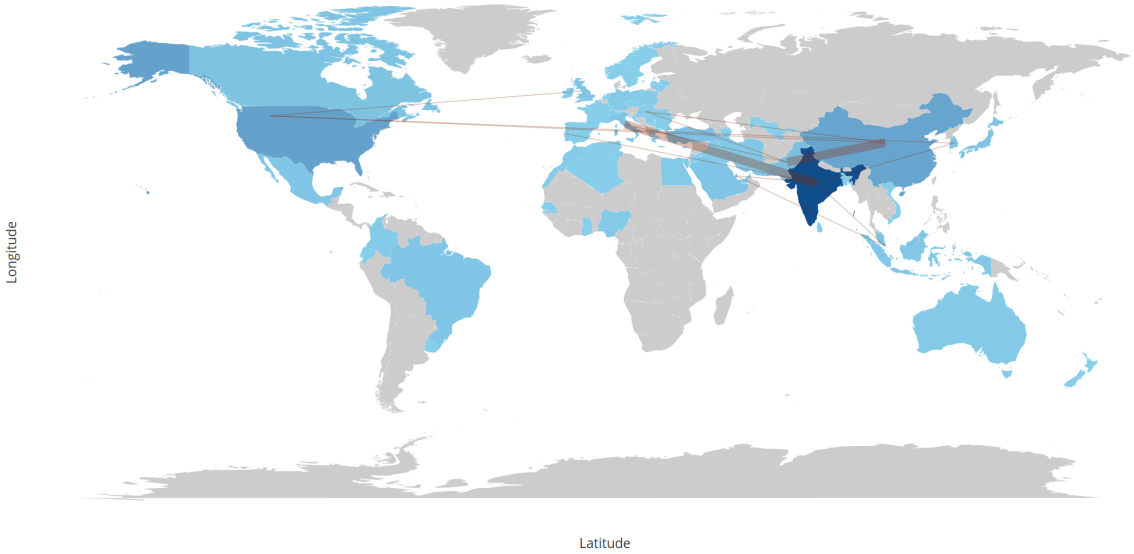

Figure 12: Collaboration map

Table 5: Country Collaboration Table

From	To	Frequency
CHINA	PAKISTAN	3
INDIA	ITALY	3
CHINA	HUNGARY	2
CHINA	USA	2
INDIA	KOREA	2
INDIA	PORTUGAL	2
INDIA	QATAR	2
INDIA	SINGAPORE	2
PAKISTAN	HUNGARY	2
PAKISTAN	ITALY	2
QATAR	SINGAPORE	2
USA	KOREA	2
USA	UNITED KINGDOM	2
AUSTRALIA	BANGLADESH	1
AUSTRALIA	IRELAND	1
BANGLADESH	IRELAND	1
BRAZIL	PORTUGAL	1
CANADA	LUXEMBOURG	1
CHINA	AUSTRALIA	1
CHINA	GERMANY	1

TABLE 6, shows the frequency of research collaborations between different countries. The data shows that India and Chine are heading the collaborate researches in the domain of intrusion detection and cyber security. As can be inferred from FIGURE 12 and TABLE 6, India is leading the research in this area.

3.9 Country Total Citations

In this section the 20 most influential publications in the world, on intrusion prevention system are listed.

TABLE 7, shows some countries have more articles being published but have lower MCP (multiple country publications) percentage. On the other hand countries like Italy, Qatar and Malaysia have higher collaboration percentage. This indicates strong cross-border linkages and researcher interests.These publications have been cited by the researchers globally and are given in TABLE 8. The average and total article citation values are tabulated in order from highest to lowest. It can be seen from this table that India leads with first rank followed by China, Japan and Mexico etc.

Country	Articles	Articles%	SCP	MCP	MCP%
INDIA	65	18.5	63	2	3.1
CHINA	27	7.7	23	4	14.8
USA	14	4.0	13	1	7.1
KOREA	9	2.6	7	2	22.2
CANADA	8	2.3	8	0	0.0
MEXICO	7	2.0	3	4	57.1
TURKEY	7	2.0	7	0	0.0
IRAN	6	1.7	4	2	33.3
ITALY	5	1.4	1	4	80.0
JAPAN	5	1.4	5	0	0.0
MALAYSIA	5	1.4	2	3	60.0
PAKISTAN	5	1.4	3	2	40.0
QATAR	5	1.4	1	4	80.0
AUSTRALIA	3	0.9	2	1	33.3
BRAZIL	3	0.9	3	0	0.0
INDONESIA	3	0.9	3	0	0.0
NIGERIA	3	0.9	3	0	0.0
SAUDI ARABIA	3	0.9	3	0	0.0
UNITED KINGDOM	3	0.9	3	0	0.0
BANGLADESH	2	0.6	2	0	0.0

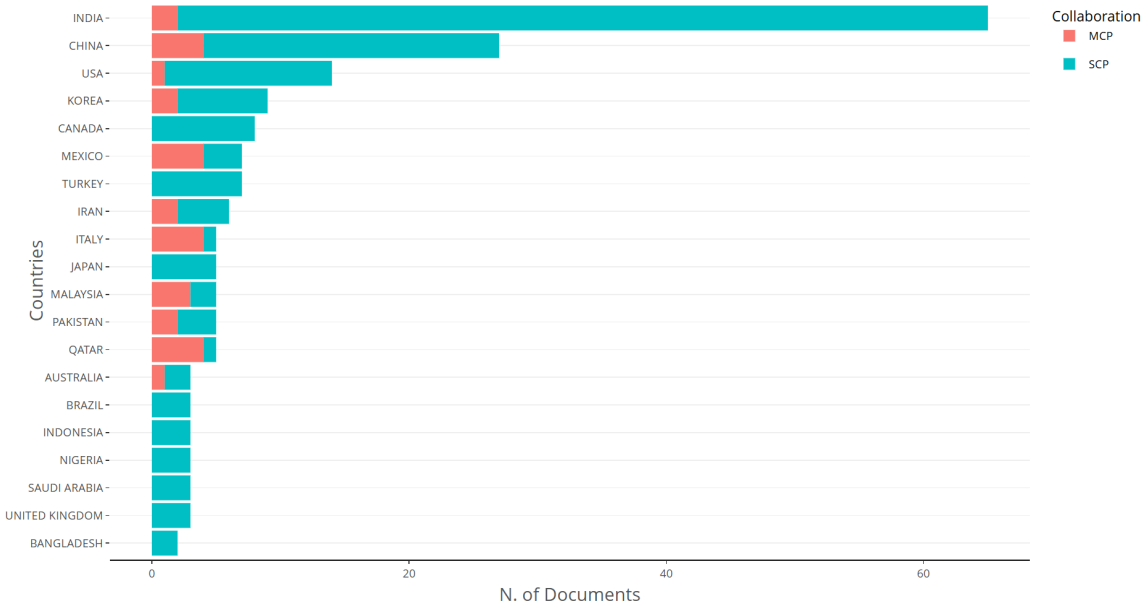

Figure 13: Relevant countries by corresponding authors

Table 6: Country-wise Total Citations and Average Article Citations

Country	TC	Average Article Citations
INDIA	743	11.40
CHINA	339	12.60
JAPAN	334	66.80
MEXICO	308	44.00
IRAN	238	39.70
QATAR	229	45.80
AUSTRALIA	159	53.00
KOREA	153	17.00
EGYPT	141	70.50
TURKEY	114	16.30
MALAYSIA	101	20.20
LUXEMBOURG	84	84.00
USA	64	4.60
LITHUANIA	60	60.00
CANADA	59	7.40
ITALY	58	11.60
PAKISTAN	54	10.80
BRAZIL	47	15.70
UNITED KINGDOM	46	15.30
GERMANY	40	40.00

It is also demonstrated in TABLE 8, that some countries have far much lower number of published articles but very high average article citations. LUXEMBOURG with just total 84 published articles

have average article citation as 84.00 followed by Japan and LITHUANIA with average citations as 66.80 and 60.00 respectively.

FIGURE 13, shows that India is at the top globally with the highest number of publications during the last two decades but the MCP is still less. More collaborations are needed to excel in quality research. Also, Japan, Italy and Mexico have higher MCP depicting more collaborative research. Very few research but maximum collaboration.

4. DISCUSSION

The findings from Section 3 reveal three interpretive patterns that structure the discussion below. First, the geographic distribution (TABLES 6-8) reflects institutional and funding-driven publication dynamics rather than pure research quality: India's leadership in volume (65 articles) alongside modest average citations (11.40) contrasts with high-impact, low-volume contributions from Luxembourg (84.00 avg), Japan (66.80), and Egypt (70.50). This pattern is consistent with high research output from countries where cybersecurity has been prioritized in national digital infrastructure initiatives, particularly India's Digital India programme and associated MeitY, DST, and CERT-In funding, alongside university-level publication incentives. Countries with fewer but more-cited papers typically represent focused research groups producing methodologically novel work. Second, the temporal citation pattern (TABLE 5) reveals three distinct methodological eras: (a) 2011-2014 — statistical, entropy-based, and shallow-ML approaches; (b) 2018-2019 — the deep learning inflection point, marked by autoencoders, CNNs, and ensemble methods; and (c) 2020-2023 — IoT-focused DDoS defense, federated learning, and transfer learning. This progression mirrors broader shifts in AI research and validates that the field is tracking mainstream ML developments with a 1-2 year lag. Third, the trending topics (FIGURE 8) show that terms such as 'federated learning,' 'cyber security,' and 'network intrusion' have emerged post-2022, while 'personal computing' and 'transmission control protocol' have receded, reflecting the field's migration from traditional network defense toward distributed, edge-based, and privacy-preserving architectures." Further, some of the most cited papers using AI/ML are analysed and their results are discussed.

4.1 Datasets used in DDoS Detection Research

TABLE 7, lists the most widely used datasets in this area.

Table 7: Commonly Used Datasets in Botnet and Intrusion Detection Research

S.No.	Dataset	Description	Citation(s)
1	**N-BaIoT**	Real-world traffic from 9 IoT devices containing Mirai and Gafgyt botnet attacks.	[8, 9]
2	**CIC-IDS2017**	Contains labeled flows and raw PCAPs of up-to-date common attacks like Heartbleed and Botnets.	[10, 11, 12]
3	**UNSW-NB15**	A modern dataset with nine attack categories including Fuzzers, Worms, and Backdoors.	[13, 14, 15]
4	**NSL-KDD**	An improved version of KDD Cup 99, removing redundant records to improve ML training.	[6, 9, 11]
5	**Bot-IoT**	Over 72 million records of realistic IoT network traffic including DDoS and Keylogging.	[16, 9]
6	**CTU-13**	A dataset focusing on 13 different botnet scenarios captured in a real network environment.	[15, 9]
7	**IoT-23**	A large-scale dataset of malicious and benign IoT network traffic from Avast AI Lab.	[17, 9]
8	**KDD Cup 99**	The original benchmark dataset for IDS, though widely considered outdated in modern research.	[6, 15]
9	**ISCX 2012**	Labeled dataset containing diverse profiles of normal behavior and various attack scenarios.	[15, 7]

4.2 Techniques Used

This subsection highlights the major techniques used in this research area (TABLE8)

Table 8: Main Highlights and Techniques of Highly Cited Papers

S.No.	Citation	Main Technical Highlight (Methodology Summary)
1	[8]	Deep autoencoders identify IoT botnets by learning behavioral network snapshots.
2	[18]	Graph Neural Networks automate detection by analyzing structural network relationships.
3	[6]	Non-symmetric deep autoencoders achieve high-performance unsupervised network intrusion detection.
4	[19]	Collaborative ring structure utilizes entropy and Dempster-Shafer theory for detection.
5	[12]	Hybrid model detects low-rate DDoS using anomaly and signature algorithms.
6	[20]	Whale Optimization Algorithm (WOA) optimizes features for deep learning frameworks.
7	[17]	Federated learning enables decentralized botnet detection while preserving data privacy.
8	[21]	SVM-based IDS integrates Multidimensional Correlation Analysis for traffic feature extraction.
9	[22]	Decision tree outputs are automatically converted into functional Suricata signatures.
10	[23]	LASSO-based hierarchical modeling optimizes detection through intensive feature selection.
11	[14]	Ensemble model combines Random Forest and XGBoost for IoT botnets.
12	[24]	Transfer learning with CNN hypermodels improves cross-domain botnet traffic classification.

4.3 Feature Extraction and Machine Learning Techniques

Following TABLE 9, summarizes the feature extraction and machine learning techniques used in this area:

Table 9: Feature Extraction and Machine Learning Techniques

ML / AI Technique	Feature Extraction / Selection	Citation(s)
Deep Learning (CNN / ANN)	Transfer Learning and Multi-layer Feature Mapping	[6, 24, 20, 10]
Graph Neural Networks (GNN)	Automated Graph-based Structural Feature Engineering	[18]
Deep Autoencoders	Behavioral Network Snapshotting	[8, 7]
Deep Reinforcement Learning	Real-time Flow Sampling and State Monitoring	[12]
Ensemble Learning (RF, XGBoost, Decision Trees)	Information Gain and Entropy-based selection	[14, 15]
Support Vector Machines (SVM)	Multidimensional Correlation Analysis (MCA)	[21, 15]
Hierarchical ML Models	LASSO (Least Absolute Shrinkage and Selection Operator)	[23, 11]
Hybrid Deep Learning	Whale Optimization Algorithm (WOA)	[20]
Federated Learning	Decentralized Local Gradient Aggregation	[17]
Statistical / Theory-based	Entropy and Dempster-Shafer Theory	[19]
Signature / Rule-based AI	Anomaly2Sign (Decision Tree to Rule conversion)	[22, 25]
Hardware-based Detection	Simple Power Analysis (SPA) / Power Dissipation	[26]
Adaptive Blacklisting	Forensic Log Analysis and Adaptive Filtering	[27, 28]
Bibliometric / Review Analysis	Keyword and Citation Co-occurrence Analysis	[15, 29, 9]

The methodological patterns in TABLES 10 and 11, correlate with the geographic clusters identified in Section 3.8. Research groups from Mexico [30] concentrate on SDN-based mitigation, reflecting their institutional focus at Tecnológico de Monterrey. Federated learning contributions [31] align with growing IoT security research in the Middle East and South Asia.

4.4 Evaluation Metrics

Most of the research work have evaluated the techniques based on accuracy, precision, recall and F1-score. However, some the researches are using other metrics which are mentioned in TABLE 10.

Table 10: Evaluation Metrics Used in Botnet and Intrusion Detection Literature

S.No.	Evaluation Metrics	Citation(s)
1	Accuracy, Precision, Sensitivity, False Alarm Rate (FAR)	[15, 6, 9]
2	Accuracy, Detection Rate, Delay (ms), Resource Usage (SRAM/TCAM)	[18, 19]
3	Accuracy, Recall, F1-score, False Positive Rate (FPR), False Negative Rate (FNR)	[8, 7, 32]
4	Hybrid Performance Accuracy, Efficiency	[16, 20]
5	Average Detection Rate, Mitigation Success Rate (%), Flow Sampling Rate	[12]
6	Accuracy, ROC Curve, AUC, Precision, Recall, F1-score	[10, 17]
7	Throughput (%), Availability, Abnormal Traffic Detection Accuracy	[33, 34]
8	Resource Consumption, Latency, Detection Accuracy	[25, 26]
9	Mean Absolute Error (MAE), Accuracy, F-measure	[21, 14]
10	Recall, Precision, Accuracy, F1-score (Hyperparameter Optimization results)	[23, 22]
11	Packet Drop Ratio, Network Resilience, Detection Accuracy	[11, 13]
12	Mitigation Efficiency (%), Success Rate, Real-world Physical Assessment	[27, 12]

4.5 Challenges in Intrusion Prevention Systems for DDoS Attacks

The development and deployment of effective intrusion prevention systems (IPS) for Distributed Denial of Service (DDoS) attacks encounter several significant challenges that warrant thorough examination. Some of those challenges are listed below:

1. Firstly, real-time detection and mitigation capabilities remain critical bottlenecks, as many machine learning and deep learning models, such as Long Short-Term Memory (LSTM) networks and Convolutional Neural Networks (CNN), impose substantial computational overhead. This situation renders them impractical for use in high-speed or resource-constrained environments, such as Software-Defined Networking (SDN) controllers and Internet of Things (IoT) gateways.

2. The training and evaluation process for IPS models faces obstacles because there are not enough complete real-world datasets available for their development. The NSL-KDD and CICDDoS2019 datasets which researchers use most often fail to include modern attack signatures and application-layer communication protocols and encrypted data streams and authentic network interference which makes their models less applicable to real network environments.

3. The third point shows that IPS systems fail to operate properly because they generate incorrect alerts which create false positives and false negatives. The system blocks real network traffic because it produces high false positive rates which damages Quality of Service (QoS) but false negatives allow dangerous network traffic to go undetected. The process of detecting threats while reducing system interruptions to their lowest level creates an extremely difficult situation.

4. The lack of explanation in machine learning and deep learning-based IPS systems creates trust problems which block their implementation in operational settings. Most models operate as unknown systems which block investigators from understanding their detection mechanisms because they function as unexplainable systems that prevent effective troubleshooting and system optimization and emergency management. The domain of adversarial robustness remains an underexplored area within IPS development. Organizations do not check their existing systems for resistance against evasion techniques and adversarial samples which makes their systems open to attacks that target these vulnerabilities.

5. Organizations face ongoing difficulties when they try to link different IPS systems which operate in various mixed technology environments. The deployment process becomes more complex because organizations need to manage host-based systems and network-based systems and cloud-based systems while they protect data privacy and follow all regulatory rules during federated and international operations.

4.6 Future Directions and Research Opportunities

The reviewed literature shows that smart Intrusion Prevention Systems (IPS) now protect Distributed Denial of Service (DDoS) attacks better than before yet multiple existing obstacles still need to be solved. Next-generation systems need to find the proper equilibrium between their ability to produce exact results and their ability to explain their decisions and their operational speed. The deployment of lightweight models together with pruning methods will improve deployment capabilities for Software-Defined Networking (SDN) and Internet of Things (IoT) networks. The design process needs to make explainability its top priority for all development activities. Security analysts achieve better trust through decision trees and neural network models which produce understandable outputs that help them respond to incidents. The research method from [28] operates effectively in this particular situation. The system requires solid benchmarking to maintain its operational standards. The development of standardized DDoS datasets which include multiple attack types and encrypted data and real-world deployment artifacts will enable proper evaluation of these datasets. The research field will progress faster through its use of online competitions and simulators which function as collaborative platforms for real-time testing. Researchers need to concentrate on generating synthetic traffic through generative artificial intelligence systems which will also help them develop better defensive systems against attacks and improve machine learning models. The combination of multi-layered IPS systems which unite host-based detection with

network-based detection and cloud-based detection systems will create a more resilient security system.

Finally, it is critical to address the legal and ethical considerations associated with IPS design, particularly in federated and host-involved frameworks. The actual deployment of automated protection systems needs organizations to maintain data authority while showing their users all their system operations and they must follow their current policy requirements.

5. CONCLUSION

The research paper examines security systems that protect against DDoS attacks through their system structures and their use of machine learning algorithms and their methods for extracting features and their performance evaluation systems and their current research developments. The security domain has advanced from basic rule-based systems to modern detection systems which use SDN technology together with machine learning ensembles and edge computing systems.

The combination of machine learning with deep learning techniques has brought superior detection results through the use of Long Short-Term Memory (LSTM) and Convolutional Neural Networks (CNN) models. The system faces two major obstacles because it requires expensive computing resources while developers struggle to understand its operations. The fusion of federated learning with hardware-based detection systems offers new possibilities for building scalable IPS solutions. The process faces multiple challenges because it requires various data types and needs better systems that explain their decisions and sustain their performance. The majority of research studies focus on detection performance, but they do not perform complete assessments under authentic operational conditions.

The research needs to develop lightweight IPS models that explain their decisions while adapting to different network environments of various systems. Organizations need to develop strong IPS defense systems through generative artificial intelligence progress and collaborative defense system development, while creating standardized evaluation methods for tracking DDoS threat changes.

As the number of papers available ON IPS is comparatively less and the research is done only on 352 records obatined from Scopus and WOS databases, limiting the scope of this bibliometric analysis. The study can be extended to incorporate papers from other databases.

References

[1] Kolias C, Kambourakis G, Stavrou A, Voas J. DDoS in the IoT: Mirai and Other Botnets. Computer. 2017;50:80-84.

[2] Antonakakis M, April T, Bailey M, Bernhard M, Bursztein E, et al. Understanding the Mirai Botnet. In: 26th USENIX Security Symposium (USENIX Security 17). 2017:1093-1110.

[3] https://www.paloaltonetworks.in/cyberpedia/what-is-a-fast-flux-network.

[4] Andriesse D, Rossow C, Bos H. Reliable Recon in Adversarial Peer-to-Peer Botnets. In: Proceedings of the 2015 Internet Measurement Conference. 2015:129-140.

[5] Lee W, Wang C, Dagon D. Botnet Detection: Countering the Largest Security Threat. Springer Science & Business Media. 2007.

[6] Shone N, Ngoc TN, Phai VD, Shi Q. A Deep Learning Approach to Network Intrusion Detection. IEEE Trans Emerg Top Comput Intell. 2018;2:41-50.

[7] Gaonkar S, Dessai NF, Costa J, Borkar A, Aswale S, et al. A Survey on Botnet Detection Techniques. In: 2020 International Conference on Emerging Trends in Information Technology and Engineering (IC-ETITE). 2020:1-6.

[8] Meidan Y, Bohadana M, Mathov Y, Mirsky Y, Breitenbacher D, et al. N-BaIoT: Network-Based Detection of IoT Botnet Attacks Using Deep Autoencoders. 2018. ArXiv preprint: https://arxiv.org/pdf/1805.03409.

[9] Liu H, Lang B. Machine Learning and Deep Learning Methods for Intrusion Detection Systems: A Survey. Applied Sciences. 2019;9:4396.

[10] Yungaicela-Naula NM, Vargas-Rosales C, Perez-Diaz JA. SDN-Based Architecture for Transport and Application Layer DDoS Attack Detection by Using Machine and Deep Learning. IEEE Access. 2021;9:108495-108512.

[11] Maseer ZK, Yusof R, Al-Bander B, Saif A, Kadhim QK. Meta-Analysis and Systematic Review for Anomaly Network Intrusion Detection Systems: Detection Methods, Dataset, Validation Methodology, and Challenges. 2023. ArXiv preprint: https://arxiv.org/pdf/2308.02805

[12] Perez-Diaz JA, Valdovinos IA, Choo KK, Zhu D. A Flexible SDN-Based Architecture for Identifying and Mitigating Low-Rate DDoS Attacks Using Machine Learning. IEEE. 2020;8:155859-155872.

[13] Syaifuddin S, Kusumawardani SS, Widyawan W. Tackling DDoS Attacks in IoT: A Synthesis of Literature 2018 to 2022. Int J Intell Syst Appl Eng. 2024;12:802-809.

[14] Alabdulatif AA, Thilakarathne NN, Aashiq M. Machine Learning Enabled Novel Real-Time IoT Targeted DoS/DDoS Cyber Attack Detection System. Comput Mater Continua. 2024;80:3655-3683.

[15] Amina S, Vera R, Dargahi T, Dehghantanha A. A Bibliometric Analysis of Botnet Detection Techniques. Handbook of Big Data and IoT Security. 2019:345-365.

[16] Ali M, Shahroz M, Mushtaq MF, Alfarhood S, Safran M, et al. Hybrid Machine Learning Model for Efficient Botnet Attack Detection in IoT Environment. IEEE Access. 2024;12:40682-40699.

[17] de Caldas Filho FL, Soares SC, Oroski E, de Oliveira Albuquerque R, Da Mata RZ, et al. Botnet Detection and Mitigation Model for IoT Networks Using Federated Learning. Sensors. 2023;23:6305.

[18] Zhou J, Xu Z, Rush AM, Yu M. Automating Botnet Detection With Graph Neural Networks. 2020. ArXiv preprint. https://arxiv.org/pdf/2003.06344.

[19] François J, Aib I, Boutaba R. FireCol: A Collaborative Protection Network for the Detection of Flooding DDoS Attacks. IEEE/ACM Transactions on Networking. 2012;20:1828-1841.

[20] Fadel MM, El-Ghamrawy SM, Ali-Eldin AM, Hassan MK, El-Desoky AI. HDLIDP: A Hybrid Deep Learning Intrusion Detection and Prevention Framework. Comput Mater Contin. 2022;73:2293-2312.

[21] Kumar J, PJ AL. Mitigate Volumetric DDoS Attack Using Machine Learning Algorithm in SDN-Based IoT Network Environment. Int J Adv Comput Sci Appl. 2023;14.

[22] Coscia A, Dentamaro V, Galantucci S, Maci A, Pirlo G. Automatic Decision Tree-Based NIDPS Ruleset Generation for DoS/DDoS Attacks. J Inf Secur Appl. 2024;82:103736.

[23] Dasari S, Kaluri R. An Effective Classification of DDoS Attacks in a Distributed Network by Adopting Hierarchical Machine Learning and Hyperparameters Optimization Techniques. IEEE Access. 2024;12:10834-10845.

[24] Amitha M, Srivenkatesh M. Design of a Hypermodel Using Transfer Learning to Detect DDoS Attacks in the Cloud Security. Int J Adv Comput Sci Appl. 2023;14.

[25] Hatzivasilis G, Soultatos O, Chatziadam P, Fysarakis K, Askoxylakis I, et al. WARDOG: Awareness Detection Watchdog for Botnet Infection on the Host Device. IEEE Trans Sustain Comput. 2019;6:4-18.

[26] Myridakis D, Myridakis P, Kakarountas A. A Power Dissipation Monitoring Circuit for Intrusion Detection and Botnet Prevention on IoT Devices. Computation. 2021;9:19.

[27] Timcenko VV. An Approach for DDoS Attack Prevention in Mobile Ad Hoc Networks. Elektronika ir Elektrotechnika. 2014;20:150-153.

[28] Zargar ST, Joshi J, Tipper D. A Survey of Defense Mechanisms Against Distributed Denial of Service (Ddos) Flooding Attacks. IEEE Commun Surv Tutor. 2013;15:2046-2069.

[29] Goranin N, Hora SK, Čenys HA. A Bibliometric Review of Intrusion Detection Research in IoT: Evolution, Collaboration, and Emerging Trends. Electronics. 2024;13:3210.

[30] Yungaicela-Naula NM, Vargas-Rosales C, Pérez-Díaz JA, Carrera DF. A Flexible SDN-Based Framework for Slow-Rate Ddos Attack Mitigation by Using Deep Reinforcement Learning. J Netw Comput Appl.2022;205:103444.

[31] Rashid MM, Khan SU, Eusufzai F, Redwan MA, Sabuj SR, et al. A Federated Learning-Based Approach for Improving Intrusion Detection in Industrial Internet of Things Networks. Network. 2023;3:158-179.

[32] García-Teodoro P, Díaz-Verdejo J, Macía-Fernández G, Vázquez E. Anomaly-Based Network Intrusion Detection: Techniques, Systems and Challenges. Comput Secur. 2009;28:18-28.

[33] Eliyan LF, Di Pietro R. DoS and DDoS Attacks in Software Defined Networks: A Survey of Existing Solutions and Research Challenges. Future Gener Comput Syst. 2021 Sep 1;122:149-71.

[34] Suthar F, Patel N, Khanna S. A Signature-Based Botnet (Emotet) Detection Mechanism. Int J Eng Trends Technol. 2022;70:185-193.