Project Completion Report

Scientometric mapping of global research on cyber security with special reference to India's status

Study implemented by

Dr. B. ELANGO
Principal Investigator
Librarian, IFET College of Engineering,
Gangarampalayam – 605 108. Villupuram – Dt.

Project No. DST/NSTMIS/05/236/2017-18

Sponsored by

National Science and Technology Management Information System (NSTMIS)
Department of Science & Technology (DST)
Government of India

DECEMBER 2021

©NSTMIS Division 2021

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying, recording or otherwise without the prior permission of NSTMIS(DST). Readers may make verbatim copies of this document for non-commercial purposes by any means, provided that the above copyright notice appears on all copies.

NSTMIS Division
Department of Science & Technology
Ministry of Science & Technology
Technology Bhawan, New Mehrauli Road, New Delhi-110016
Phone: 91-011-26567373
Website: www.nstmis-dst.org/

About NSTMIS:

The National Science and Technology Management Information System (NSTMIS), a division of Department of Science and Technology (DST) has been entrusted with the task of building the information base on a continuous basis on resources devoted to scientific and technological activities for policy planning in the country.

Citation:

The report may be cited as DST (2021): Report on "Scientometric mapping of global research on cyber security with special reference to India's status"; Dr. B. Elango, IFET College of Engineering, Villupuram. New Delhi, Govt. of India.

Disclaimer:

Every care has been taken to provide the authenticated information. However, the onus of authenticity of data rests with the PI of the project.

PREFACE

Cyber Security is a distinct domain that pertains to and is a component of new technologies such as artificial intelligence, the internet of things, big data, advanced mobile computing, cloud computing, e-commerce and other developing technologies. As vast volumes of data have been kept on computers and other devices by financial, government, military, medical and corporate or other institutions in the digital era, it is essential to protect sensitive information from the intruders. Hence this study has been undertaken to explore the publication pattern in the field of cybersecurity with special reference to India's status. For this purpose, the bibliographic records have been collected from various databases such as Web of Science, Scopus, Indian Citation Index and Indian Science Abstracts. Analysed data has been presented in the form of tables and figures.

The research report has been presented in the following chapters:

Chapter 1	-	Introduction
Chapter 2	-	Review of literature
Chapter 3	-	Methodology
Chapter 4	-	Data analysis and interpretation
Chapter 5	-	Results and discussion
Chapter 6	-	Findings and recommendations

The report concludes with a bibliography and appendices

The investigators believe that this research report will create awareness, provide help to the scientists working in this domain, and serve as an eye-opener for student community.

Principal Investigator

ACKNOWLEDGEMENT

I extend my sincere thanks to DST officials for approving my project proposal and provide financial support to carry out this project. The comments and suggestions provided by the members of the Expert Committee on Bibliometrics have been gratefully acknowledged.

Co-Principal Investigators Dr. S. Matilda, Professor & Mr. M Arul Pugazhendhi, Associate Professor of IFET College of Engineering, Villupuram, Tamilnadu are also acknowledged for their constant support.

I would like to thank LPAC members who offered their suggestions during the meeting.

My sincere thanks to the subject experts Dr. Sandeep K Shukla of Indian Institute of Technology Kanpur and Prof. M. Sethumadhavan of Amrita Vishwa Vidyapeetham for their constructive comments and suggestions in the keyword design.

My thanks to the authorities of Anna University Library and Central University of Karanataka Library for providing the access to databases for data collection.

I am also thankful to my colleagues those who have helped directly or indirectly during the study period.

Finally, I am grateful to the project assistants worked in this project in the different periods. Especially the work of recent project staff Ms.M.Martina Jose Mary has been acknowledged.

Dr. B. ELANGO
Principal Investigator

Table of Contents

List of Tables

List of Figures

List of Abbreviations

B – Book

BC – Book Chapter

C – Conference

CAGR – Compound Annual Growth Rate

D – Discontinued

DT – Document Type

ICI – Indian Citation Index

IIT – Indian Institute of Technology

J – Journal

NIT – National Institute of Technology

PY – Publication year

TC – Total Citations

TP – Total publications / papers / documents

UK – United Kingdom

USA – United States of America

Executive Summary

Cyber Security is a distinct domain that pertains to and is a component of new technologies such as artificial intelligence, the internet of things, big data, advanced mobile computing, cloud computing, e-commerce and other developing technologies. Because vast volumes of data have been kept on computers and other devices by financial, government, military, medical and corporate or other institutions in the digital era, it is critical. Hence this study has been undertaken to explore the publication pattern in the field of cybersecurity with special reference to India's status, with the following specific objectives:

- To examine the status of global S&T in the field of cyber security.
- To identify the size and growth of publication output in the field cyber security.
- To measure the publication quality
- To assess the status of India's contribution and compare it with selected countries using publication and citation metrics.
- To examine the prolific authors as well institutions in the field of cyber security in terms of publication output.
- To rank the Indian institutions in terms of publication output.
- To analyze the publications according to the following different perspectives: security component, application domain, objective and intervention level.

To achieve the above objectives, the present study uses the publication data from various databases such as Web of Science, Scopus, Indian Citation Index, and Indian Science Abstracts. To provide a recent trend, the study covers publications up to 2020,

as suggested by Local Project Advisory Committee. The project has been carried out in a phased manner.

In the first phase, delineation of keywords has been undertaken. Available keywords have been collected manually from the existing literature, thesaurus (IEEE Thesaurus and WordNet) and international standards organization (NIST, Glossaries of British Standards Institution, and the National Initiative for Cyber Security Careers and Studies). Among the sources, WordNet has been suggested by the expert committee during the review meeting held on 13.09.2019 @ CEPT University, Ahmadabad. Clustering of the collected keywords has been made and the unrelated or unwanted keywords have been discarded. For example, the keyword "*voting system*" is not directly related to cybersecurity. Finally, potential keywords have been sent to the subject experts and after getting the suggestions and opinions, the following keywords have been identified: "Cybersecurity" OR "Cyber Security" OR "Cyber-Security" OR "Cyber crisis management" OR "Cyber incident management" OR "Cyber threat management" OR "Cyber Safety" OR "Cybersafety" OR "Cyber defense". The procedure followed in this phase has been documented and presented in a conference.

In the second phase, three pilot studies have been undertaken to understand the publication structure: top-cited publications, role of cybersecurity in smart grid and India's contribution to cybersecurity. These studies help us to narrow down the research further.

In the final phase, data collection and analysis have been undertaken.

Major findings of this study are:

- India is ranked 4th in terms of number of publications behind the USA, the UK and China.
- Indian researchers have been publishing on this topic since 2004, six years after global researchers who began in 1998. It clearly evidences that it takes six years to raise awareness among the Indian researchers.
- Fifty-four percent of the global research output has been published in the recent three years (2018-2020) where as it is 67% for the Indian output, which reveals that Indian researchers recognized this topic very recently.
- In India, 46% publications were in the form of conference papers, which is relatively low compared to the global trend.
- The value of the collaboration index (CI) for Indian publications is greater than the value for global output, indicating that Indian researchers prefer to do research in groups.
- In the global environment, the USA was the most productive country as well as the most preferred partner country for international collaboration by Indian researchers.
- In terms of citation impact, Indian publications received 5.66 citations per paper on average, which is somewhat lower than the global output. It is mostly due to research findings being published in the non-standard sources. For example, almost one-third of the most preferred sources by Indian researchers have been discontinued its coverage by Scopus.

- Indian researchers commonly collaborate within institutions, as seen by research teams made up of Indian researchers.
- The focus of Indian researchers has shifted from data-related issues to smart grid and emerging technologies, to AI-related technologies.
- Indian institutions having centers of excellence in cybersecurity produces more publications than others.

Based on the data analysis and interpretation, the following suggestions have been made:

- It is suggested to create centers of excellence in all the universities and higher educational institutions in order to active participation in research in the field of cybersecurity.
- It is recommended to establish a national level center for creating the awareness about predatory journals as well as providing training in doing the high quality research.
- It is strongly recommended to concentrate on the application of emerging technologies into cybersecurity.
- More research must be carried out on the smart grid cybersecurity as well as the role of cybersecurity in the energy sector.
- A database has been created with the classification of publications based on the cybersecurity taxonomy: research domains, technology & use cases, and sector, which needs follow-up action.

- Only publications indexed in Scopus, Web of Science, Indian Citation Index and Indian Science Abstracts were used in this study. An exclusive study of non-indexed publications could be undertaken.
- Compared to other topics, there are only few publications by Indian researchers. In this regard, an awareness program is needed.

Chapter I

INTRODUCTION

Preamble

The term *Cybersecurity* is the state of being protected from cyber-criminals or unauthorized access of electronic data or the measures taken to achieve this. Cyber Security is a distinct domain that pertains to and forms a part of new technologies such as artificial intelligence, internet of things, big data, advanced mobile computing, cloud computing, e-commerce and other emerging technologies. Because vast volumes of data being stored on computers and other devices by financial, government, military, medical and corporate or other organizations in the digital era, it is critical.

Cybersecurity is a key feature in the digital world, where information and communication technology (ICT) is increasingly used. Security measures do not evolve as data accessibility. Hackers employ a technique known as cyber-attacks to gain access to data without the user's permission. Cyber-attacks are distributed through phishing and spam emails, infected websites, macros, web attacks, exploit kits, lateral movement, and botnet. Virus, worms, ransomware, backdoors, downloaders, botnets, key loggers, remote access tools (RAT), and password stealers are all examples of cyber-attacks. Big data, artificial intelligence, machine learning, block chain, internet of things, e-commerce, and data analytics and so on is the most emerging technology related to the field of cybersecurity which are able to protect the network, infrastructure, hardware, software, mobile and, system.

"Cybersecurity" refers to the process of decreasing the danger of hostile assaults on software, computers and networks. It comprises tools for detecting intrusions, preventing viruses and unauthorized access, imposing authentication and enabling encrypted communications, among other things (Amoroso 2007). Technology shifts, outsourcing, and economic mode shifts all have an impact on cybersecurity. Cloud computing, big data, and the internet of things are developing technologies in the fourth industrial revolution that will automate and exchange data. Cyber-attacks are attracted to automation, necessitating security measures. The smart grid is an automated system that is vulnerable to cyber-attacks.

Governments all over the world are conducting continual cybersecurity research in order to protect against new and emerging threats. In India, cyber technologies are now widely used in a range of fields. Because India is a key outsourcing destination for IT and Business Process Management (BPM) services, security and privacy must be considered. In the domain of cybersecurity, research and development (R&D) is in great demand (Mallick 2019). Some Indian universities have attempted to combine cybersecurity with artificial intelligence in their research. The Indian Institute of Technology Kanpur, for example, has established a National Interdisciplinary Centre for Cyber Security and Cyber Defence of Critical Infrastructures, with a goal of creating and deploying technical safeguards to protect critical infrastructures. According to the Indian Computer Emergency Response Team (CERT-In), 607,220 cybersecurity incidents were reported in the first half of 2021, nearly double the number of incidents reported in 2019. In May 2021, the FBI and the Australian Cyber Security Centre issued a warning about the ongoing Avaddon ransomware outbreak, which is hitting a range of industries in several countries, including India. In this context, the purpose of this study is to assess the cybersecurity research done by global researchers with special reference India's status in response to cyber-related concerns.

Project objectives

The primary goal of this project is to map cybersecurity research output with the following specific objectives:

- To examine the status of global S&T in the field of cyber security.
- To identify the size and growth of publication output in the field cyber security.
- To measure the publication quality
- To assess the status of India's contribution and compare it with selected countries using publication and citation metrics.
- To examine the prolific authors as well institutions in the field of cyber security in terms of publication output.
- To rank the Indian institutions in terms of publication output.
- To analyse the publications according to the following different perspectives: security component, application domain, objective and intervention level

Scope and limitations

The project is being carried out using the bibliographic records retrieved from various databases such Scopus, Web of Science, Indian Citation Index and Indian Science Abstracts. Publications indexed on the date of access of databases were considered for this project. Other databases as well as non-indexed publications were not considered for this study.

Chapter II

REVIEW OF LITERATURE

Few researchers have attempted to investigate the impact of cybersecurity on the other areas in the past. For example, Parvin et al. (2019) evaluated the research trends of information security in the Middle-East and the world from a scientometric perspective and found that the majority of scientific publications in the field of information security were produced by the United States of America and China. Jalali et al. (2019) did a bibliometric review of the literature to describe the state of research on various aspects of cybersecurity in health care system using the data from PubMed and Web of Science. Rahima et al. (2020) looked at the literature trends in the areas of cyber security and higher education and discovered that the majority of documents were conference papers. In order to uncover contemporary topics, authors, and themes in the field of cyber security, Furstenau et al. (2020) conducted a science mapping analysis.

In the past, very few scholars attempted to map the field of Cyber Security with narrower keywords or unrelated keywords. For example, Chang (2016) used the following narrower keywords to retrieve the bibliographic records: *"Cyber- security", "Cyber Security", or "cybersecurity".* Similarly, Cojocaru and Cojocaru (2019) used the keywords "Cybersecurity" or "cyber security" to illicit bibliographic records from three databases like Web of Science, SCOPUS and IBN - National Bibliometric Instrument from the Republic of Moldova. On the other hand, Jalali et al. (2019) used a set of keywords from the viewpoint of cyber-attack such as "cyber crisis", "cyber incidence", cyber infrastructure", "cyber operation", "cyber risk", etc. In the same way, Christen et al. (2017) and Abbas et al. (2019) used a set of keywords from the viewpoint of broader concept such as “IT Security”, “computer security”, “hardware security”, “mobile security”, “system security”, “security software”, etc. Surprisingly, Rai, Singh and Varma (2019) searched the title in SCOPUS with “cyber” and “security”. Dhawan et al. (2021) used the two narrower keywords “cybersecurity” OR “cyber security” in the title and keywords fields in Scopus to retrieve the bibliographic records.

From the literature review, various authors use a different set of keywords. It is to be noted that even though the terms *information security* and *cybersecurity* have been used interchangeably by the researchers in general, they are different concepts with commonalities (Von Solms and Van Niekerk 2013). For example, information security is concerned with the protection of any type of information, whereas computer security is concerned with the protection of a standalone computing, and cybersecurity if concerned with the prevention of information in cyberspace, which is a personal as well as global concept (Figure 1).

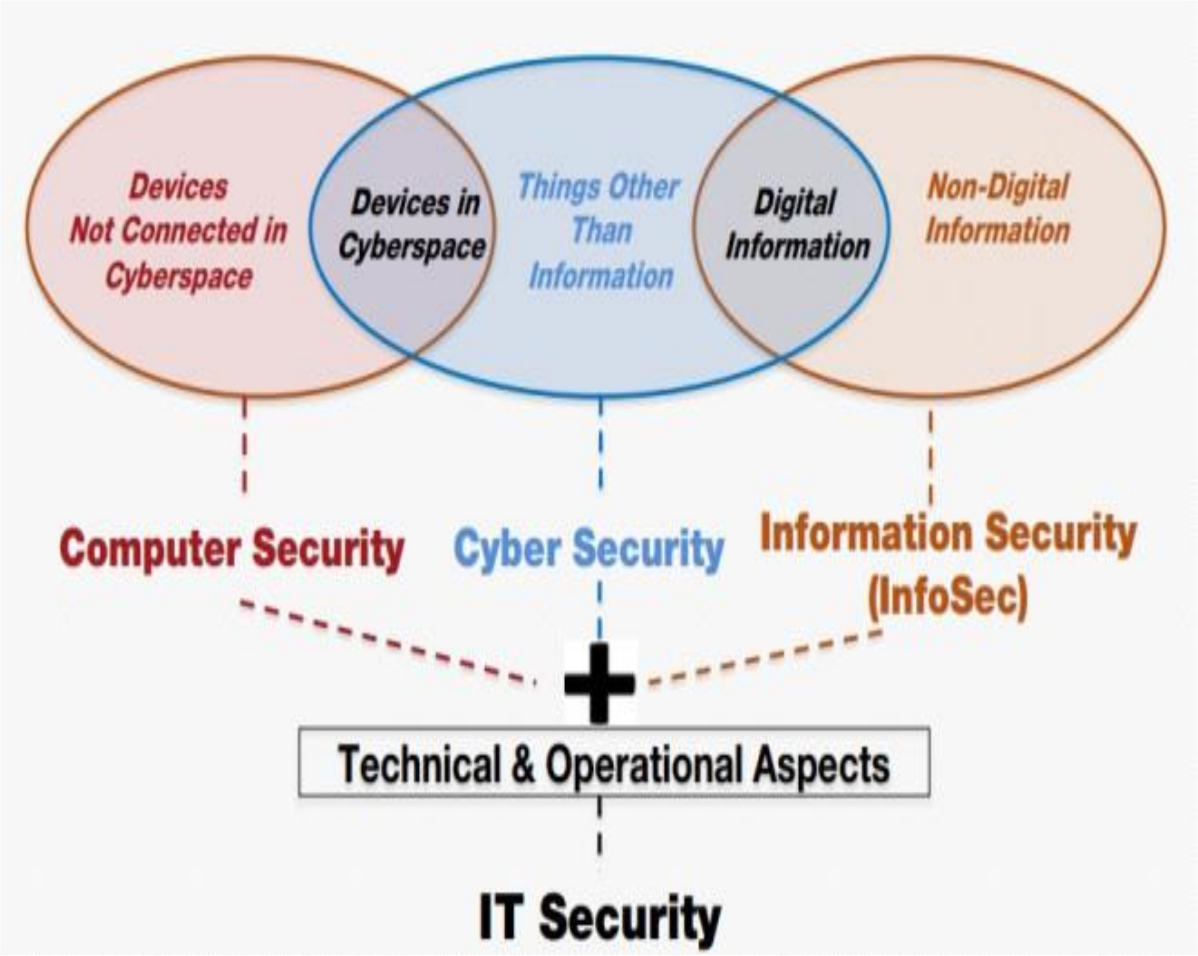

Devices Not Connected in Cyberspace
Devices in Cyberspace
Things Other Than Information
Digital Information
Non-Digital Information
Computer Security
Cyber Security
Information Security (InfoSec)
Technical & Operational Aspects
IT Security

Figure 1 – Cybersecurity and interrelated domains
(Source: https://vncybersecu.com/2018/10/29/cyber-security-vs-it-security/)

The following observations are made based on the literature published up to 2020: (1) there has no exclusive study in this domain (2) there has no detailed analysis on developing a set of keywords for the emerging field of cybersecurity.

Chapter III

METHODOLOGY

Keywords Delineation

The process of defining the keywords has been done in phases. In the first phase, ambiguity of Cyber Security domain has been discussed with interrelated domains. In the second phase, available keywords were manually collected from existing literature, thesaurus (IEEE Thesaurus and WordNet) and international standards organizations (NIST, Glossaries of British Standards Institution, and the National Initiative for Cyber Security Careers and Studies). Among the sources, WordNet has been suggested by the expert committee during the review meeting held on 13.09.2019 @ CEPT University, Ahmadabad. The collected keywords were clustered in the third phase, and the unrelated or unwanted keywords were removed in the final phase (Table 1). The term "*voting system*", for example, has nothing to do with cybersecurity.

Table 1 – Clustering of available keywords

Cluster	**Related Keywords**	**Definition**	**Keyword meets the definition of Cyber Security**
Cyber Security	Cyber crisis management, cyber safety, cyber security management, cyber threat management, Cyber incident Management, Cyber defense	Ability to protect the cyberspace from cyber-attacks.	Yes
	Worm, virus, Malicious code ,spam, Trojan, malware, malvertising,	An incident or circumstance which has the	No

Cyber threat – cyber crime	Scare ware, Adware, Backdoor, Amplication attack, browser hijacker, Botnet, command and control server, Data breach / leakage / data loss, data theft, security breach, security threat, Security Incident, exfiltration, Denial of service, digital sabotage, exploit kit, Advanced Persistent Threat, fast flux, hacktivism, identity theft, security vulnerability, spoofing, spyware, Advanced Persistent Threat, Cyber Bullying, cyber harassment, Cyber espionage or Cyber Spy, cyber risk, cyber conflict, cyber terrorism, Vishing, pharming, twishing, Reflection attack, Rootkit, Root to local Attack, User to root attack	potential to cause a loss of assets and the unintended effects or effect of such loss.	
Hackers	Black hat, white hat, Botnet, crypto-locker, dropper, key logger, script kiddies, Bot herder, botmaster, Hacktivist ,Threat Hunters, Phisher	Hackers are unauthorized users who hack into computer systems in order to steal, modify or damage information.	No
Broader concept	Computer security, Information IT security, System security, Web security, Online security, Internet security, Mobile security, Telecommunication security, Cybernetics	There is a wider scope.	No

Network security	Cryptography or cryptology, IP Security, Packet Sniffing prevention, Firewall, Intrusion prevention systems, Pen test, Intrusion Detection and Prevention system	It is related to preventing and tracking unapproved access, violence, alteration, or denial of a network of computers and resources that are available in the network.	No
Cyber Security Framework	Cyberspace, Cyber Infrastructure, Cyber Ecosystems, cyber-physical system	This framework consists of an interdependent network of infrastructures for information systems which includes the Internet, telecommunication and computer systems.	No

Finally, identified keywords have been sent to the following subject experts:

1. Dr. Sandeep K. Shukla,
   Head, Department of Computer Science and Engineering
   Indian Institute of Technology Kanpur,
   Kanpur, India.

2. Prof. M. Sethumadhavan
   Head of TIFAC-Centre of Relevance and Excellence in Cyber Security,
   Amrita Vishwa Vidyapeetham,
   Coimbatore Campus.

Among these two subject experts, Dr. Sandeep K Shukla is one of the top authors (refer Table 22). After getting the opinions and suggestions from the above experts, the keywords listed in table 2 have been finalized.

Table 2 – Identified potential keywords

Keywords	Definition
Cyber Security	Prevention of cyberspace from cyber-attacks.
Cyber crisis management	It provides the strategic framework and guidelines to prepare for, respond to, and begin to coordinate recovery from cyber incident or Crises (Readiness, response, and recovery)
Cyber incident management	Monitoring and detection of security incidents on-computer or network.
Cyber threat management	It is an early detection of risks, situational awareness driven by evidence, timely decision-making and threat mitigation actions.
Cyber safety management	It is safe practices to prevent attacks or threats on Internet.
Cyber defense	Prevention, detection and timely response to attacks or threats.

A comparison of number of keywords between the existing studies and potential keywords identified in this project has been provided in table 3.

Table 3 - Comparison of number of papers

Article	Keywords	# Documents in SCOPUS as on 20.11.2020	Remarks
Chang (2016)	“information security" OR “cybersecurity" OR “cyber security", OR “cyber-security"	39151	Author used broader term "information security" along with cybersecurity
Rai, Singh & Varma (2019)	cyber AND security	26367	Authors split the word cybersecurity
Abbas et al. (2019)	"Cyber Security" OR "Network Security" OR "Information	366676	Authors used broader terms as

	Security" OR "Telecommunication Security" OR "Data Privacy" OR "cyber attribution" OR "Intrusion Detection System" OR ids OR cryptography OR "Intrusion Prevention Systems" OR ips OR "Internet Security" OR "IP Security" OR "Preventing Packet Sniffing" OR "attack prevention mechanisms" OR "User to root attack" OR "Root to local Attack" OR "Cyber Attack" OR "IT security"		well as unwanted keywords
Jalali et al. (2019)	"Cybersecurity" OR "Cyber Security" OR "Cyber Attack" OR "Cyber Crisis" OR "Cyber Incident" OR "Cyber Infrastructure" OR "Cyber Operation" OR "Cyber Risk" OR "Cyber Threat" OR "Cyberspace" OR "Data Breach" OR "Data Security" OR "Firewall" OR "Information Security" OR "Information Systems Security" OR "Information Technology Security" OR "IT Security" OR "Malware" OR "Phishing" OR " Ransomware" OR "Security Incident" OR "Information Assurance"	94429	Authors used broader terms as well as unwanted keywords
Christen et al. (2019)	*botnet* OR "computer crim*" OR "computer security” OR cryptography OR cyberattack OR "cyber attack" OR cyber conflict OR "cyber conflict" OR "cyber crim*" OR cyber defense OR "cyber defense" OR cybersecurity OR "cyber-security" OR "cyber security" OR “cyber terrorism” OR "cyberterrorism" OR cyber thread* OR "cyber threat*" OR cyberwar* OR "cyber war*"OR "data leak*" OR "data security" OR "denial of service" OR DDoS OR firewall OR "hardware security"OR "information security" OR "internet security" OR "IT security" OR malware OR "mobile security" OR"network security" OR "non-repudiation" OR "security breaches" OR "security of data" OR "security requirement*" OR "security software" OR "security system*" OR "security threat*" OR "security vulnerabilit*" OR sigint OR "system security" OR "voting system" OR "web security"	401412	Authors used broader terms as well as unwanted keywords
This study	"Cybersecurity" OR "Cyber Security" OR "Cyber-Security" OR " Cyber crisis management" OR "Cyber incident management" OR "Cyber threat management" OR "Cyber Safety" OR "Cybersafety" OR "Cyber defense"	**18851**	**All the ambiguities are eliminated**

Search Strategy

The following set of keywords has been used to retrieve the bibliographic records related to cybersecurity:

"Cybersecurity" OR "Cyber Security" OR "Cyber-Security" OR "Cyber crisis management" OR "Cyber incident management" OR "Cyber threat management" OR "Cyber Safety" OR "Cybersafety" OR "Cyber defense"

Table 4 provides the information about the number records retrieved from the four databases.

Table 4 – Number of records retrieved

Database	Search	Period	No. of records
Global output			
Web of Science	Topic	1980-2020	3744
Scopus	Title-Abstract-Keywords	1999-2020	19932
Indian output			
Web of Science	Topic	2005-2020	124
Scopus	Title-Abstract-Keywords	1999-2020	994
Indian Citation Index	Topic	2004-2019	130
Indian Science Abstracts	Term	Any time	02

Data Collection

Bibliographic records have been downloaded in BibTex format from Web of Science and Scopus, and in Excel format from Indian Citation Index and Indian Science Abstracts.

Pre-processing

Downloaded individual (Scopus and Web of Science) BibTex files have been merged in the DOS command prompt and then uploaded separately in Biblioshiny, a web interface of Bibliometrix R Package and saved as .xlsx files with common tag fields. Then open in MS-excel and merge the data with common columns (table 5).

Table 5 – Examples of fields tags

Field Tag	Description
AB	Abstract
AU	Authors
C1	Author Address
CR	Cited References
DE	Author Keywords
DT	Document Type
LA	Language
PY	Year Published
SO	Publication Name / Source
TC	Number of citations
TI	Document Title

There are some variations among the two databases. For example, document type "editorial" in Scopus is denoted as "editorial materials" in Web of Science. Similarly, the country name "United States" in Scopus is denoted as "USA" in Web of Science. After merging the two datasets, automatic duplication removal has been done in MS-Excel followed by screening of each record has been done manually. For example, the following two titles are identical and duplicate record has been removed.

Zombies, Sirens, and Lady Gaga - Oh My! Developing a Framework for Coordinated Vulnerability Disclosure for <u>US</u> Emergency Alert Systems

Zombies, Sirens, and Lady Gaga – Oh My! Developing a Framework for Coordinated Vulnerability Disclosure for <u>U.S.</u> Emergency Alert Systems

Additional to the records from Web of Science and Scopus, records from Indian Citation Index and Indian Science Abstracts have been included for the Indian output. Final corpus of the dataset has been provided in table 6.

Table 6 – Final corpus

Item	No. of records
Global output	20039
Indian output	1108

Data Analysis

The merged file was uploaded in Biblioshiny, a user interface of Bibliometrix R Package to analyse the data into two categories:

Publication analysis

- Annual trend & growth of papers
- Scattering of literature
- Author productivity & Lotka's law
- Authors, institutions & countries

Science mapping

- Co-authorship
- Co-occurrence
- Word cloud
- Trend topics
- Factorial analysis

For science mapping, VOSViewer has also been used. A detailed research methodology has shown in the figure 2.

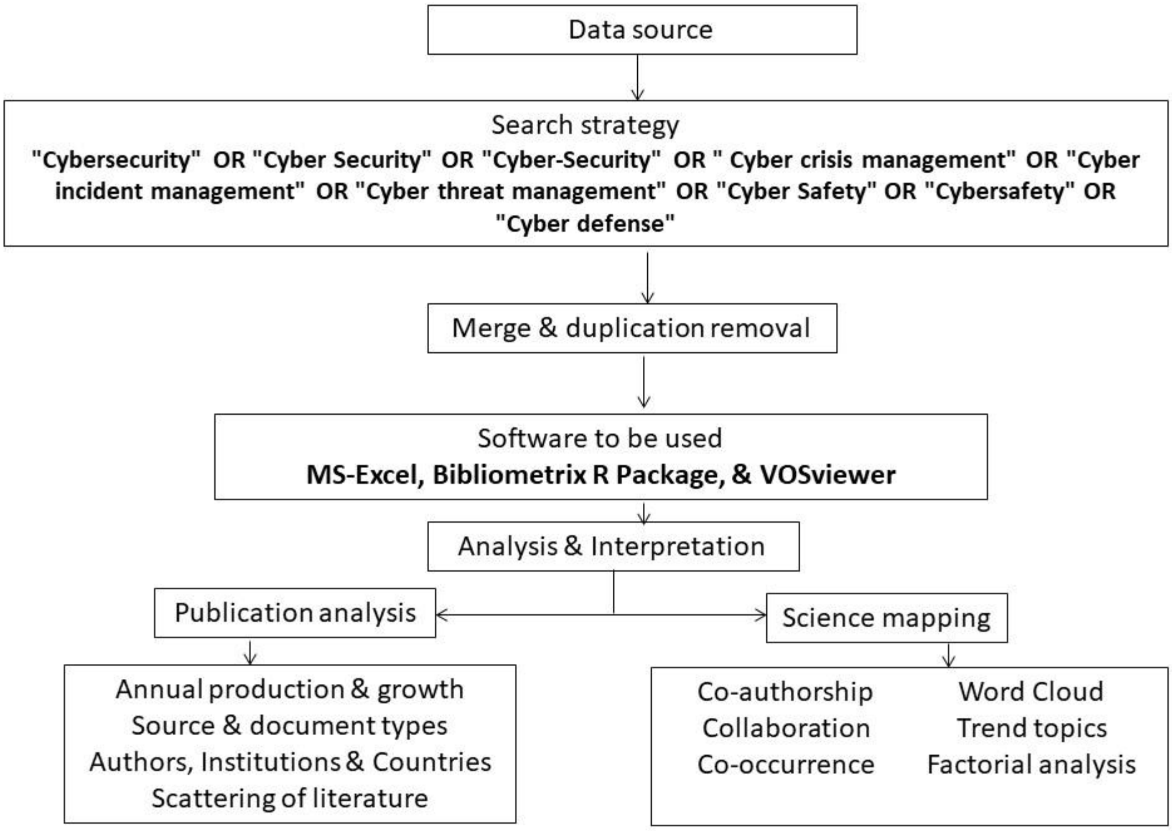

Data source
Search strategy
"Cybersecurity" OR "Cyber Security" OR "Cyber-Security" OR " Cyber crisis management" OR "Cyber incident management" OR "Cyber threat management" OR "Cyber Safety" OR "Cybersafety" OR "Cyber defense"
Merge & duplication removal
Software to be used
MS-Excel, Bibliometrix R Package, & VOSviewer
Analysis & Interpretation
Publication analysis
Science mapping
Annual production & growth
Source & document types
Authors, Institutions & Countries
Scattering of literature
Co-authorship
Collaboration
Co-occurrence
Word Cloud
Trend topics
Factorial analysis

Figure 2 – Systematic research methodology

Further, the following bibliometric tools have been employed:

- Compound Annual Growth Rate (CAGR)
- Collaboration Index (CI)
- Bradford Law of Scattering
- Lotka's Law
- Citations per Paper (CPP)
- h-index

Chapter IV

DATA ANALYSIS AND INTERPRETATION

In this chapter, a detailed analysis and its interpretation has been discussed in two parts: Global output and Indian output

GLOBAL OUTPUT

General overview

Table 7 provides the information on document types, document content, individual authors, and authors' collaboration. Global researchers have published 20039 papers in the field of cybersecurity in 5864 different sources between 1980 and 2020. In terms of citation impact, publications received an average of 7.41 citations per publication. A total of 27401 keywords have been appended by the authors in the 20039 papers with 1.36 keywords per publication. Most of the publications were contributed with co-author(s) (75%) which denotes that strong collaboration exists among the researchers in this field. There are 33156 authorships in the 20039 publications and the Collaboration Index (Elango and Rajendran 2012) is evaluated to 1.99, which indicates that the research team consists of approximately two authors in this field. The Collaboration Index (CI) is calculated by dividing the total number of authors of multi-authored papers by total number of multi-authored papers.

Table 7 – General information on global output

Description	**Results**
Timespan	1980:2020
Sources (Journals, Books, etc.)	5864
Documents	20039
Average citations per documents	7.41
Average citations per year per doc	1.61
References	567615
Keywords Plus (ID)	37221
Author's Keywords (DE)	27401
Authors	33156
Authors of single-authored documents	3080
Authors of multi-authored documents	30076
Single-authored documents	4925
Collaboration Index	1.99

Document type

Global researchers published their research findings in seventeen different document types which shown in table 8. There are different formats of document types followed by the databases, for instance of article in Scopus and research article in Web of Science. Hence, normalization has been done manually. Researchers had mainly preferred the conference paper (53%) because it takes short duration for feedback, presenting the work done so far and easy to interact with people for working on a same field (table 8). Next, article (30%) is another important document type to published their research findings on cyber security. Below 17% of total papers has been published in the long list of another 19 document types.

Table 8 – Document type of global output

Sl.	Document Type	TP	Share
1	Conference Paper	10674	53.27
2	Article	5990	29.89
3	Book Chapter	1010	5.04
4	Conference Review	728	3.63
5	Review	607	3.03
6	Editorial	286	1.43
7	Book	248	1.24
8	Note	178	0.89
9	Short Survey	148	0.74
10	News Item	75	0.37
11	Letter	21	0.10
12	Meeting Abstract	18	0.09
13	Book Review	17	0.08
14	Business Article	14	0.07
15	Erratum	9	0.04
16	Article in Press	8	0.04
17	Others	8	0.04
	Total	20039	100

Annual Production

The first publication was identified in 1980 as conference type of "Proceedings of the 18th Annual Southeast Regional Conference, ACM-SE 1980". After a long gap of 18 years, next publication was published in 1998 and from that researchers had focused on cybersecurity research continuously. The highest number of papers had published in 2020 with 4798 documents (figure 3) and highest growth has been observed (300%) in 1999. Almost 80% of the total papers were published between 2015 and 2020. Particularly, 54% of the total papers were published in the recent three-years (2018-2020).

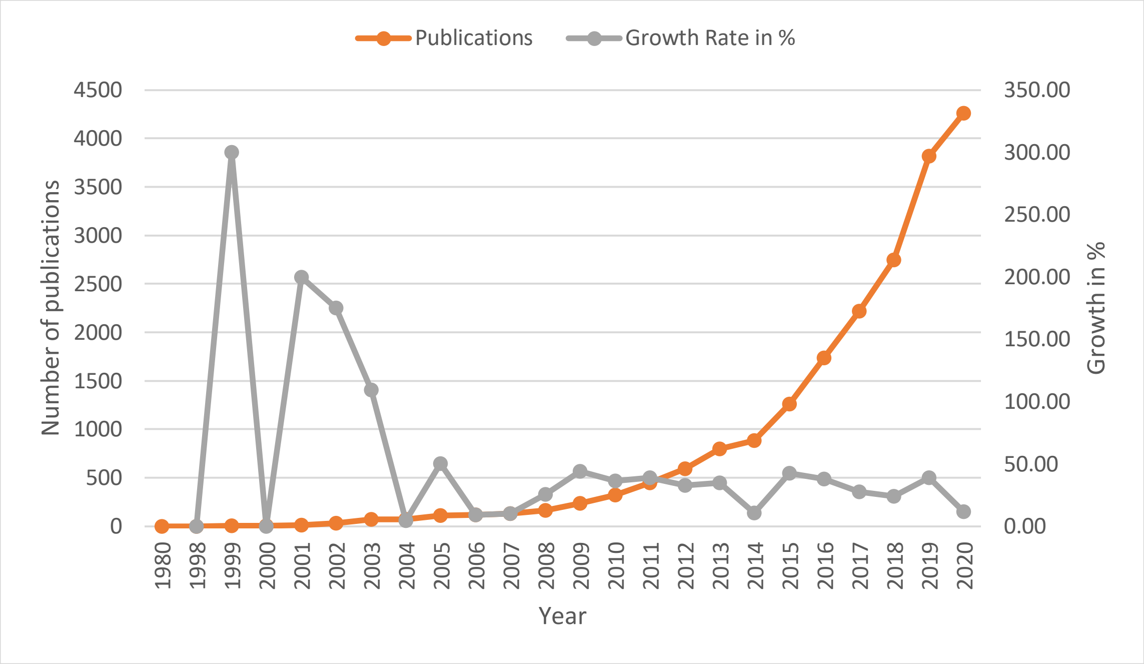


Figure 3 – Annual productivity and its growth of global output

Scattering of literature

The study of sources aids in the selection of the most appropriate sources for the publication of relevant studies. Total papers 20039 papers were published in 5864 different sources and table 9 provides the top most twenty sources in the field of cybersecurity research from 1980 to 2020. These top twenty sources published the 20% of total papers. We have ranked the sources based on the number of papers. Lecture Notes in Computer Science (including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) is the one of the top most productive source and had published a 795 documents followed by ACM International Conference Proceeding Series (n=540) and Advances in Intelligent Systems and Computing (n=364): these top three sources are book series and conference proceedings. Of the top 20 sources, conference proceedings were the most preferred one accounting to 45% followed by journals (30%) and book series (25%). All the top 20 sources were being published from the three countries, viz. USA (65%), Germany (20%) and UK (15%).

Table 9 – Top twenty sources of global output

Sources	Type	Country	TP	Rank
Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence And Lecture Notes in Bioinformatics)	BS	Germany	795	1
ACM International Conference Proceeding Series	C	USA	540	2
Advances in Intelligent Systems and Computing	BS	Germany	364	3
IEEE Access	J	USA	267	4
Communications in Computer and Information Science	BS	Germany	230	5
IEEE Security and Privacy	J	USA	205	6
European Conference on Information Warfare And Security ECCWS	C	USA	193	7
CEUR Workshop Proceedings	C	USA	164	8
Computers & Security	J	United Kingdom	140	9
Proceedings of SPIE - The International Society for Optical Engineering	C	USA	135	10
IFIP Advances in Information and Communication Technology	BS	USA	119	11
Proceedings of the Annual Hawaii International Conference on System Sciences	C	USA	105	12
Computer	J	USA	104	13
Proceedings - IEEE Military Communications Conference Milcom	C	USA	98	14
IEEE Transactions on Smart Grid	J	USA	94	15
ASEE Annual Conference and Exposition Conference Proceedings	C	USA	92	16
IEEE Power and Energy Society General Meeting	C	USA	92	16
Computer Fraud and Security	J	United Kingdom	91	17
IET Conference Publications	C	United Kingdom	88	18

Advanced Sciences and Technologies for Security Applications	BS	Germany	71	19
B = Book; BS = Book Series; C = Conference; J = Journal				

Top authors

A total of 33156 authors from 11277 institutions located in 115 countries contributed to the 20039 papers. Table 10 reveals the top 20 authors in the field of cybersecurity, these top authors having published papers between 32 and 79.

Table 10 – Most productive global authors

Sl.	Authors	TP	TC	h-index	PY-Start
1	Zhang Y	79	1394	18	2009
2	Wang L	61	1017	14	2006
3	Wang Y	59	643	14	2010
4	Zhang J	51	937	15	2008
5	Liu Y	45	575	12	2010
6	Wang J	43	1265	19	2013
7	Li Z	43	968	17	2012
8	Xiang Y	42	902	15	2011
9	Li J	42	508	10	2010
10	Liu X	40	720	15	2015
11	Zhu Q	40	643	17	2012
12	Chen H	40	569	15	2012
13	Kozik R	37	313	10	2012
14	Xu S	36	510	14	2009
15	Govindarasu M	36	2112	17	2007
16	Li Y	35	769	10	2008
17	Chora M	33	269	9	2012
18	Choo KKR	32	611	12	2014
19	Joshi A	32	553	9	2012
20	Liu CC	32	1766	19	2007

Top author Zhang Y had the highest number of papers with 79 while Wang J is top author in terms of h-index. In terms of period, Wang L started the research career earliest in the field of cybersecurity in the year 2006 while Liu X started the research career in the recently, 2015. In terms of h-index, Wang J has highest value of h-index (n = 19) and lowest by Joshi A (n = 9).

Top institutions

In total, authors from 11277 institutions were responsible for the 20039 papers. Table 11 lists the top twenty institutions based on the number of papers.

Table 11 – Most productive global institutions

Sl.	Institutions	TP	Country	Type
1	University of Maryland	160	USA	Education
2	Carnegie Mellon University	158	USA	Education
3	University of California	142	USA	Education
4	Pacific Northwest National Laboratory	135	USA	R & D
5	George Mason University	133	USA	Education
6	Purdue University	106	USA	Education
7	Sandia National Laboratories	104	USA	R & D
8	Arizona State University	95	USA	Education
9	University of Oxford	95	England	Education
10	University of Texas At San Antonio	94	USA	Education
11	University of Johannesburg	86	South Africa	Education
12	University of Virginia	85	USA	Education
13	Iowa State University	84	USA	Education
14	Old Dominion University	79	USA	Education
15	Pennsylvania State University	78	USA	Education
16	Norwegian University of Science and Technology	73	Norway	Education
17	University of Arizona	73	USA	Education
18	Tsinghua University	72	China	Education
19	Masaryk University	68	Czech Republic	Education
20	New York University	68	USA	Education

With 160 publications, University of Maryland had the greatest level of influence. This university has the Maryland Cybersecurity Center which produces research output in the area of cryptography, privacy, programming-language and software security, empirical security, hardware security, network security, behavioural and economic of cybersecurity. Followed by, Carnegie Mellon University, and University of California with 158 and 142 papers respectively. Of the most productive institutions, only two institutions (Pacific Northwest National Laboratory and Sandia National Laboratories) are belonging to research and development while others are educational institutions. Both the R & D institutions are from USA. Of the top 20 institutions, majority of the institutions were from USA (75%) followed by China, Czech, England, Norway and South Africa, each one representation.

Top countries

A total of 115 countries have produced the 20039 global publications in the field of cybersecurity. Table 12 lists the top twenty countries with their rank. USA is the most productive country in the field of cybersecurity because variety of cyber-attacks had occurred in various sectors of energy, transportation, hospital and so on, followed by the United Kingdom, China, India and Australia. Figure 4 depicts the frequently collaborated countries which were grouped into three: (1) Red cluster contains 10 countries: USA, China, India, Canada, Israel, Korea, Qatar, Thailand, Finland and Pakistan. (2) Blue cluster is the largest network with 11 countries: Singapore, Australia, Japan, Spain, France, United Kingdom, Italy, Belgium, Greece, Austria and South Africa (3) Green cluster is the smallest network with 3 countries: Saudi Arabia, Portugal and Brazil. Among these three groups, red cluster was centre of the cooperation for the other groups.

Table 12 – Most productive countries

Country	TP	Rank
USA	7472	1
UK	1512	2
China	1269	3
India	1017	4
Australia	727	5
Italy	616	6
Germany	567	7
Canada	561	8
Japan	418	9
South Korea	403	10
Spain	392	11
France	384	12
Russia	383	13
South Africa	287	14
Saudi Arabia	262	15
Norway	261	16
Malaysia	244	17
Poland	232	18
Sweden	226	19
Ukraine	223	20

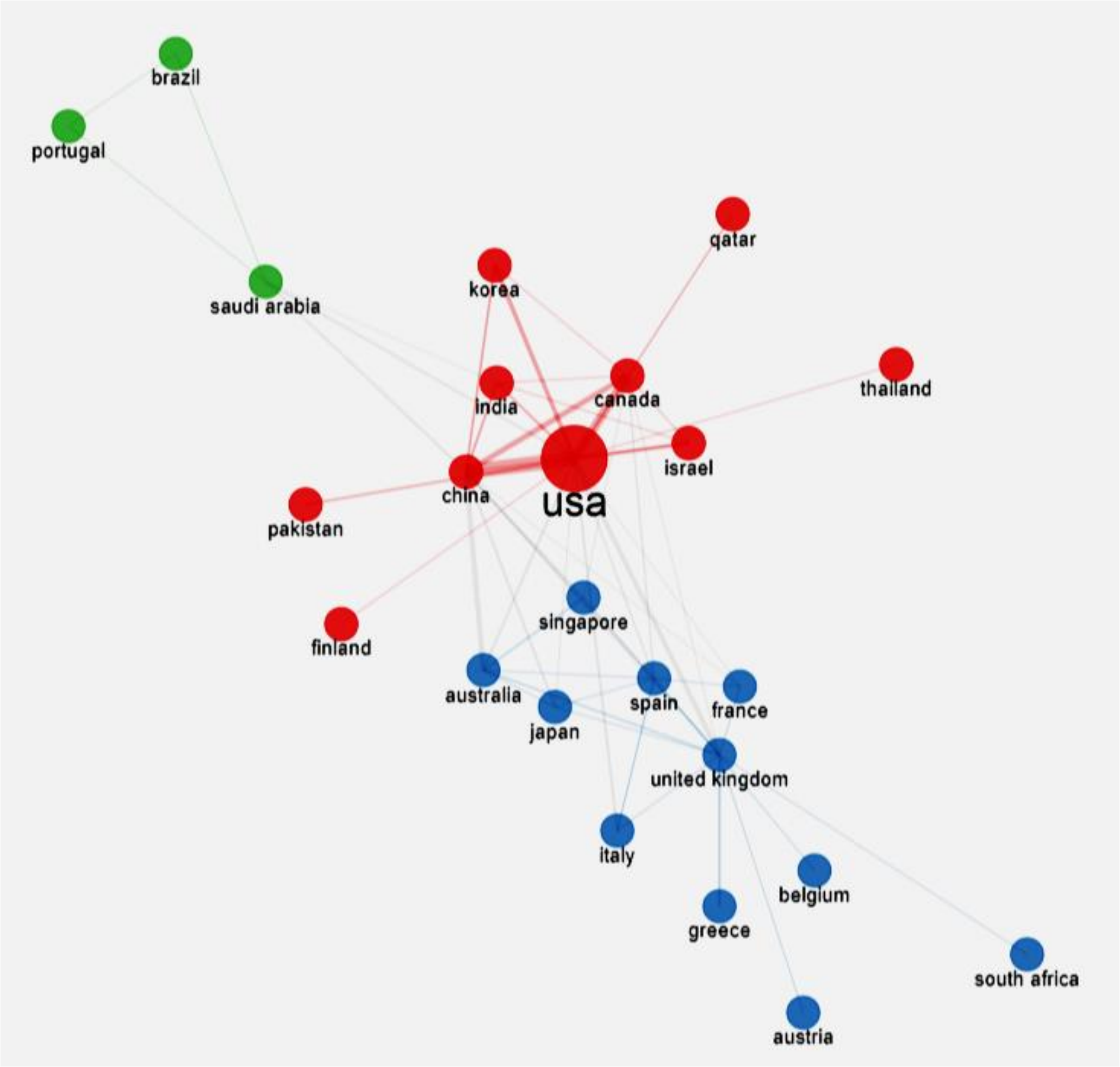


Figure 4 – International collaboration network
Parameter setting: Field – Countries, No. of Nodes – 50, Remove Isolated Nodes – Yes, Layout – Automatic, Clustering – Louvain, Normalization – Association, Minimum No. of Edges – 2, Repulsion Force – 0.1

Frequently used author keywords

Table 13 provides the list of high frequency author keywords in the field of cybersecurity during the study period. The keyword 'machine learning' is the most frequently used keyword apart from the search terms 'cyber security' and security by the authors. It is one of the emerging technologies that plays an essential role in cybersecurity for the detection and protection of cyber-attacks.

Table 13 – High frequency author keywords in global output

Sl.	Keywords	Occurrence
1	Cybersecurity	4309
2	Cyber security	3696
3	Security	1201
4	Machine learning	818
5	Cyber-security	620
6	Smart grid	560
7	Information security	506
8	Privacy	452
9	Internet of things	440
10	Intrusion detection	389
11	Iot	373
12	Deep learning	356
13	Anomaly detection	348
14	Malware	310
15	Blockchain	291
16	Cyber-physical systems	291
17	Scada	275
18	Cybercrime	272
19	Artificial intelligence	271
20	Big data	258
21	Network security	254
22	Cloud computing	252
23	Risk assessment	222
24	Risk management	221
25	Computer security	204

Cybersecurity plays a vital role in identifying, detecting and protecting against cyber-attacks, particularly detection of false data attacks, in the most emerging area in the electrical sector ‘smart grid’, which has been the topic of 560 papers. Among the top keywords, *anomaly detection* has been used to detect the cyber-attacks and *intrusion detection* helps to identify anomalies and prevent attacks. It can be seen that emerging technologies such as machine learning, deep learning, internet of things, big data, block chain, artificial intelligence and cloud computing are listed among the frequently used author keywords. Broader terms such as information security network security, and computer security were also listed among the frequently used keywords. Process control system ‘SCADA’ is also listed among the top keywords. Risk assessment that aims to identify various information assets which could be affected by cyber-attack was attracted by global researchers. The term ‘privacy’ has attracted by the researchers, which is more important in the context of data protection in the cyberspace.

Holistic taxonomy based classification of papers

Various organizations such as NIST CSRS (NIST Computer Security Resource Center), ETSI TC, European Union Cybersecurity Taxonomy, ACM (Association for Computing Machinery) and IEEE have proposed the cybersecurity taxonomy in different ways. Among these taxonomies, European Union cybersecurity taxonomy provides all perspectives on research in cybersecurity activities, entities by sector or industry, and technology and use cases that are well-defined. Application-centric, technology-centric, cyber-defense, cyber-awareness, cybersecurity implementation strategy and performance evaluation are all topics covered. To recognize the research gap in cybersecurity domains, it must determine the needs of cybersecurity in large/small scale organization, government and business based on these classifications. Table 14 details the three realms of cybersecurity research: research domains, sectors and technological and use. Domains of knowledge related to various aspects of cybersecurity are known as research domains. Because of the multifaceted nature of cybersecurity, such domains intended to cover a wide range of themes, including technological, education and legal issues. Technologies and use cases dimension has provided technologies (digital system) that are facing cybersecurity challenges in a variety of industries (or sectors).

Table 14 - EU Cybersecurity Taxonomy

Research domains	**Technological & Use Case**	**Sector**
Assurance, Audit, Certifications	Artificial Intelligence	Audio visual and Media
Cryptography	Big Data	Chemical
Data Security and Privacy	Block chain and Distributed Ledger Technology	Defence
Education and Training	Cloud, Edge and Virtualisation	Digital Service and Platforms
Human Aspects	Critical Infrastructure Protection	Energy
Identity Management	Protection of Public Spaces	Financial
Incident Handling and Digital Forensics	Disaster Resilience and Crisis Management	Food and Drink
Legal Aspects	Fight Against Crime and Terrorism	Government
Network and Distributed Systems	Border and External Security	Health
Security Management and Governance	Local / Wide Area Network and Surveillance	Manufacturing and Supply Chain
Security Measurements	Hardware Technology	Nuclear
Software and Hardware Security Engineering	High-performance Computing	Safety and Security
Steganography, Steganalysis and Watermarking	Human Machine Interface	Space
Theoretical Foundations	Industrial IOT and Control Systems	Telecomm Infrastructure
Trust Management and Accountability	Information Systems	Transportation
	Internet of Things, Embedded Systems and Pervasive System	
	Mobile Device	
	Operating Systems	
	Quantum Technologies	
	Robotics	
	Satellite System and Applications	
	Vehicular System	
	UAV	

(Source: https://www.cyberwiser.eu/news/jrc-proposal-european-cybersecurity-taxonomy)

Global researchers have emphasized the importance of Industrial IoT and control system in protecting SCADA (Supervisory Control and data Acquisition) and CPS (Cyber-Physical System) which are critical components of the Industrial Revolution 4.0. The sectors are offered to illustrate the need of assessing various cybersecurity requirements and challenges (from a human, legal, and ethical standpoint) in scenarios such as energy, transportation, and finance. Researchers in the field of cybersecurity have focused on security management and governance in the energy industry, taking into account Industrial IoT and Control Systems.

For the classification of publications, the following inclusion and exclusion criteria have been employed:

(1) Publications other than English have discarded.

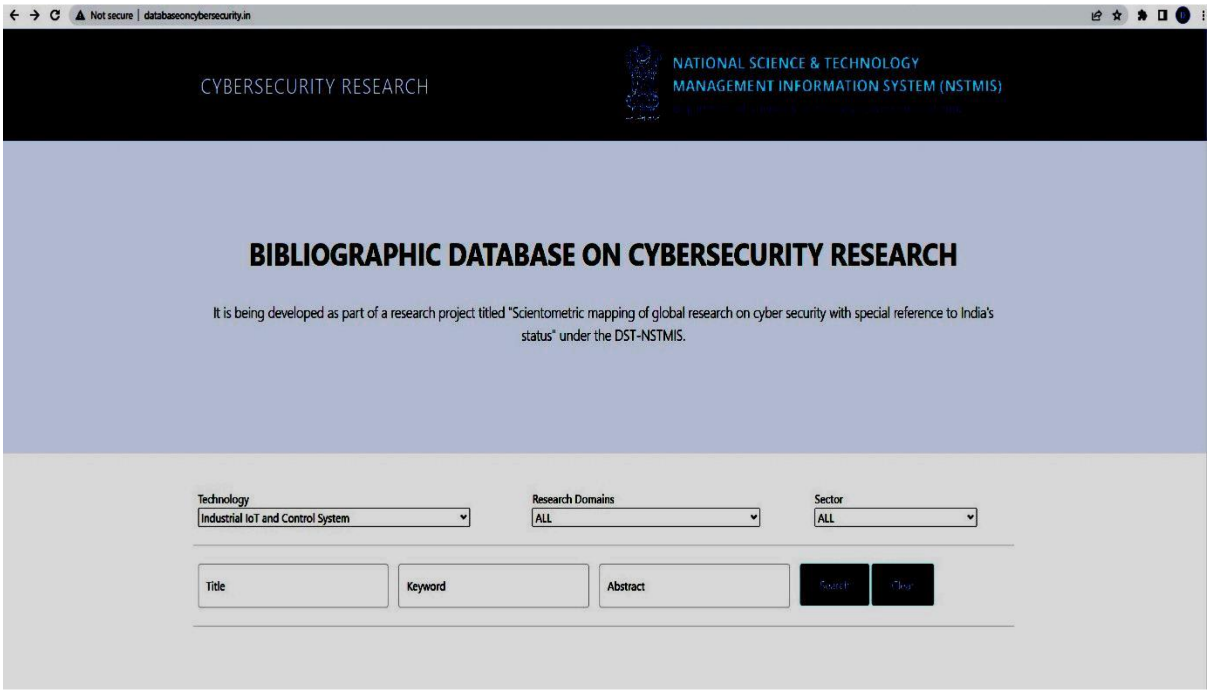
(2) Only articles and conference papers have been considered.

(3) Short reviews, workshop report, and guidelines have been discarded.

(4) Publications dealt with general concepts of cybersecurity have been discarded.

(5) Final dataset has been classified into three aspects: research domain, technology & use case, and sector.

The dataset contains classified publications has been uploaded in an exclusively developed web portal (http://databaseoncybersecurity.in) which is shown in the figure 5.



CYBERSECURITY RESEARCH

NATIONAL SCIENCE & TECHNOLOGY
MANAGEMENT INFORMATION SYSTEM (NSTMIS)

BIBLIOGRAPHIC DATABASE ON CYBERSECURITY RESEARCH

It is being developed as part of a research project titled "Scientometric mapping of global research on cyber security with special reference to India's status" under the DST-NSTMIS.

Technology: Industrial IoT and Control System

Research Domains: ALL

Sector: ALL

Figure 5 – Screenshot of newly created web portal, http://databaseoncybersecurity.in

From this web portal, researchers can be searched with a term in title, keyword and abstract as well as by selecting the topic under the three categories: technology, research domains and sector. Retrieved records can be downloaded in a spreadsheet for academic and/or research use only.

Technology & use case refers to the technical enablers that aid the development many industries and it has to do with cybersecurity research as well as technological components (Figure 6). Top five technological use cases are: IoT (industrial), artificial intelligence, hardware technology, cloud related technologies and critical infrastructure protection.

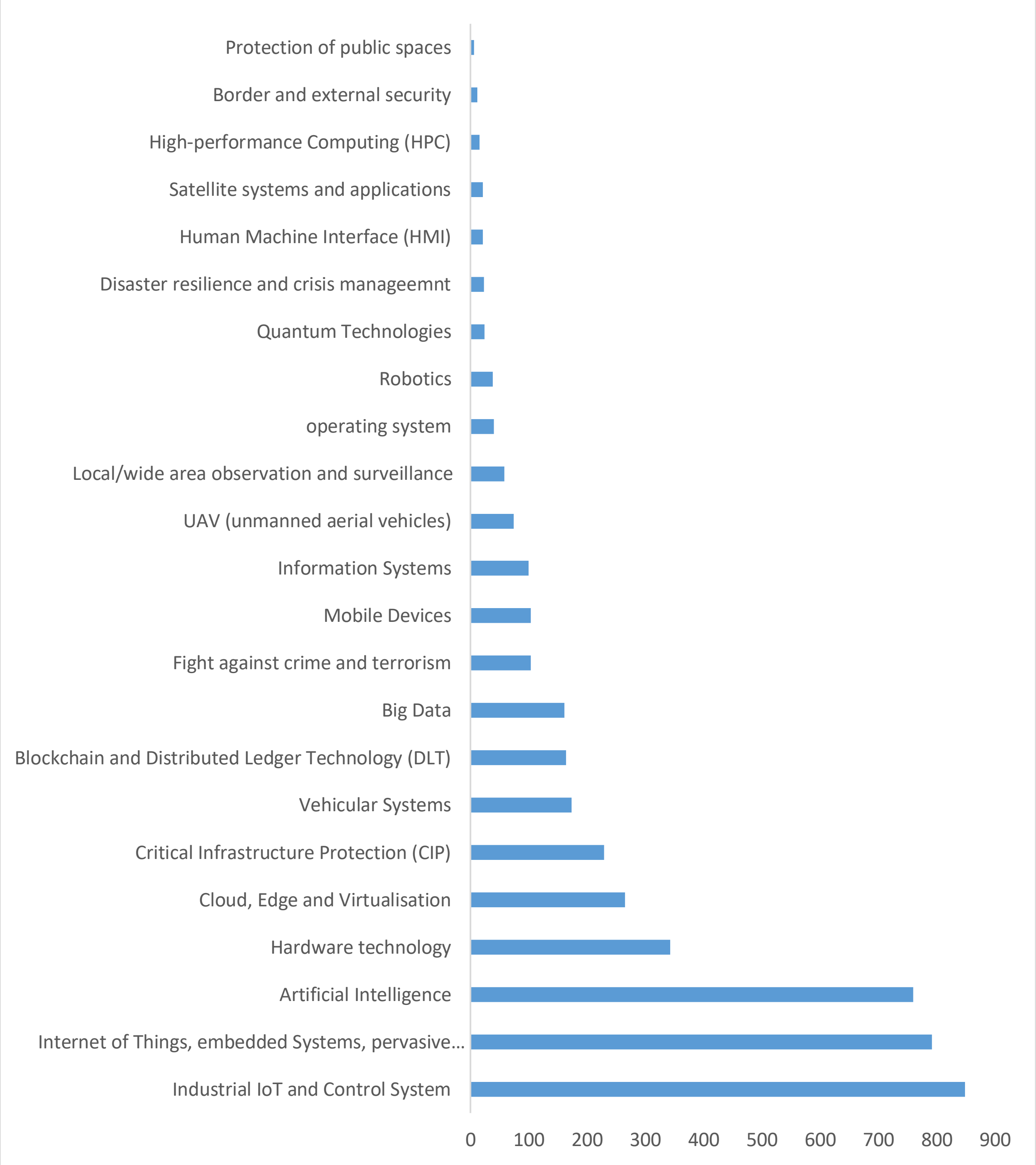


Figure 6 – Distribution of papers by technology use case

Sectors are offered to highlight the need of assessing various cybersecurity requirements and impediments (from a human, legal, and ethical standpoint) in scenarios such as energy, transportation, and finance (Figure 7). Top five sectors are: defence, energy, & digital service and platforms, government and health.

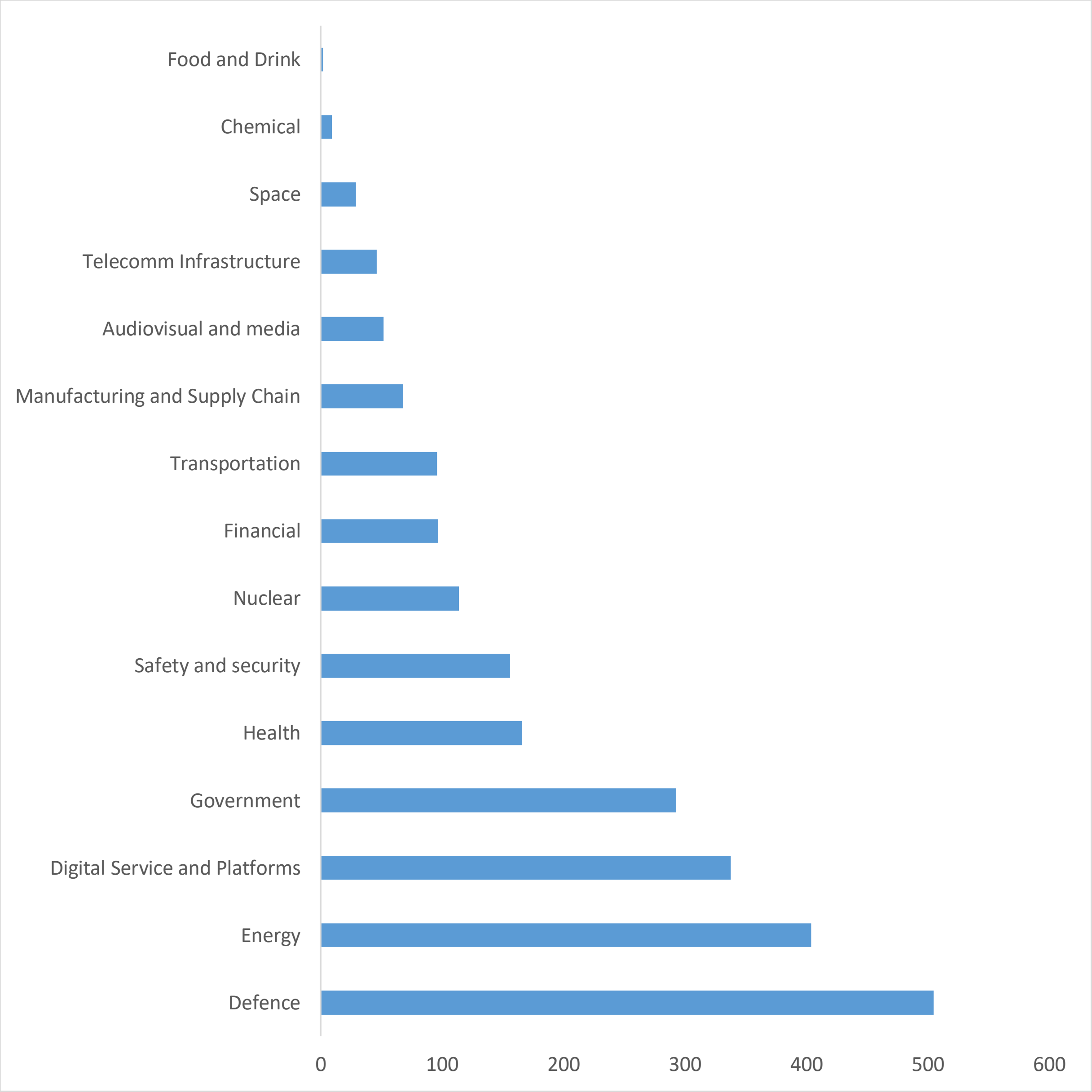


Figure 7 - Distribution of papers by sector

Research domains are areas of expertise that deal with various aspects of cybersecurity. Due to the multifaceted nature of cybersecurity, such domains are meant to encompass a wide variety of topics, including technological, educational, and legal issues (Figure 8). Top three research domains are education & training, security management and governance and software and hardware security engineering, human aspects, theoretical foundations.

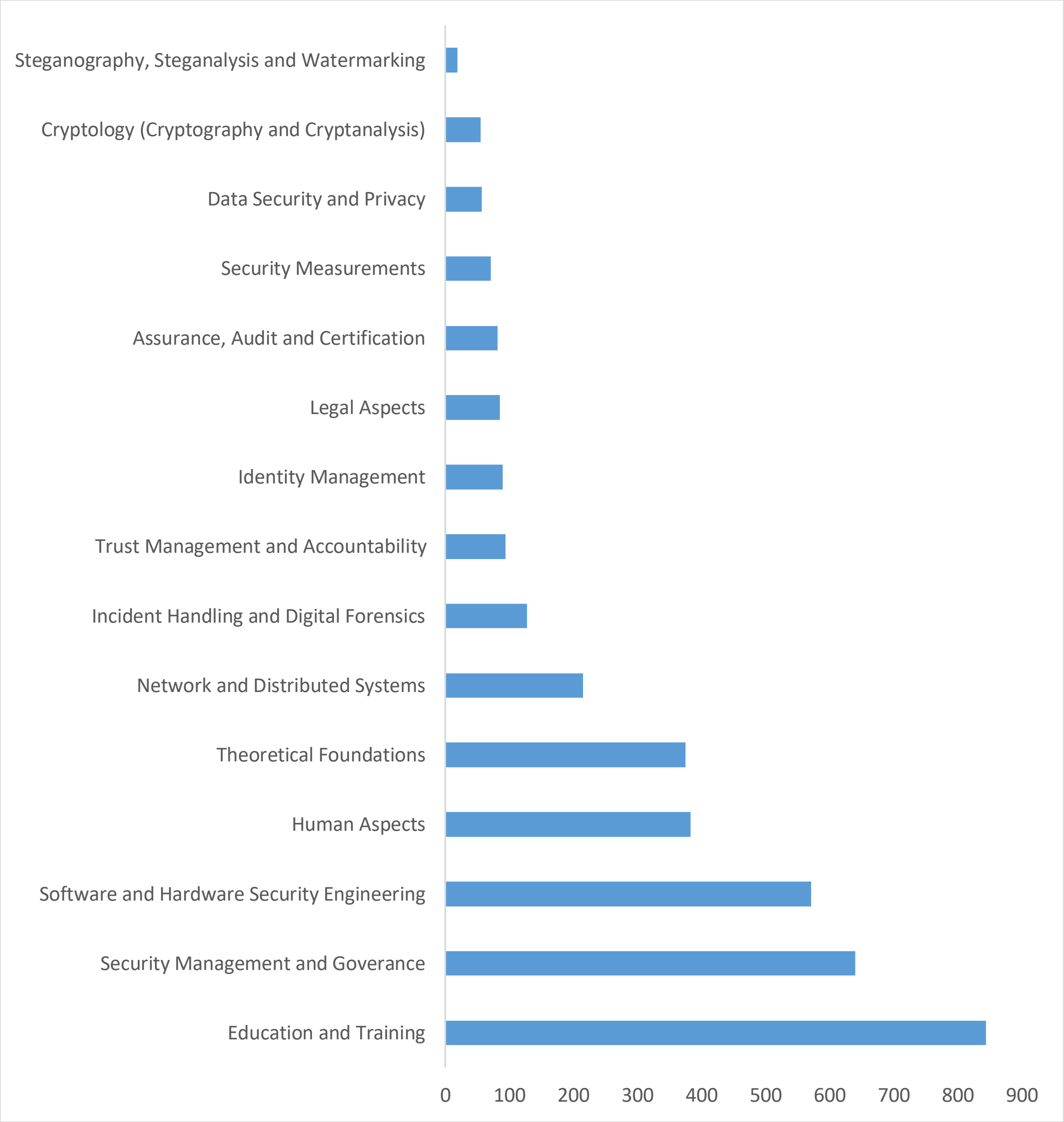


Figure 8 – Distribution of papers by research domain

INDIAN OUTPUT

Main information

Table 15 provides the information on document types, document content, individual authors, and authors' collaboration. Indian researchers have published 1108 papers in the field of cybersecurity in 577 different sources between 1999 and 2020. In terms of citation impact, Indian publications received an average of 5.66 citations per paper. A total of 3010 keywords have been appended by the authors in the 1108 papers with 2.69 keywords per paper. Most of the papers were contributed with co-author(s) (n = 1002) which denotes that strong collaboration exists among the Indian researchers in this field. There are 3305 authorships in the 1108 papers and the Collaboration Index (Elango and Rajendran 2012) is evaluated to 2.41, which indicates that the research team consists between two and three authors in this field.

Table 15 – Main information on Indian output

Description	Numbers
Timespan	1999-2020
Sources (Journals, Books, etc.)	577
Papers / publications	1108
Average citations per paper	5.66
Author's Keywords (DE)	2994
Authors	2512
Author Appearances	3305
Authors of single-authored papers	95
Authors of multi-authored papers	2417
Single-authored papers	106
Collaboration Index	2.41

Document type

The total Indian papers are classified into 8 documents types as shown in Table 16. Among the documents types, most papers are in the form of conference papers and articles, accounting to 85% together. Other document types make up a small percentage of the total.

Table 16 – Document type of Indian output

Sl.	Document Type	TP	Share
1	Conference paper	513	46.30
2	Article	437	39.44
3	Book chapter	81	7.31
4	Review article	30	2.71
5	Book	10	0.90
6	Editorial	9	0.81
7	Note	2	0.18
8	Book review	1	0.09
9	Unknown	25	2.26
	Total	1108	100

The preferred document types by the Indian researchers in the field of cybersecurity are depicted in the figure 9. The most popular document type is conference paper, followed by article. Others trail behind at a safe distance.

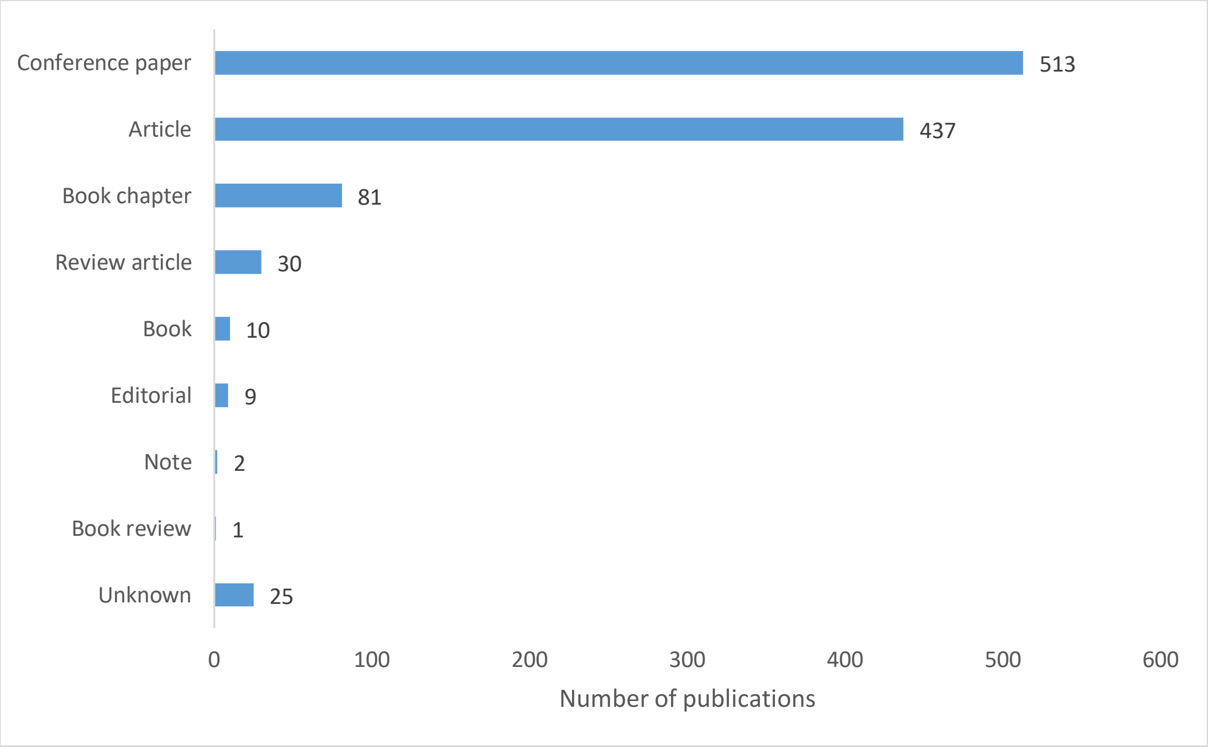


Figure 9 – Document types of Indian output

Yearly trend

Figure 10 depicts the annual growth of Indian papers from 1999 to 2020. During this period, there was a significant gap in publication, with no publication following the first publication in 1999 until 2004. In terms of the number and growth of publications, there was no discernible pattern. The year with the highest growth rate was 2005, while the years with the most publications was 2020.

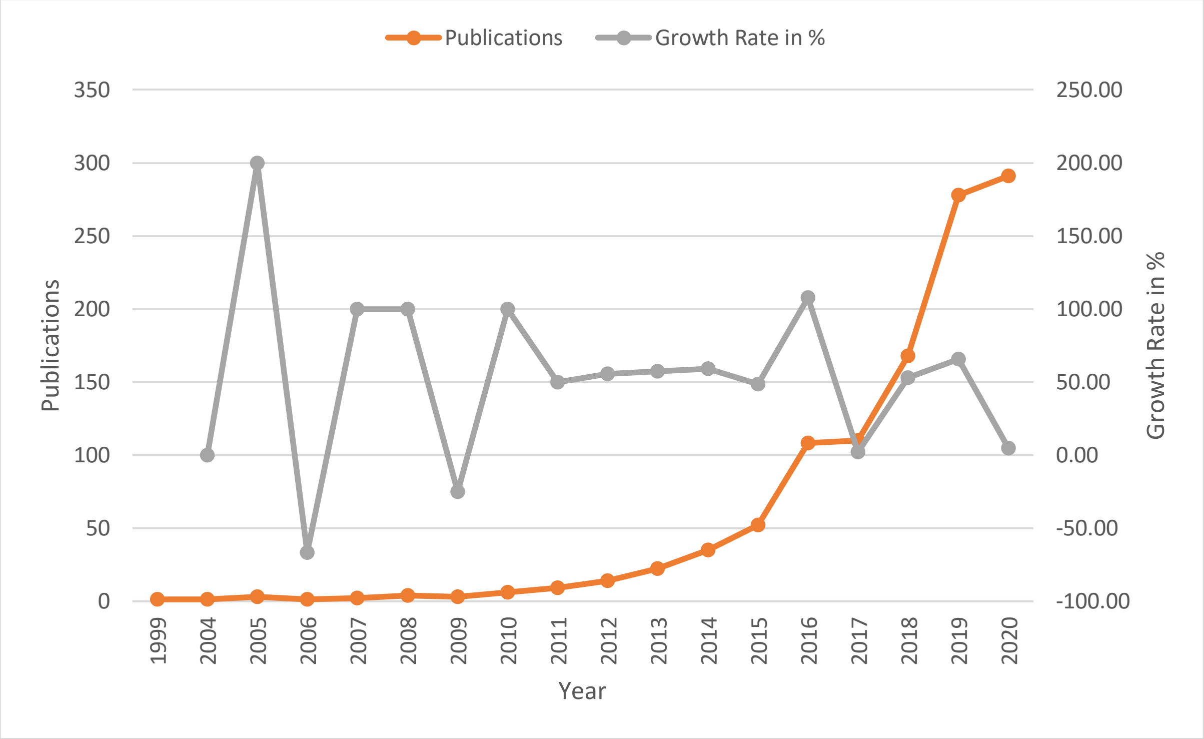


Figure 10 - Annual growth of Indian output

The historical distribution of the amount Indian literature can be used to depict and define the cybersecurity research trend at various phases of development. Hence, the entire time period has been divided into three stages (see Table 17) based on the number and growth of publications (Figure 10), with phase-1 (1999-2010) denoted as incubation, phase-2 (2011-2017) denoted as development and phase-3 (2018-2020) denoted as maturation (Briones-Bitar et al. 2020; Kastrin & Hristovski 2021):

1. Incubation phase (1999–2010) – This phase has 21 papers, accounting for only 1.87% of total papers with an average of nearly 3 papers per year and a 29.17% average growth. During this phase, cybersecurity research grew slowly, and amount of literature demonstrating the relevance of cybersecurity did not entice the Indian scholars.

2. Development phase (2011–2017) - During this phase, over 30% of total papers was published, with an average of 50 per year and a growth rate of 51.77%. The peak years (n = 115 & 118) in this period were 2016 & 2017. Indian researchers have been drawn to this topic since several cyber-attacks involving spyware and phishing have targeted the individual users and organizations, resulting in a surge in cybersecurity research.

3. Maturation phase (2018–2020) – This phase saw almost two-third (66.85%) of total papers, which is notably higher than the preceding two phases. During this phase, an average of 245 papers were produced per year, representing a 31.61% growth. The security challenges surrounding the Internet of Things, cloud computing and fog computing were studied by Indian researchers.

Table 17 – Distribution of Indian papers in phases

Phase	Period	TP	Share of 1108	TP / Year	CAGR in %
1 (Incubation)	1999-2010	21	1.87	3	29.17
2 (Development)	2011-2017	350	31.28	50	51.77
3 (Maturation)	2018-2020	737	66.85	249	31.61

Scattering of literature

The number of core sources in any particular field can be determined using Bradford’s law of scattering (Bradford 1934). Sources in a given field can be divided into three zones, each containing the same number of papers, a core zone containing the one-third of the total papers, middle zone containing the same number of papers but a greater number of sources, and tail zone containing the same number of papers but still a greater number of sources. The mathematical relationship between the number of sources in the middle zone to the core zone is a constant n and to the tail zone the relationship is n^2: 1: n: n^2. Accordingly, the relationship of sources in each zone (table 18) is observed to 31 : 31 X 5.83 : 31 X 2.02 which reveals that scattering of literature does not fit the Bradford's Law of scattering.

Table 18 – Sources and Bradford zones

Zone	No. of Sources	Share of 577	No. of Publications	Share of 1108
Core	31	5.37	369	33.30
Middle	181	31.37	374	33.75
Tail	365	63.26	365	33.96
Total	577	100	1108	100

Core sources

The total Indian papers were published in 577 different sources, including journals, conference proceedings, books etc. The top 31 sources published one-third of total papers and each of these top sources had at least 5 papers (see Table 19). In the realm of cybersecurity, Indian researchers prefer to publish their findings in journals and book series. Advances in Intelligent Systems and Computing, the most preferred source has 44 papers published. Because it covers the most recent advances in modern intelligence and computers in the areas of wireless security, trust management and artificial intelligence. Of the top 31 sources, 13 (~42%) are from India, demonstrating that Indian researchers prefer national sources. Almost half of the top 31 sources are journals and Scopus has cancelled coverage for eleven (35%) of the top 31 sources due to various publication concerns, emphasizing those Indian researchers should be cautious while choosing sources. Two journals indexed in Indian Citation Index were also among the top 31 sources.

Table 19 – Core sources in Indian output

Sources	Status	Type	Country	TP	Rank
Advances in Intelligent Systems and Computing	D	BS	Germany	44	1
International Journal of Innovative Technology and Exploring Engineering	D	J	India	28	2
International Journal of Recent Technology and Engineering	D	J	India	26	3
Communications in Computer and Information Science		BS	Germany	21	4
Lecture Notes in Networks and Systems		BS	Switzerland	17	5
IITM Journal of Management and IT	ICI	J	India	16	6
Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes/ in Bioinformatics)		BS	Germany	15	7
Lecture Notes in Electrical Engineering		BS	Germany	15	7
IEEE Access		J	United States	14	8
Journal of Advanced Research in Dynamical and Control Systems	D	J	United States	13	9
ACM International Conference Proceeding Series		C	United States	13	9
Procedia Computer Science		C	Netherlands	12	10

International Journal of Advanced Science and Technology	D	J	Australia	11	11
International Journal of Scientific and Technology Research	D	J	India	11	11
International Journal of Cyber Criminology		J	India	10	12
International Journal of Engineering and Advanced Technology	D	J	India	10	12
ICRITO 2020 - IEEE 8th International Conference on Reliability INFOCOM Technologies and Optimization (Trends and Future Directions)		C	India	9	13
International Journal of Advanced Research in Computer Science	ICI	J	India	7	14
International Journal of Advanced Trends in Computer Science and Engineering	D	J	India	7	14
IOP Conference Series: Materials Science and Engineering		C	UK	7	14
Lecture Notes on Data Engineering and Communication Technologies		BS	Germany	7	14
Computer Communications		J	Netherlands	6	15
International Journal of Pharmacy and Technology	D	J	India	6	15
Water and Energy International		J	India	6	15
2016 1st International Conference on Innovation and Challenges in Cyber Security, ICICCS 2016		C	USA	6	15
Advanced Sciences and Technologies for Security Applications		B	Germany	6	15
International Journal of Control Theory and Applications	D	J	India	6	15

2017 International Conference on Advances in Computing, Communications and Informatics, ICACCI 2017		C	USA	5	16
Detecting and Mitigating Robotic Cyber Security Risks		BS	USA	5	16
Handbook of Computer Networks and Cyber Security: Principles and Paradigms		B	Germany	5	16
Indian Journal of Science and Technology	D	J	India	5	16
B = Book; BS = Book Series; C = Conference; J = Journal					

Authorship Pattern

Table 20 shows that the majority of the papers (90.43%) were written with co-authors. Only 9.57% of the papers were contributed by single authors. Highest number of papers (36.19%) was contributed by two authors followed by three authors with 25.18%, four authors with 15.79%. The remaining 13.26% of the papers were contributed with five or more authors. Only two papers have been contributed with the highest number of authors, 16 and 25 respectively. In the field of cybersecurity, it appears that team research has surpassed solo research.

Table 20 – Authorship pattern of Indian output

No. of Authors	No. of Papers	Share
Single	106	9.57
Two	401	36.19
Three	279	25.18
Four	175	15.79
Five	93	8.39
Six	32	2.89
Seven	12	1.08
More than seven	10	0.90
Total	1108	100

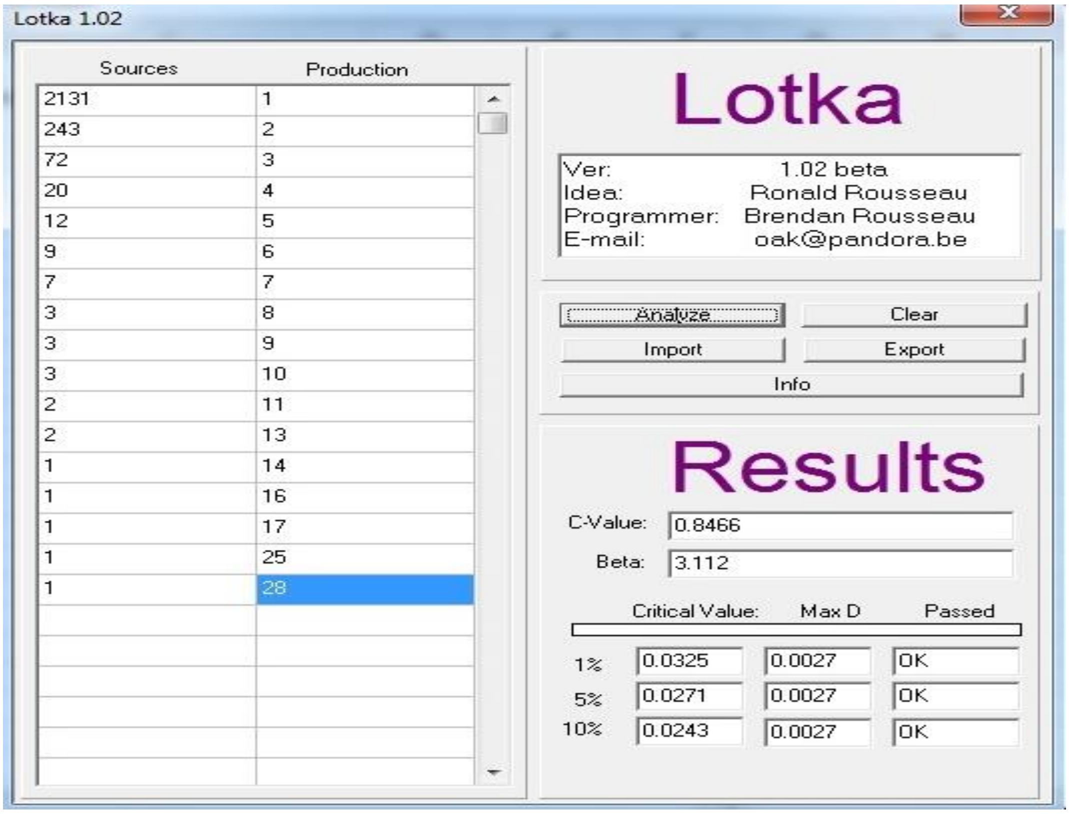
Author Productivity

The frequency distribution of author productivity in the field of cybersecurity is shown in table 21. The number of papers contributed by each author is shown in order of increasing output. According to the data, a total of 2512 unique authors contributed over the time period under consideration. Almost 85% of authors published one paper while remaining 15% of authors contributed more than one paper.

Table 21 – Author productivity in Indian output

No. of papers	No. of Authors	Share of 2512
1	2131	84.8
2	243	9.7
3	72	2.9
4	20	0.8
5	12	0.5
6	9	0.4
7	7	0.3
8	3	0.1
9	3	0.1
10	3	0.1
11	2	0.1
13	2	0.1
14	1	0
16	1	0
17	1	0
25	1	0
28	1	0

The applicability of Lotka’s law (Lotka 1926) was tested using a software program named Lotka developed by Rousseau & Rousseau (2000). Using the parameters from the table 21 (Papers column as Production and Authors column as Sources), the software generated C and n values of 0.846 and 3.112, respectively (see figure 11). Because the value of n (2.543) falls within Pao's (Pao 1985) range of values, i.e. 1.78 to 3.78, the author productivity follows Lotka's original distribution.



Sources	Production
2131	1
243	2
72	3
20	4
12	5
9	6
7	7
3	8
3	9
3	10
2	11
2	13
1	14
1	16
1	17
1	25
1	28

Ver: 1.02 beta
Idea: Ronald Rousseau
Programmer: Brendan Rousseau
E-mail: oak@pandora.be

C-Value: 0.8466
Beta: 3.112

	Critical Value:	Max D	Passed
1%	0.0325	0.0027	OK
5%	0.0271	0.0027	OK
10%	0.0243	0.0027	OK

Figure 11 – Testing the Lotka's law

Top authors

A total of 2512 authors, including international authors, participated in the Indian cybersecurity research. The most of the authors (85%) have contributed one paper, while the remaining authors produced between 2 and 28 papers. There are a few authors with the same name. For example, there are different authors for the label "kumar s": kumar sushil and kumar sanjeev. Hence, the manual verification has been done. Top 27 authors had at least 5 papers (see Table 22) and among these top authors, seven originated from the United States (n = 3), Japan (n = 2), Australia (n = 1) and Finland (n =1) that make a high impact in this field.

Table 22 – Top Indian authors

Author	Institutions	Country	h_index	TC	TP	Period
Soman KP	Amrita Vishwa Vidyapeetham	India	11	778	28	2016-20
Vinayakumar R	Amrita Vishwa Vidyapeetham	India	12	802	27	2017-20
Poornachandran P	Amrita Vishwa Vidyapeetham	India	11	626	19	2016-19
Dutt V	IIT Mandi	India	4	44	13	2015-20
Ustun TS	National Institute Of Advanced Industrial Science And Technology, Japan	Japan	9	189	13	2018-20
Hussain SMS	National Institute Of Advanced Industrial Science And Technology, Japan	Japan	8	173	12	2018-20
Mehtre BM	Institute For Development And Research In Banking Technology	India	6	124	11	2013-19
Farooq SM	Yogi Vemana University	India	6	135	10	2018-20
Gupta BB	NIT Kurukshetra		7	235	10	2015-20
Aggarwal P	Carnegie Mellon University, United States	USA	3	30	8	2015-20
Janet B	NIT Tiruchirappalli	India	2	20	8	2015-20
Alazab M	Charles Darwin University	Australia	5	407	7	2019-20
Joshi A	University of Maryland Baltimore County	USA	5	86	7	2016-20
Marqbool Z	IIT Mandi	India	3	16	7	2015-20
Gonzalez C	Carnegie Mellon University, United States	USA	3	26	6	2016-20
Khanna K	IIT Delhi	India	4	78	6	2016-20
Pammi VSC	University of Allahabad	India	3	16	6	2015-20
Panigrahi B K	IIT Delhi	India	4	78	6	2016-20
Sriram S	Amrita Vishwa Vidyapeetham	India	3	28	6	2019-20
Achuthan K	Amrita Vishwa Vidyapeetham	India	3	42	5	2014-20
Akarsh S	Amrita Vishwa Vidyapeetham	India	3	21	5	2019-20

Bhardwaj A	University of Petroleum & Energy Studies	India	2	13	5	2019-20
Jain AK	NIT Kurukshetra	India	5	192	5	2016-19
Khan RA	Babasaheb Bhimrao Ambedkar University	India	2	18	5	2018-20
Shukla SK	IIT Kanpur	India	3	20	5	2015-19
Starck J	ABB Oy - Finland	Finland	2	9	5	2013-18
Sukumara T	ABB GISLP Ltd	India	2	9	5	2013-18

To explore the research collaboration between these 27 lead authors, we generated a co-authorship network diagram (see Figure 12). There are six clusters groups among the top 27 authors (table 23). Cluster 1 (red colour) is the largest network among the others, with six authors contributing to it. Because the authors Soman K, Vinayakumar R, Poornachandran P are affiliated with the same institution, Amrita Vishwa Vidyapeetham, their collaboration is the strongest. This improves the citation impact also. Of the clusters, six authors (Achutan K, Bhardwaj A, Janet B, Khan RA, Mehtre BM and Shukla SK) did not collaborate with other lead authors.

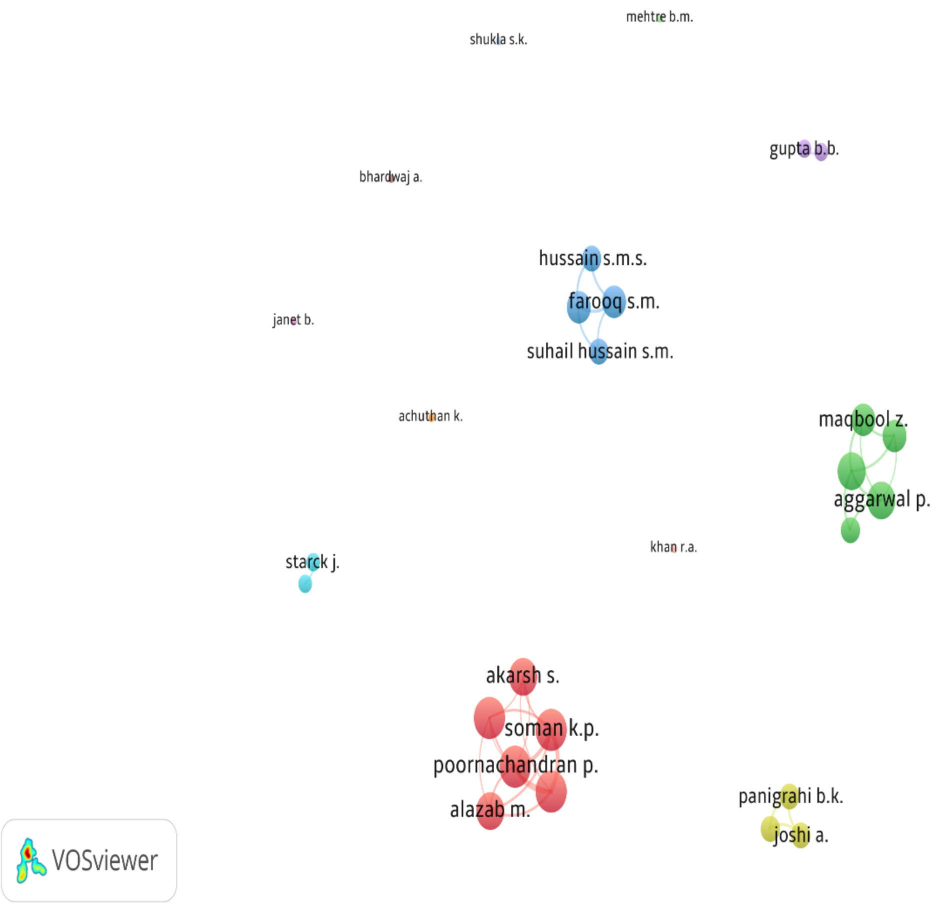


Figure 12 – Co-authorship network of Indian authors

Table 23 – Authors and clusters

Cluster	Authors
Red	Akarsh S
	Alazab M
	Poornachandran P
	Soman KP
	Sriram S
	Vinayakumar R
Green	Aggarwal P
	Dutt V
	Gonzalez C
	Marqbool Z
	Pammi VSC
Blue	Farooq SM
	Hussain SMS
	Ustun TS
Metallic Gold	Joshi A

	Khanna K
	Panigrahi B K
Purple	Gupta BB
	Jain AK
Light Blue	Starck J
	Sukumara T

Indian authors were collaborated with 53 countries during the study period and top six frequently collaborated countries are listed in the table 24. Four of the G7 countries are listed among major partner countries which indicate that India has more frequent partnerships with scientists from G7 countries.

Table 24 – Top collaborating countries for Indian authors

Country	TP
USA	96
Australia	19
Saudi Arabia	19
UK	19
Japan	15
Canada	12

Figure 13 depicts the international collaboration map between Indian researchers and international co-authors from 53 countries.

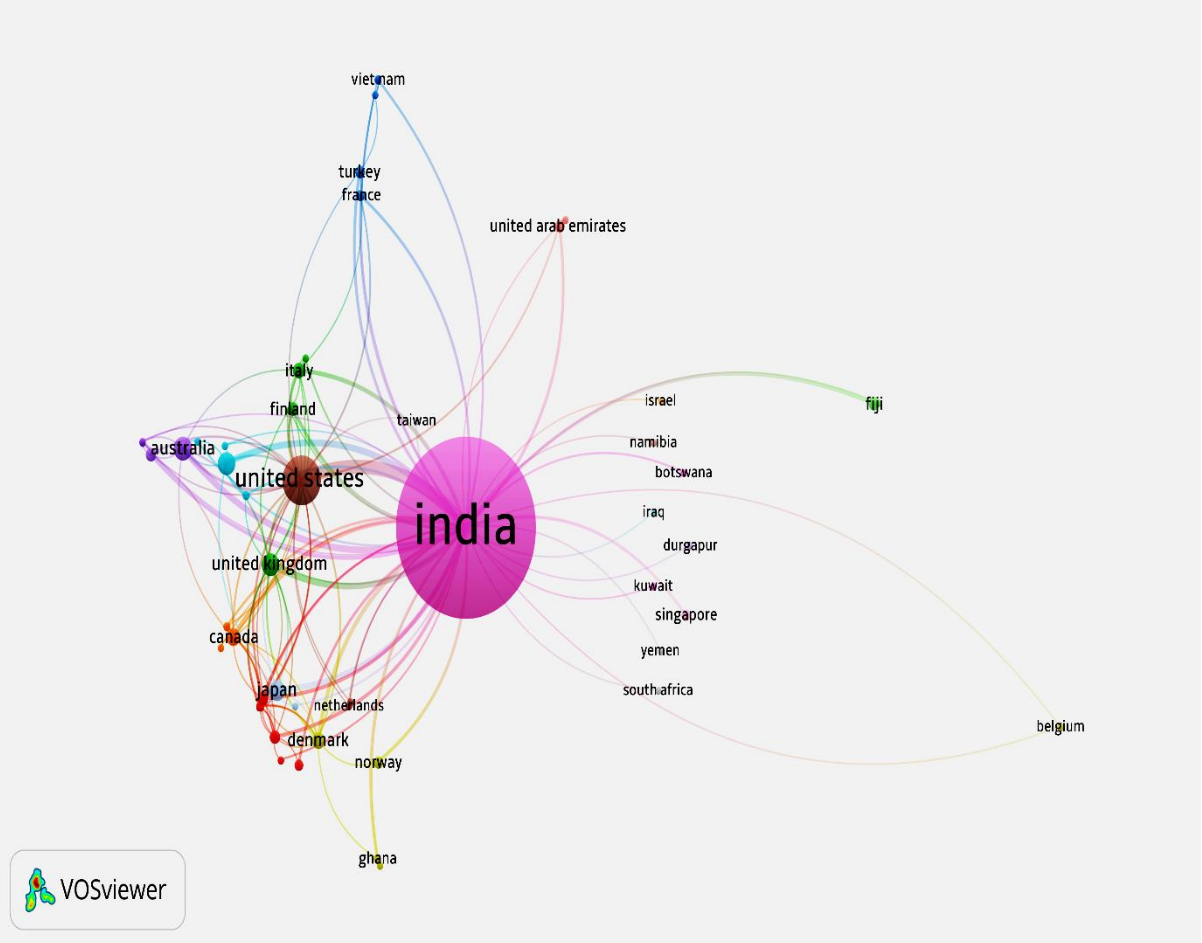


Figure 13 – Network of collaborating countries for Indian authors

Top institutes

Authors from 859 institutions contributed to the Indian cybersecurity publications (including international ones). Thirty-four institutions have published at least 7 papers to their credit (Table 25). Institutions with campuses in multiple locations are consolidated into a single entity. Amrita Vishwa Vidyapeetham, for example, has seven campuses spread across four states. Vellore Institute of Technology, on the other hand, has three campuses spread over two states. Seventeen (50%) of the top 34 institutions are private, indicating that private institutions have greater research capability in this topic. Amrita Vishwa Vidyapeetham has topped with 58 papers. The reason could be attributed to centres of excellence in cybersecurity (https://www.amrita.edu/research/centers) namely, Amrita Center for Cybersecurity Systems and Networks in Amritapuri campus and TIFAC-Core in Cyber Security in Coimbatore campus. Multinational companies, Tata Consultancy Services India and ABB Global Industries & Services Pvt. Ltd. are also named among the top institutions, because they offer a variety of cybersecurity services, Cybersecurity Implementation Services and Cyber Defense Suite. Four international institutions are also among the top institutions. Four international institutions (Fukushima Renewable Energy Institute of Japan, Carnegie Mellon University of USA, Charles Darwin University of Australia, and Cincinnati Children's Hospital Medical Center of USA) are also listed among the top institutions which denotes that Indian researchers have frequent partnership with these institutions.

Table 25 – Top Indian institutions

Affiliations	**Institution type**	**TP**	**Rank**
Amrita Vishwa Vidyapeetham	Non-government	58	1
Amity University	Non-government	36	2
VIT University	Non-government	35	3
IIT Delhi	Government	23	4
SRM Institute of Science and Technology	Non-government	19	5
University of Petroleum and Energy Studies	Non-government	19	5
KL University	Non-government	17	6
Institute for Development and Research in Banking Technology	Government	15	7
IIT Mandi	Government	14	8
National Institute of Technology, Kurukshetra	Government	14	8
Thapar Institute of Engineering and Technology	Non-government	13	9
University of Hyderabad	Government	13	9
Sathyabama Institute of Science and Technology	Non-government	12	10

Fukushima Renewable Energy Institute*	Japan	12	10
Tata Consultancy Services India**	Non-government	12	10
ABB GISL LTD**	Non-government	11	11
IIT Kharagpur	Government	10	12
Delhi Technological University	Government	9	13
Carnegie Mellon University*	USA	9	13
Symbiosis International	Non-government	9	13
Yogi Vemana University	Government	9	13
Birla Institute of Technology and Science, Pilani	Non-government	8	14
Anna University	Government	8	14
Jamia Millia Islamia	Government	8	14
Kalasalingam University	Non-government	8	14
IIT Kanpur	Government	8	14
NMIMS	Non-government	8	14
KIIT University	Non-government	7	15
Charles Darwin University*	Australia	7	15
SASTRA University	Non-government	7	15
Veermata Jijabai Technological Institute	Non-government	7	15
Cincinnati Children's Hospital Medical Center*	USA	7	15
Defence Institute of Advanced Technology	Government	7	15
NIT Jalandar	Government	7	15
*Foreign institution **Multinational company			

Citation analysis

Alarmingly, more than 40% of papers did not receive a single citation from the date of publication to the date of access (table 26). Almost 60% of papers received a citation between 1 and 279. However, majority of the papers (37%) received the citations between 1 and 5. Only a meagre amount of papers (21%) received more than 5 citations.

Table 26 – Citation pattern of Indian output

Citations	No. of papers	Share of 1108
0	460	41.52
1-5	415	37.45
6-10	104	9.39
11-50	106	9.57
51-100	14	1.26
100+	9	0.81
Total	1108	100.00

A total of 6130 citations have been received by Indian papers from its time of publication up to the date of access (table 27). Average citations per paper is 5.66 during the study period, papers published in the years 2008, 2014 – 2018 received the higher citations than the average.

Table 27 - Annual Indian output and citation impact

Year	TP	TC	CPP
1999	1	10	10.00
2004	1	0	0.00
2005	3	0	0.00
2006	1	3	3.00
2007	2	4	2.00
2008	4	24	6.00
2009	3	6	2.00
2010	6	4	0.67
2011	9	26	2.89
2012	14	54	3.86
2013	22	93	4.23
2014	35	290	8.29
2015	52	619	11.90
2016	108	771	7.14
2017	110	857	7.79
2018	168	1146	6.82
2019	278	1430	5.14
2020	291	793	2.73
1999-2010	21	51	2.43
2011-2017	350	2710	7.74
2018-2020	737	3369	4.57

When it comes to three phases, papers published in development phase received more citations compared to other two phases. Even though the maturation phase is very recent, the papers published in the phase received some impact.

Top 23 papers are termed as the highly- cited Indian papers in this topic (table 28) and have received a total of 2212 citations (36% of all citations received by the total 1108 papers) with an average of 96 citations per paper which is almost 17 times higher than the average (5.66). Fourteen of the 23 highly cited papers were articles, five were conference papers, and four were reviews. Similarly, most of the top-cited papers have been published as articles (Elango et al. draft). Eleven papers were collaborated with international co-authors and remaining twelve have published with national collaborators. All highly cited papers were published in the journals publishing from other countries.

Table 28 – Most influential Indian papers

Title	Year	Source	DT	Country	TC	Rank
Deep learning approach for intelligent intrusion detection system	2019	IEEE Access	Article	India, Australia, UK	279	1
Applying convolutional neural network for network intrusion detection	2017	2017 International conference on advances in computing, communications and informatics, ICACCI 2017	Conference Paper	India	170	2
Ddos attacks in cloud computing: issues, taxonomy, and future directions	2017	Computer Communications	Review	India, Italy, Australia	129	3
Internet of things forensics: recent advances, taxonomy, requirements, and open challenges	2019	Future Generation Computer Systems	Article	South Korea, Malaysia, India	119	4
A survey of the applications of text mining in financial domain	2016	Knowledge-Based Systems	Article	India	113	5
A survey towards an integration of big data analytics to big insights for value-creation	2018	Information Processing and Management	Article	India	108	6

Impact of covid-19 pandemic on information management research and practice: transforming education, work and life	2020	International Journal of Information Management	Article	UK, Denmark, USA, India	108	6
8A c9ybersecurity framework to identify malicious edge device in fog computing and cloud-of-things environments	2018	Computers and Security	Article	India, China	103	7
A review of integration, control, communication and metering (iccm) of renewable energy based smart grid	2014	Renewable and Sustainable Energy Reviews	Review	India, UK	101	8
Blockchain: future of financial and cyber security	2016	Proceedings of the 2016 2nd International conference on contemporary computing and informatics, IC3I 2016	Conference Paper	India	99	9
Performance analysis of smart metering for smart grid: an overview	2015	Renewable and Sustainable Energy Reviews	Review	India	96	10
A review and comparative analysis of various encryption algorithms	2015	International Journal of Security and its Applications	Article	India	83	11
Robust intelligent malware detection using deep learning	2019	IEEE Access	Article	India, Australia	78	12
A novel approach to protect against phishing attacks at client side using auto-updated white-list	2016	Eurasip Journal on Information Security	Article	India	77	13
Dual watermarking framework for privacy protection and content authentication of multimedia	2019	Future Generation Computer Systems	Article	India, Egypt, South Korea	72	14
A detailed analysis of cicids2017 dataset for designing intrusion detection systems	2018	International Journal of Engineering and Technology(UAE)	Article	India	70	15
Lightweight classification of iot malware based on image recognition	2018	Proceedings - international computer software and applications conference	Conference Paper	India, UK	69	16
A survey on intrusion detection systems and honeypot based proactive security mechanisms in vanets and vanet cloud	2018	Vehicular Communications	Review	India	63	17
Joint-transformation-based detection of false data injection attacks in smart grid	2018	IEEE Transactions on Industrial Informatics	Article	India, USA	58	18
A framework for fast and efficient cyber security	2016	Procedia Computer Science	Conference Paper	India	56	19

network intrusion detection using apache spark						
Botnet detection via mining of traffic flow characteristics	2016	Computers and Electrical Engineering	Article	India	54	20
Technical aspects of cyber kill chain	2015	Communications in Computer and Information Science	Conference Paper	India	54	20
Survey on collaborative smart drones and internet of things for improving smartness of smart cities	2019	IEEE Access	Article	Yemen, USA, India, Saudi Arabia	53	21

Research themes and topics

Keyword analysis can help researchers in determining a research domain's thematic trend and focus (Su et al. 2010). Author keyword analysis, in particular, lays the groundwork for analysing cybersecurity research trends (Dhawan et al. 2021). Table 29 lists the most often used author keywords over the three time periods. During the incubation phase (1999-2010), Indian researchers focused on data-related issues. During the development phase (2011-2017), their attention has changed from development to application (smart grid) and integration of emerging technologies such as machine learning into cybersecurity. During the maturation phase (2018-2020), they focused on AI-related technologies and IoT applications.

Table 29 – Most frequent author keywords in three phases

1999-2010		**2011-2017**		**2018-2020**	
Words	**Freq.**	**Words**	**Freq.**	**Words**	**Freq.**
Cyber crime	2	Cyber security	100	Cyber security	226
Cyber security	2	Smart grid	24	Cybersecurity	92
Computer forensics	1	Security	17	Machine learning	85
Cryptography	1	Cybersecurity	14	Deep learning	55
Data fusion	1	Machine learning	14	Security	40
Data mining	1	Botnet	10	Malware	31
Data protection	1	Network security	10	Cybercrime	27
Data security	1	Malware	9	Iot	27

Digital evidence	1	Cloud computing	7	Intrusion detection	25
e-health	1	Cyber crime	10	Cyber-security	24
Electric power systems	1	Cyber-security	9	Blockchain	24
Gas industry	1	Cryptography	7	Artificial intelligence	23
Global cyber security	1	Penetration testing	6	Intrusion detection system	21
Information Technology	1	Phishing	8	Smart grid	20
Internationalization	1	Privacy	6	Phishing	19
IT (amendment) act 2008	1	Clustering	6	Internet of Things	18
Oil industry	1	Encryption	6	Big data	16
Principal Component Analysis	1			Cyber-attacks	15
Privacy	1			Cloud computing	14
Hacking	1			Cryptography	14
				Information security	14
				Network security	14

As the quantity of papers between 1999 and 2010 is insufficient, additional analyses such as word clouds, co-occurrence networks, trend topics, and factorial analysis were conducted for two phases, 2011-2017 and 2018-2020.

The word cloud makes it simple to understand the most important issues while also exploring the prominent research topics in the field of cybersecurity. The frequency of author keywords that appear in Indian publications is shown in Figures 14-15. Between 2011 and 2017, the researchers focused on “smart grid” and “security”. In the years 2018-2020, the focus has shifted to “machine learning” and “deep learning”. These are the most commonly used author keywords in bibliographic analysis, indicating that research trend is moving towards Industry 4.0.

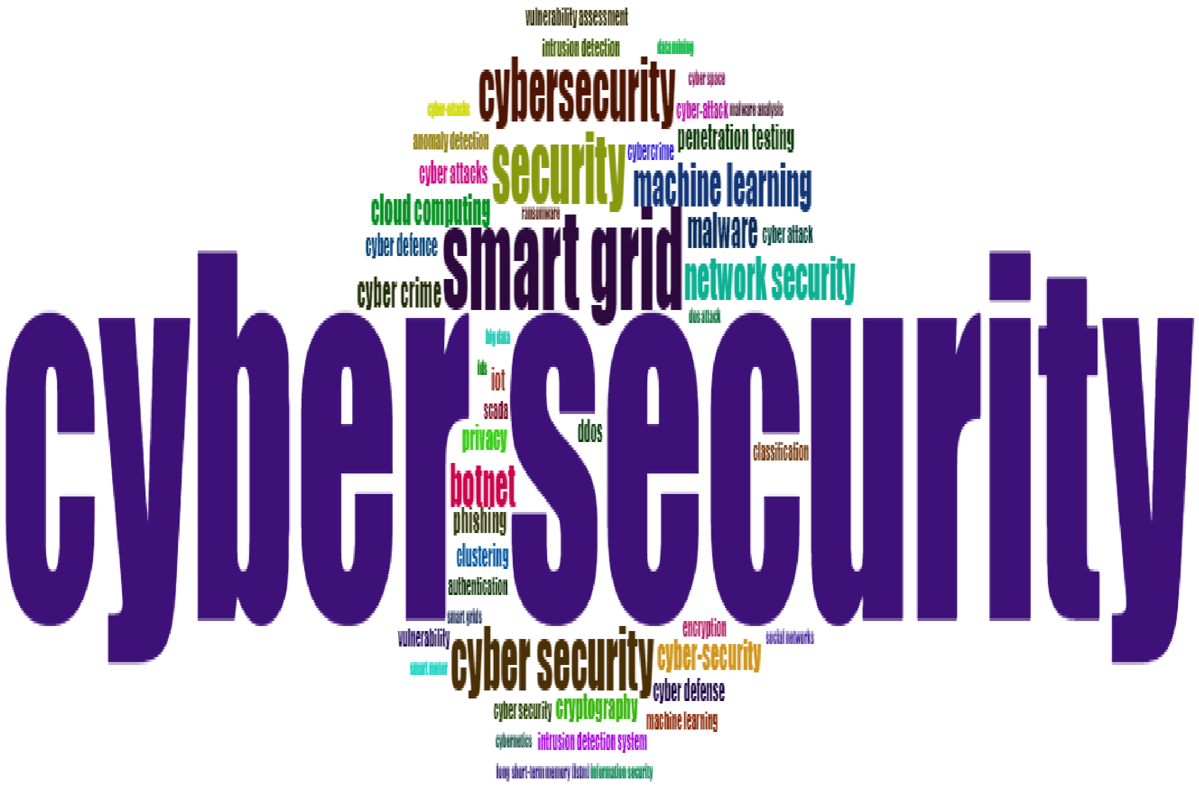


Figure 14 - Word cloud of author keywords (2011-2017): font size denotes the frequency
Parameter setting: Field – Author's keywords, No. of words – 50, Measure – Frequency

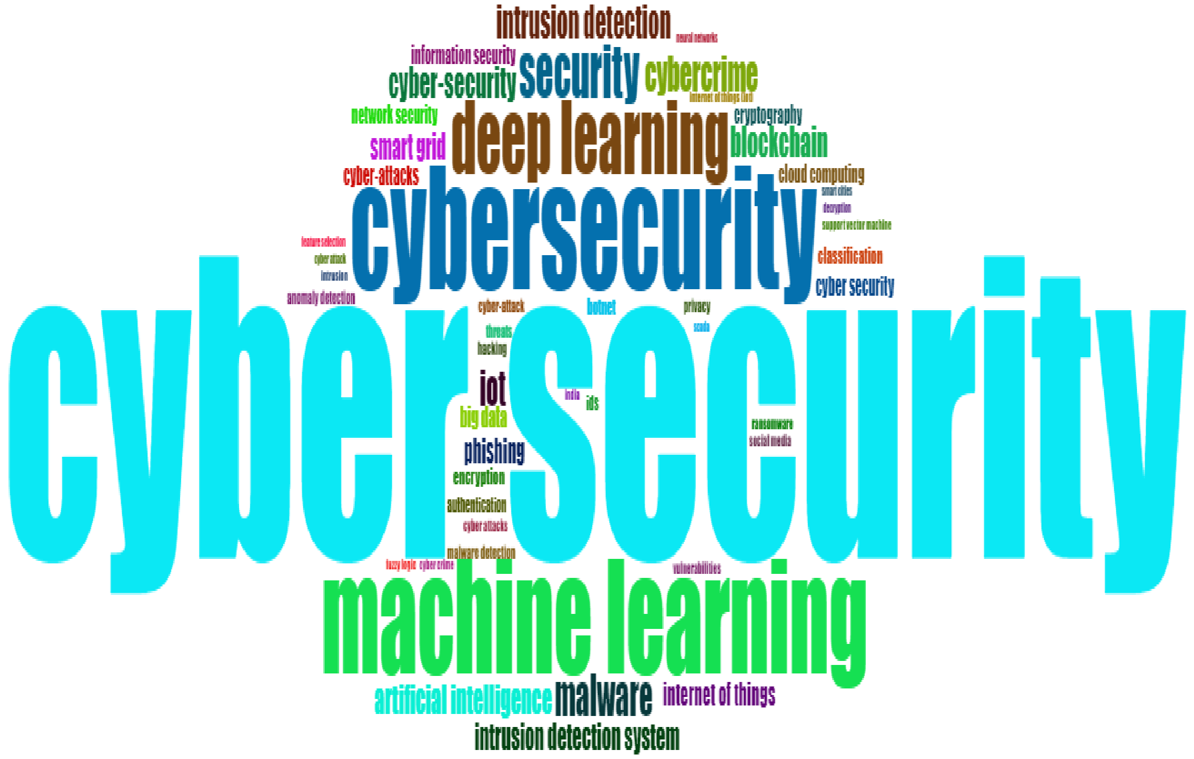


Figure 15 - Word cloud of author keywords (2018-2020): font size denotes the frequency
Parameter setting: Field – Author's keywords, No. of words – 50, Measure – Frequency

A co-occurrence network was created in VOSviewer to investigate the relationship between the most frequently used author keywords (see Table 29). It is text-analysis technique that incorporates a graphic representation of potential relationship between entities such as keywords, authors, organizations, etc. The relationships between the keywords (2011-2017) are depicted in Figure 16, which were divided into five clusters based on their proximity to one another.

The red cluster (security, smart grid, cloud computing) describes how smart grid is linked to cloud computing (Ezhilarasi 2016). Botnet detection (machine learning, botnet, clustering, network security) is referred to as blue cluster. The term is used to describe the process of detecting a botnet in network traffic using machine learning algorithms (Garg & Sharma 2017). Cyber-crime on phishing and malware was highlighted in the green cluster (cyber-crime, malware, phishing). The yellow cluster (privacy, cyber security, cybersecurity) represents privacy, which is an important feature of cybersecurity that prevents unauthorized access to confidential information. The violet cluster (encryption, cryptography) emphasises the cryptographic methods.

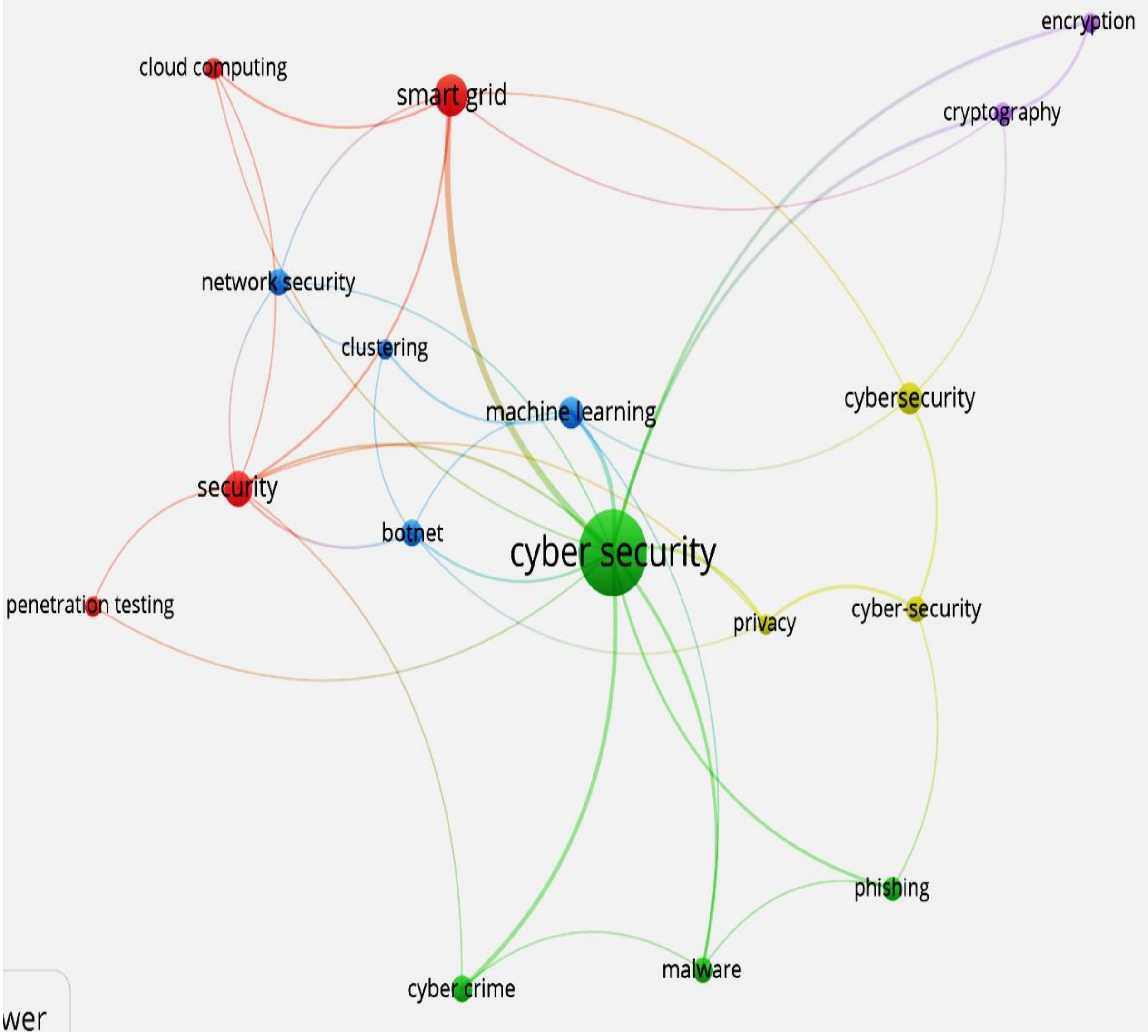


Figure 16 – Co-occurrence network of author keywords (2011-2017)
Parameter setting: Unit – Author's keywords, Counting method – Full, Min. frequency – 6, Bubble size refers to the number of publications, thickness of lines refers to link strength, color refers to cluster

Figure 17 depicts the associations between the keywords (2018-2020), which were sorted into four groups based on their proximity to one another.

The red cluster (cybersecurity, internet of things, security, smart grid, block chain) is focussed on using block chain methods to improve the privacy, authentication and encryption in internet of things cybersecurity. The green cluster (cybersecurity, cyber-security, machine learning, malware, cyber-attacks, artificial intelligence, intrusion detection system) implemented a machine algorithm in intrusion detection used to identify, detect and protect against cyber-attacks. The intrusion detection that employs deep learning to detect and identify unusual activities in network traffic was detailed in the yellow cluster (big data, intrusion detection, deep learning) (Maitham & Al-sultany 2021). The blue cluster (cryptography, information security, phishing) referred to the information security approaches used to protect the personal information from phishing.

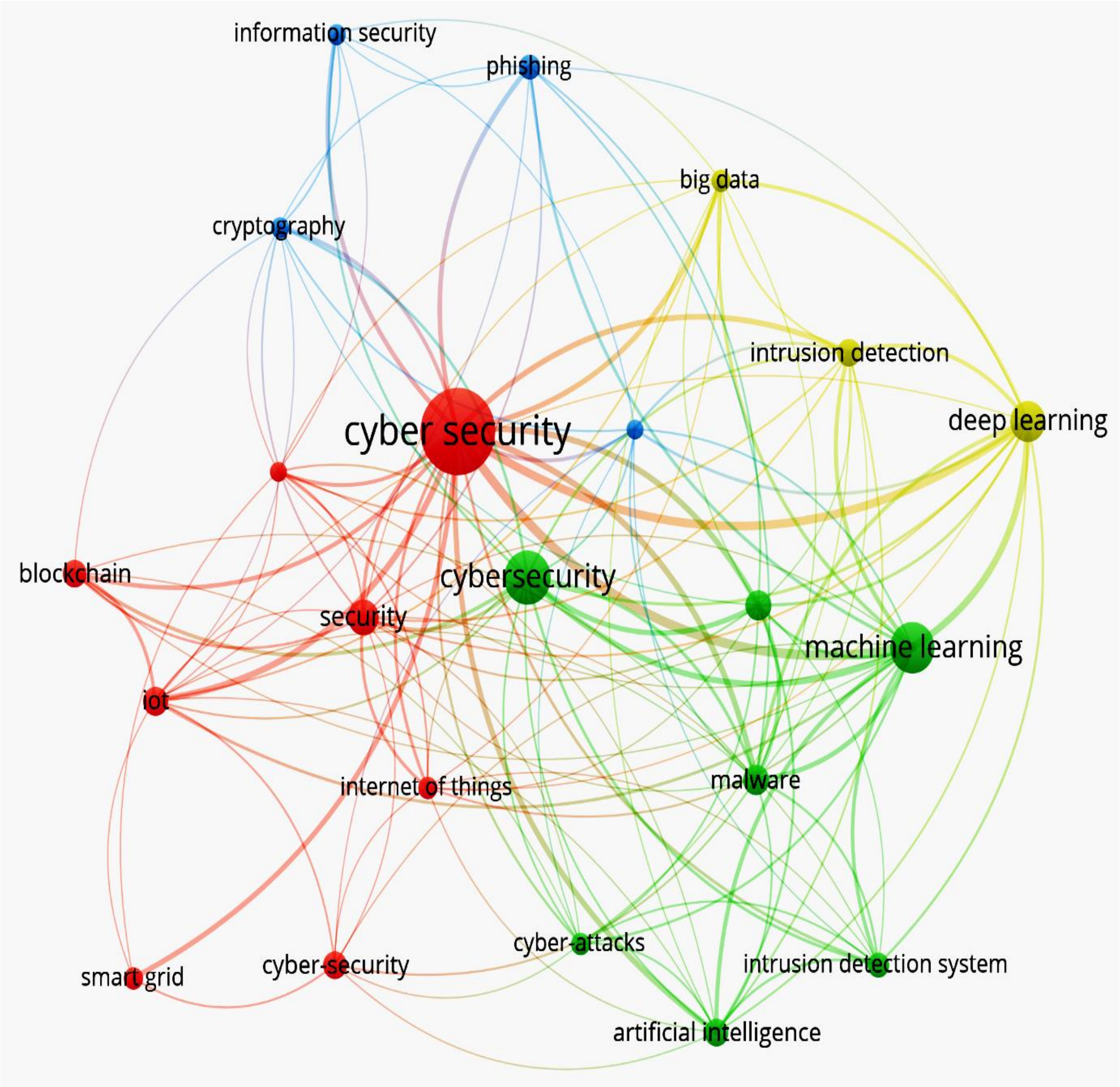


Figure 17 – Co-occurrence network of author keywords (2018-2020)
Parameter setting: Unit – Author's keywords, Counting method – Full, Min. frequency – 14, Bubble size refers to the number of publications, thickness of lines refers to link strength, color refers to cluster

From trend-topic analysis, changing topics / themes can be identified over the years in terms of highest frequent author keywords (see figures 18-19). In 2011-2017, the themes have been changed from network security (2014) to smart grids (2015) to cloud computing (2017). Similarly, in 2018-2020, the focus has changed from SCADA in the year 2018 to emerging technologies of deep learning, IOT, artificial intelligence, block chain and intrusion system in the year 2020. It is observed from the figure 19 that Indian researchers has focused on industrial revolution 4.0.

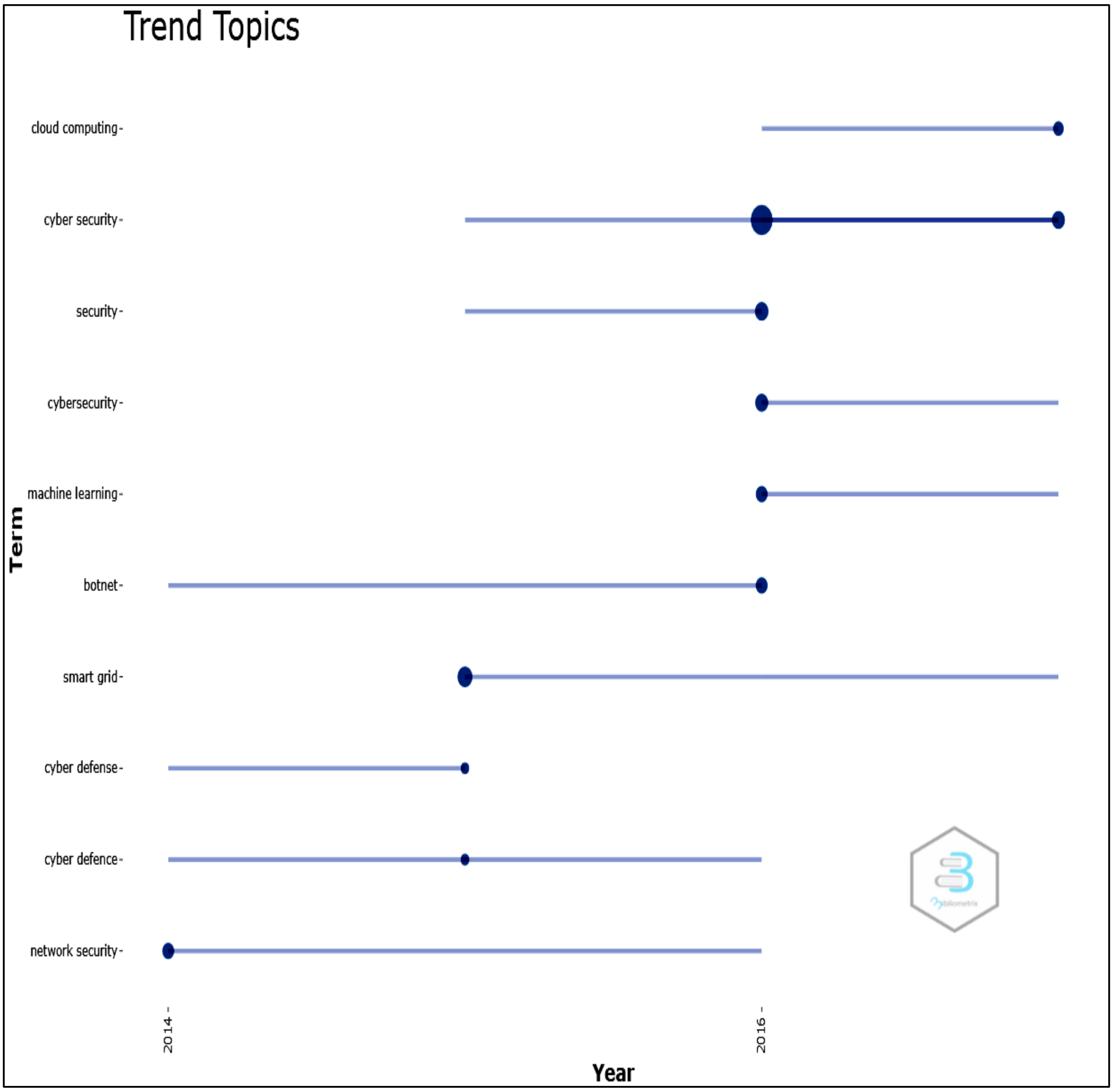


Figure 18 – Trend topics (2011-2017)
Parameter setting: Field – Author's keywords, Min. frequency – 5, No. of words per year – 5

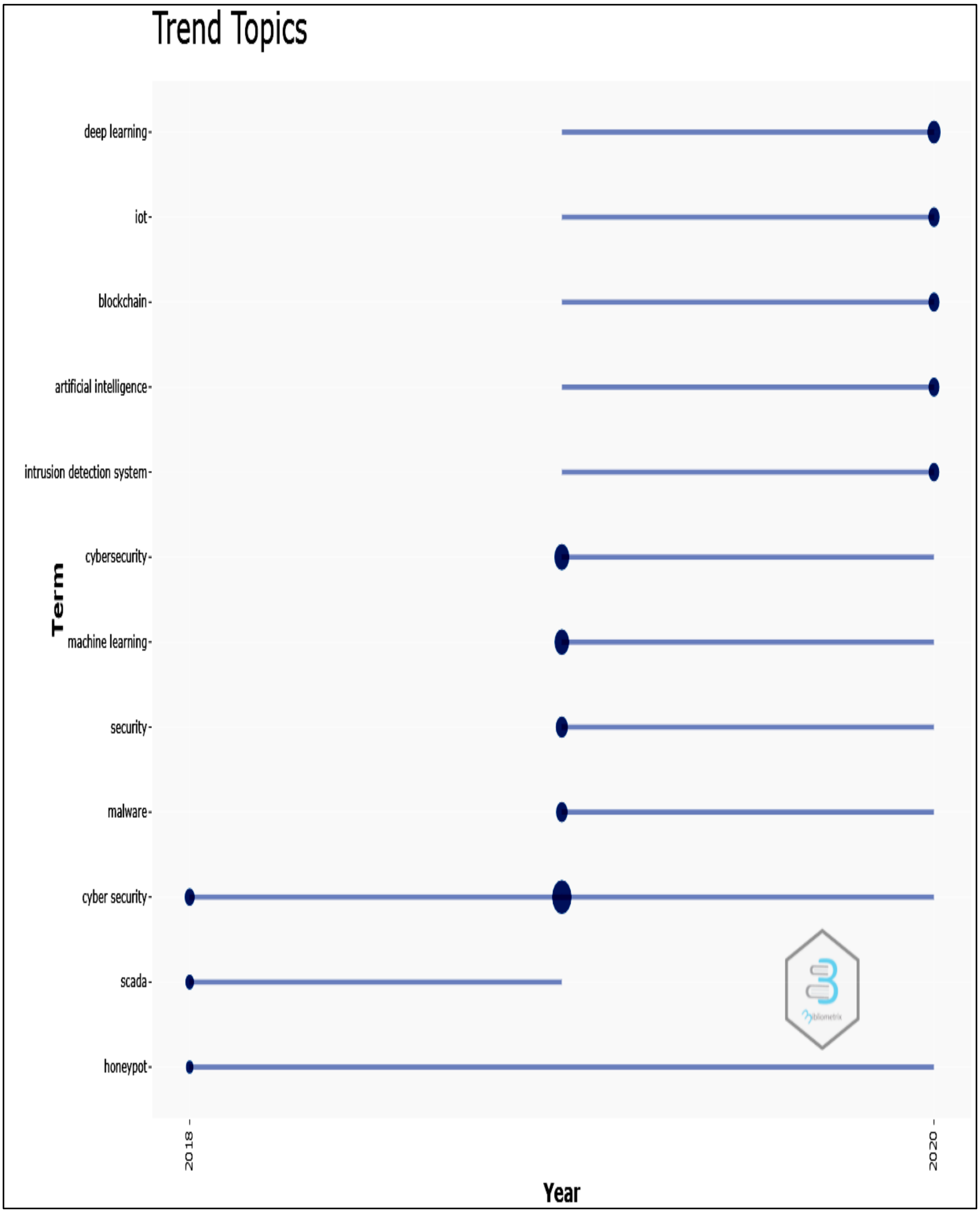


Figure 19 – Trend topics (2018-2020)

Parameter setting: Field – Author’s keywords, Min. frequency – 5, No. of words per year – 5

Multiple correspondence analysis was used to build conceptual structure maps of the author keywords (2011-2017 and 2018-2020) and the resulting clusters were shown in two-dimensional maps (see Figures 20-21): thus, two clusters were formed in the two phases. Multiple Correspondence Analysis (MCA) is a commonly used technique for the analysing categorical data with the purpose of reducing a large collection of data into smaller sets of components. The closer the dots on the graph represent each keyword, the more similar the distribution of keywords is, meaning that they co-occur more frequently in the publications. Keywords near the centre are of great interest to the research community, whereas keywords on the periphery are of low degree of relevance to other research topics (Shi et al. 2021; Mori et al. 2016).

In 2011-2017 (figure 20), contains two clusters: red cluster represents the cyber-attack detection and protection techniques including intrusion detection system, anomaly detection, static analyses and machine learning, and blue cluster mainly focuses on security testing (vulnerability testing) such as vulnerability assessment and penetration testing in information security that protect the confidential data from adversary (Goel & Mehtre 2015). Only few publications have discussed these topics because security testing was a disappearing technology.

In 2018-2020 (figure 21), blue cluster refers to encryption and decryption, both are important cybersecurity strategies that secure private and confidential data, and red cluster represents the defense mechanisms that must be used to detect cyber-attacks. Emerging technologies such as big data and AI are being used to detect the malware while maintaining effectiveness and efficiency.

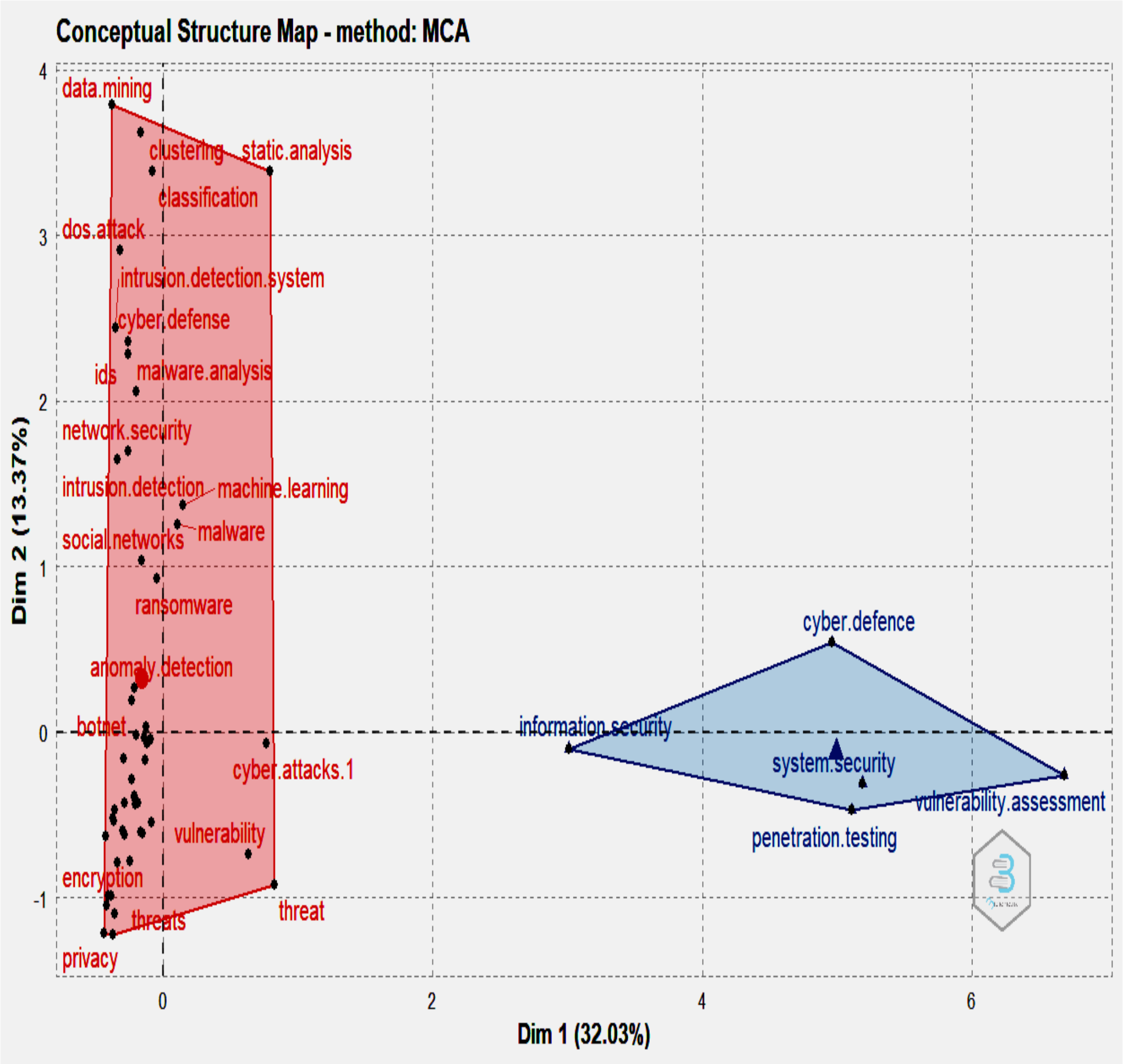


Figure 20 – Factor analysis of author keywords (2011-2017)

Parameter setting: Method – Multiple Correspondence Analysis, Field – Author's keywords, No. of terms – 50, No. of clusters - Auto

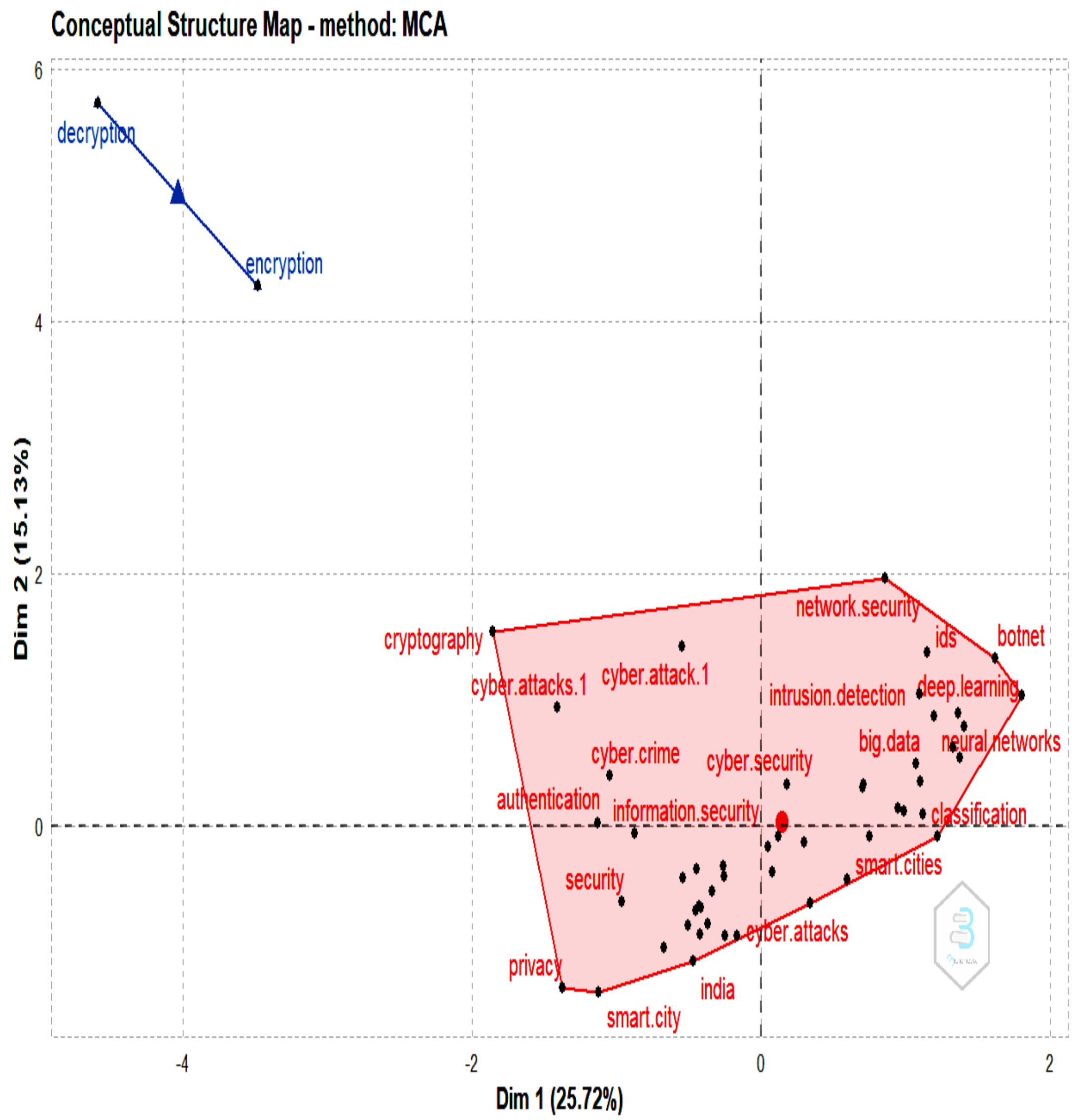


Figure 21 – Factor analysis of author keywords (2018-2020)
Parameter setting: Method – Multiple Correspondence Analysis, Field – Author's keywords, No. of terms – 50, No. of clusters - Auto

Chapter V

RESULTS AND DISCUSSION

The study brings out the scientometric mapping of global research output with special reference to India’s status.

Cyber Security is a distinct domain that pertains to and forms a part of new technologies such as artificial intelligence, internet of things, big data, advanced mobile computing, cloud computing, e-commerce and other emerging technologies. With this background, this study attempted to map the publication pattern in the field of cybersecurity. Bibliographic data from the two international indexing and abstracting databases viz. Web of Science and Scopus have been retrieved for the global mapping, and two Indian databases viz. Indian Citation Index and Indian Science Abstracts have been retrieved for the Indian mapping along with the international databases.

Global scenario

Fifty-four percent of total papers (refer Figure 2) were published in the recent three-years (2018-2020). It demonstrates that scholars have been focusing on the study topic “cybersecurity” in recent years.

Globally, more than 50% of total papers (n = 20039) have been published in the form of conference papers (refer Table 8). This result is in consistent with the findings of Fiala & Tutoky (2017) who found that 56% of all computer science related papers were published as proceedings papers and Dhawan et al. (2021) who also found that 57% of papers related to cybersecurity were published as conference papers.

Global researches were published in 5864 different sources with approximately 3.5 publication per source and of which, top three sources were book series and conference proceedings, together accounting to nearly 8.5% of total papers (refer Table 9). Forty-five percent of top 20 sources were conference proceedings and all the top 20 sources were being published in the three countries, viz. USA, Germany and UK.

A total 33156 authorships were observed in the 20039 total papers with a collaboration index of 1.99 (refer Table 7), which reveals that the research team comprises 2 in the field of cybersecurity. Single authored publications account for 25% of total papers.

In terms of citation impact, global research output had received an average of 7.41 citations per paper.

USA is the most productive country followed by UK, China, India and Australia (refer Table 12). This result is in consistent with Dhawan et al. (2021). A strong collaboration has been observed between USA, China and Canada (refer Figure 4). Following that, a duo collaboration has also been observed between USA and Korea, USA and Israel.

Most of the top 20 institutions are from USA (refer Table 11) which indicates that USA has paid a strong attention in this topic. All the most productive institutions are the types of academics except two institutions viz. Sandia National Laboratories and Pacific Northwest National Laboratory, which are research oriented and product development organizations.

Apart from the searching keywords cybersecurity, cyber security and related term security, machine learning (refer Table 13) has been listed at the global level, indicating that there are more possibilities of applying machine learning models to intrusion detection or detection and classification of attacks, to name a few (Martínez

Torres et al. 2019). In the most developing area in the electrical sector ‘smart grid’, which has been the frequent topic where cybersecurity plays a crucial role in identifying, detecting and protecting against cyber-attacks, particularly detection of false data attacks. Apart from these two topics, emerging technologies such as machine learning, deep learning, internet of things, big data, block chain, artificial intelligence and cloud computing are among the most commonly used author keywords.

Indian scenario

Considering the number of publications and its growth, the entire publication period has been divided into three phases: incubation (1999-2010), development (2011-2017) and maturation (2018-2020). Almost 67% of Indian publications (refer Figure 10) were published in the last three years (2018-2020) i.e. maturation phase which is very high compared to the global trend. It demonstrates that the study topic “cybersecurity” has attracted by the Indian researchers very recently.

Since 2004, Indian researchers have been publishing in this topic which is six years behind by global researchers who started their work in 1998. It reveals that it took six years to create awareness among the Indian researchers.

Forty-six percent of Indian publications (refer Figure 9) were published in the form of conference papers which is somewhat low compared to the global trend.

A total of 577 different sources published Indian research, with an average of two publications per source. Thirty-one core sources published almost one-third of the total papers (1108) accounting to 5% of total sources (refer Table 17) and sixteen (50%) of which were journals, demonstrating that Indian researchers prefer to publish their research findings in journals (refer Table 18). In contrast, conference proceedings accounted for 45% of the top 20 sources for global researchers. Forty-two percent of core sources are being published from India, suggesting that Indian researchers prefer to publish their research findings in their own country. It should be mentioned that Scopus stopped covering 11 sources due to publication concerns, implying that Indian researchers should focus on selecting the appropriate sources. Notably, 8 of 11 discontinued sources were published from India.

A total 3305 authorships were observed in the 1108 total papers with a collaboration index of 2.41 (refer Table 7), which reveals that the research team comprises between two and three, which is some-what higher than the global trend (1.99).

Only 10% of the Indian publications were single authored which is 40% low compared to the global trend, demonstrating that Indian researchers prefer collaborative works than solo.

Almost 85% of Indian authors have single publication in the field of cybersecurity. However, author productivity reveals that Lotka’s law is applicable to Indian contribution to the global cybersecurity research (refer Figure 11).

The study has identified top 27 Indian authors as most productive authors in the field of cybersecurity: those have at least 5 publications. Among these top authors, Soman KP from Amrita Vishwa Vidyapeetham was the top author in terms of number of publications where as Vinayakumar R was the top author in terms of highest h-index value, he was also from Amrita Vishwa Vidyapeetham. Six authors were from Amrita Vishwa Vidyapeetham whish reveals the institute’s focus on this topic.

Of the 27 top authors, seven authors from foreign countries have also been listed which shows that Indian researchers have strong collaboration with international counterparts.

Co-authorship network illustrates that there exist six research teams among the top 27 authors and mostly within institutional collaboration has been observed.

The top 34 institutions with at least 7 publications were ranked in terms of number of publications. Amrita Vishwa Vidyapeetham came out on top with 58 papers: the reason could be attributed that the institution has centres of excellence in cybersecurity (https://www.amrita.edu/research/centers) namely, Amrita Center for Cybersecurity Systems and Networks in Amritapuri campus and TIFAC-Core in Cyber Security in Coimbatore campus.

Four Indian Institutes of Technology (Delhi, Mandi, Kharagpur and Kanpur) were also list among the top 34 institutions. Among these four IITs, IIT Kharagpur has a specialized centre "Interdisciplinary Centre for Cyber Security and Cyber Defense of Critical Infrastructures". Similarly, IIT Mandi has "Applied Cognitive Science Lab".

If we group all the IITs (n = 72) and NITs (n = 62) into single cluster, these institutions will be in the top positions.

Half of the top 34 institutions belong to the category of non-government institutions, which shows their active participation in research and development in the field of cybersecurity. Two multinational companies (Tata Consultancy Services India and ABB GISL Ltd) were listed among the top 34 institutions. Four foreign institutions (two from USA and one each from Australia and Japan) were also listed among the top

34 institutions which shows that Indian researchers have frequent international collaborations in this topic.

USA is the most preferred country by Indian researchers to collaborate in this topic. A similar trend has been observed in many areas (Rajendran, Elango & Manickaraj 2014; Elango, Kozak & Rajendran 2019) including computer science (Parmar & Siwach 2016).

Citation analysis reveals that nearly 42% of total papers did not receive any citation while others have received citations between 1 and 279. Papers published in 2015 have received the highest citations per paper which shows that the year 2015 had the impactful research / publications by Indian researchers. Among the three phases, the publications in the development phase have highest citation impact than others.

Indian publications received an average of 5.66 citations per paper which is somewhat lower than the global output. It is primarily due to the publication of research findings in the non-standard sources.

Of the top 23 most-cited Indian publications, 14 (60%) were published as articles which shows that articles receive more impact than other types. This result is in consistent with previous studies (Dhawan et al. 2021; Rajendran et al. 2014). Notably, all the most-cited publications were published with co-authors which shows that co-authored publications received more citation impact than single authored ones.

During the incubation period, Indian researchers pay close attention to data-related issues (1999-2010). Then, during the development period, they concentrated on smart grid-related issues and began to integrate emerging technologies into cybersecurity (2011-2017). They have shifted their focus to AI-related technologies such as deep learning and machine learning, as well as IoT, during the maturation period (2018-2020). The changing focus of the Indian researchers has been depicted below for the easy understanding:

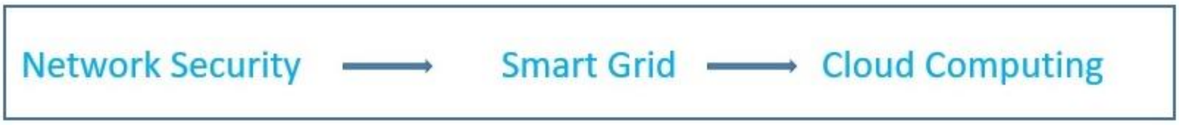


Figure - Indian researchers’ changing focus (2011-2017)

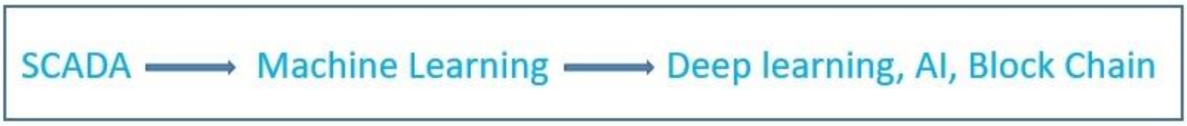


Figure - Indian researchers’ changing focus (2018-2020)

Chapter VI

FINDINGS AND RECOMMENDATIONS

In this chapter, major findings and the future research directions have been summarized.

Findings

- India was ranked 4th in terms of number of publications after the USA, the UK and China.
- Indian researchers have been publishing on this topic since 2004, six years after global researchers who began in 1998. It clearly evidences that it takes six years to raise awareness among the Indian researchers.
- In a span of seven years, only a meagre amount of papers (n = 1108) was published by Indian researchers.
- Fifty-four percent of the global research output has been published in the recent three years (2018-2020) where as it is 67% for the Indian output, which reveals that Indian researchers recognized this topic very recently.
- In India, 46% publications were in the form of conference papers, which is relatively low compared to the global trend.
- The value of the collaboration index (CI) for Indian publications is greater than the value for global output, indicating that Indian researchers prefer to do research in groups.
- In the global environment, the USA was the most productive country as well as the most preferred partner country for international collaboration for Indian researchers.

- In terms of citation impact, Indian publications received 5.66 citations per paper on average, which is somewhat lower than the global output. It is mostly due to research findings being published in the non-standard sources. For example, almost one-third of the most preferred sources by Indian researchers have been discontinued its coverage by Scopus.
- Indian researchers commonly collaborate within institutions, as seen by research teams made up of Indian researchers.
- The focus of Indian researchers has shifted from data-related issues to smart grid and emerging technologies, to AI-related technologies.
- Institutions having centers of excellence in cybersecurity produces more publications than others.

Recommendations

The results of this study will be useful for the policy decision makers as well as researchers in finding the most productive contributors (authors and institutions), most preferred sources, and hot themes.

Based on the analysis and interpretations, the following suggestions have been made:

- It is suggested to create centers of excellence in all the universities and higher education institutions in order to active participation in research in the field of cybersecurity.
- It is recommended to establish a national level center for creating the awareness about predatory journals as well as providing training in the high quality research.

- It is strongly recommended to concentrate on the application of emerging technologies into cybersecurity.
- More research must be carried out on the smart grid cybersecurity as well as in the energy sector.
- A database has been created with the classification of publications based on cybersecurity taxonomy: research domains, technology & use cases, and sector, which needs follow-up action.
- This study is based only on the publications indexed in Scopus, Web of Science, Indian Citation Index and Indian Science Abstracts. An exclusive study may be undertaken covering the non-indexed publications.
- Compared to other topics, there are only few publications by Indian researchers. In this regard, an awareness program is needed.

REFERENCES

Abbas, N. N., Ahmed, T., Shah, S. H. U., Omar, M., & Park, H. W. (2019). Investigating the applications of artificial intelligence in cyber security. *Scientometrics*, 121(2), 1189-1211.

Amoroso, E. G. (2007). *Cyber Security*. Silicon Press.

Bradford, S. C. (1934). Sources of information on specific subjects. *Engineering*, *137*, 85-86.

Briones-Bitar, J., Carrión-Mero, P., Montalván-Burbano, N., & Morante-Carballo, F. (2020). Rockfall research: a bibliometric analysis and future trends. *Geosciences*, 10(10), 403.

Chang, H. C. (2016). The synergy of scientometric analysis and knowledge mapping with topic models: modelling the development trajectories of information security and cyber-security research. *Journal of Information & Knowledge Management*, *15*(04), 1650044.

Christen, M., Gordijn, B., Weber, K., van de Poel, I., & Yaghmaei, E. (2017). A review of value-conflicts in cybersecurity: an assessment based on quantitative and qualitative literature analysis. *The ORBIT Journal*, *1*(1), 1-19.

Cojocaru, I., & Cojocaru, I. (2019). A bibliomeric analysis of cybersecurity research papers in Eastern Europe: Case study from the Republic of Moldova. In *Central and Eastern European eDem and eGov Days* (pp. 151-162).

Dhawan, S. M., Gupta, B. M., & Elango, B. (2021). Global cyber security research output (1998-2019): a scientometric analysis. *Science & Technology Libraries*, 40(2), 172-189.

Elango, B., Kozak, M., & Rajendran, P. (2019). Analysis of retractions in Indian science. *Scientometrics*, *119*(2), 1081-1094.

Elango, B., Matilda, S., Martina Jose Mary, M., & Arul Pugazhendhi, M. (**Draft**). Top-cited publications in cybersecurity: a bibliometric and content analysis.

Elango, B., & Rajendran, P. (2012). Authorship trends and collaboration pattern in the marine sciences literature: a scientometric study. *International Journal of Information Dissemination and Technology*, *2*(3), 166-169.

Furstenau, L. B., et al. (2020). 20 years of scientific evolution of cyber security: A science mapping. In *International Conference on Industrial Engineering and Operations Management* (pp. 314-325). IEOM Society International.

Jalali, M. S., Razak, S., Gordon, W., Perakslis, E., & Madnick, S. (2019). Health care and cybersecurity: bibliometric analysis of the literature. *Journal of medical Internet research*, *21*(2), e12644.

Kastrin, A., & Hristovski, D. (2021). Scientometric analysis and knowledge mapping of literature-based discovery (1986–2020). *Scientometrics*, *126*(2), 1415-1451.

Lotka, A. J. (1926). The frequency distribution of scientific productivity. *Journal of the Washington academy of sciences*, *16*(12), 317-323.

Mallick, P. K. (2019). *Research and development in cyber domain and Indian perspective*. Vivekananda International Foundation. New Delhi.

Martínez Torres, J., Iglesias Comesaña, C., & García-Nieto, P. J. (2019). Machine learning techniques applied to cybersecurity. *International Journal of Machine Learning and Cybernetics*, *10*(10), 2823-2836.

Pao, M. L. (1985). Lotka's law: A testing procedure. *Information Processing and Management*, *21*, 305–320.

Parmar, S., & Siwach, A. (2016). Indian Research output in Computer Science during 2004-2013: a bibliometric analysis. *International Journal of Digital Library Services*, *6*(2), 20-31.

Parvin, S., Sadoughi, F., Karimi, A., Mohammadi, M., & Aminpour, F. (2019). Information security from a scientometric perspective. *Webology*, *16*(1), 196-209.

Rahima, N., Othmanb, Z., Hamidc, F. Z., & Yeop, O. (2020). Cyber security and the higher education literature: A bibliometric analysis. *International Journal of Innovation, Creativity and Change*, *12*(12), 852-870.

Rai, S., Singh, K., & Varma, A. K. (2019). Global research trend on cyber security: A scientometric analysis. *Library Philosophy and Practice (e-journal)*, *3339*.

Rajendran, P., Elango, B., & Manickaraj, J. (2014). Publication trends and citation impact of tribology research in India: A scientometric study. *Journal of Information Science Theory and Practice*, 2 (1), 22-34.

Rousseau, B. & Rousseau, R. (2000). LOTKA: A program to fit a power law distribution to observed frequency data. *Cybermetrics*, *4*(1), paper 4.

Su, H. N., & Lee, P. C. (2010). Mapping knowledge structure by keyword co-occurrence: a first look at journal papers in technology foresight. *Scientometrics*, 85(1), 65-79.

Von Solms, R., & Van Niekerk, J. (2013). From information security to cyber security. *Computers & Security*, 38, 97-102.

Annexure 1 – Search strategy used

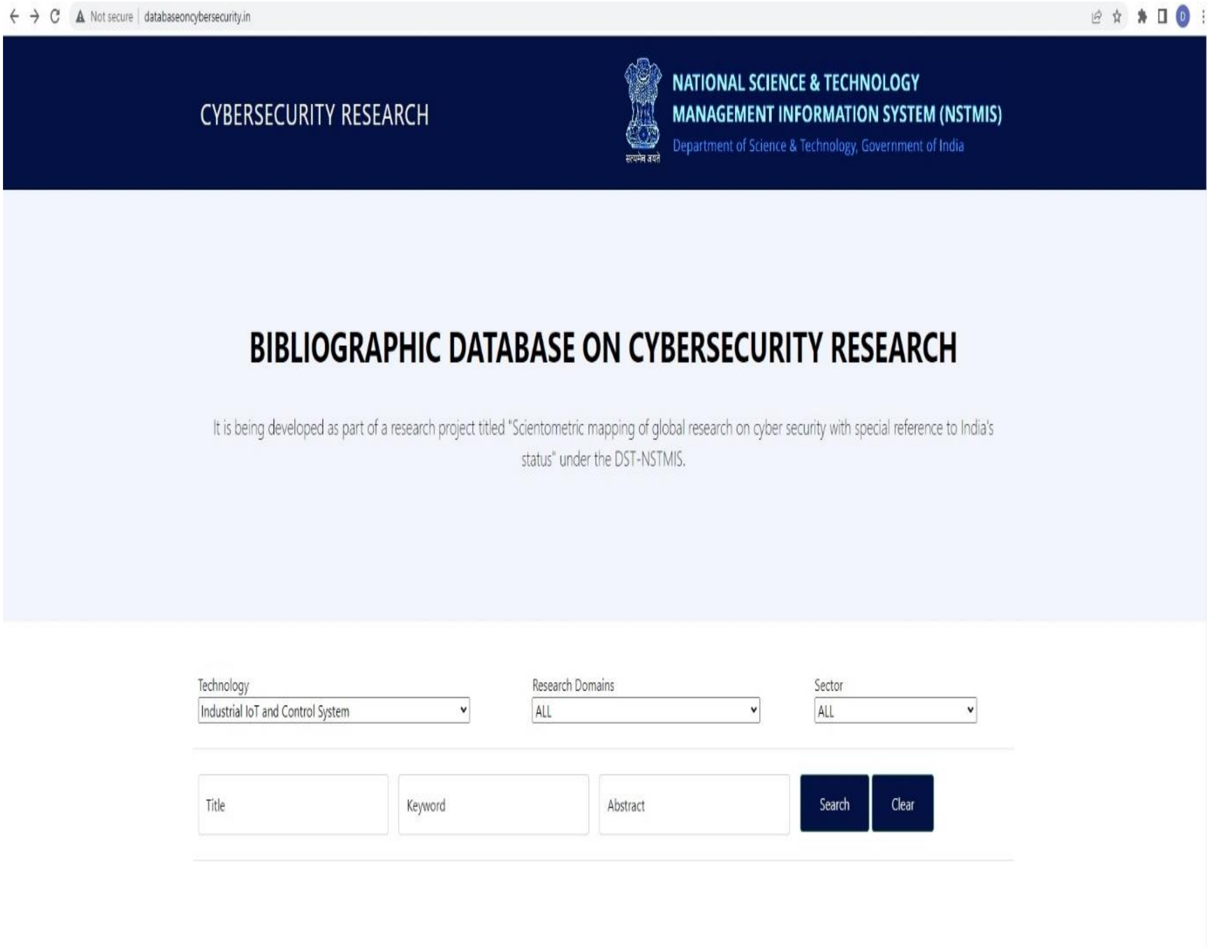
“Cybersecurity” OR “Cyber Security” OR “Cyber-Security” OR “Cyber crisis management” OR “Cyber incident management” OR “Cyber threat management” OR “Cyber Safety” OR “Cybersafety” OR “Cyber defense”

Annexure 2 – Publications out of the project

- Elango, B., Matilda, S., & Jeyasankari, J. (2020). Redefining search terms for Cybersecurity: a bibliometric perspective. In *Proceedings of the International Conference on Recent Advances in Computational Techniques (IC-RACT)*.
- Elango, B., Matilda, S., Martina Jose Mary, M., & Arul Pugazhendhi, M. (**2022**). Mapping the cybersecurity research: a scientometric analysis of Indian publications. *Journal of Computer Information Systems*, doi: 10.1080/08874417.2022.2058644. (**IF = 3.4**).
- Elango, B., Matilda, S., Martina Jose Mary, M., & Arul Pugazhendhi, M. (In Review). The role of cybersecurity in smart grid: a systematic literature review. Wireless Personal Communications (IF = 1.671).
- Elango, B., Matilda, S., Martina Jose Mary, M., & Arul Pugazhendhi, M. (In Review). Top-cited publications in cybersecurity: a bibliometric and content analysis. Wireless Personal Communications (IF = 1.671).

Annexure 3 - End Project Deliverables

- Planned five publications; one presented in a conference, three submitted to journals, one in drafting. Of three submitted, one article has been published in the *Journal of Computer Information Systems* (**IF = 3.4**).
- Exclusive database with classification of publications has been developed and it has been interlinked with the website, http://databaseoncybersecurity.in. The screen of the same has been provided below:


Not secure | databaseoncybersecurity.in
CYBERSECURITY RESEARCH
NATIONAL SCIENCE & TECHNOLOGY
MANAGEMENT INFORMATION SYSTEM (NSTMIS)
Department of Science & Technology, Government of India

BIBLIOGRAPHIC DATABASE ON CYBERSECURITY RESEARCH

It is being developed as part of a research project titled "Scientometric mapping of global research on cyber security with special reference to India's status" under the DST-NSTMIS.

Technology: Industrial IoT and Control System

Research Domains: ALL

Sector: ALL

Title
Keyword
Abstract
Search
Clear

Annexure 4

Beneficial to the various stakeholders

The major goal of this study is to identify and characterize the cybersecurity research with special reference to India's status with a focus on the yearly trend, top authors and institutions, collaborating countries, preferred sources, and most cited publications along with topic trends. The results of this study will:

- create awareness among the research community.
- provide help to the scientists in understanding the publication pattern / structure.
- provide help to decision makers in which topics may provide more funds.
- be an eye-opener to the student / research scholar community.

Research Summary

Scientometric mapping of global research on cyber security with special reference to India’s status, by Dr. B. Elango. IFET College of Engineering, Villupuram. 2021.

The goal of this study is to examine the cybersecurity research conducted by global researchers with special reference to India’s status in response to cyber-related problems, based on the bibliographic data obtained from various databases such as Web of Science, Scopus, Indian Citation Index and Indian Science Abstracts. The project was carried out in phases, as follows: (1) Keyword delineation was completed in the first phase. (2) Three pilot studies were conducted in the second phase. (3) Data collection and analysis were carried out in the final phase. Indian researchers have taken six years to recognize the topic and India is ranked 4th in terms of number of publications at the global level. Compared to publications in the latest emerging domains, there are only a few publications related to cybersecurity by Indian authors. Two-third of the total Indian output were published in the recent three years (e.g. 2018-2020). Half of the top 34 institutions belongs to the category of non-government institutions. Top three institutions are: Amrita Vishwa Vidyapeetham, Amity University and VIT University. Four Indian Institutes of Technology (Delhi, Mandi, Kharagpur and Kanpur) were also listed among the top institutions. Indian institutions having centres of excellence in cybersecurity produces more publications than others. Eleven of top sources were discontinued its coverage by Scopus. If we group all the IITs and NITs into single cluster, these institutions will be in the top positions. The focus of Indian researchers has shifted from data-related issues to smart grid and emerging technologies, to AI-related technologies.