Ransomware and its future trends: A scientometric analysis

Scientific Modelling and Research
Vol. 9, No. 1, 15-43, 2024
e-ISSN: 2523-952X

Kuldeep Mohanty[1]
Veena Goswami[2]
Shahazad Niwazi Qurashi[3]
Rabindra Kumar Barik[4]

(*Corresponding Author*)

[1,2,4]*School of Computer Applications, Kalinga Institute of Industrial Technology, Bhubaneswar, India.*
[1]*Email: 2147007@kiit.ac.in*
[2]*Email: veena@kiit.ac.in*
[4]*Email: rabindrafca@kiit.ac.in*
[3]*Department of Public Health, College of Nursing and Health Sciences, Jazan University, Jazan, Kingdom of Saudi Arabia.*
[3]*Email: squrashi@jazanu.edu.sa*

ABSTRACT

The research article investigates an ever-evolving knowledge mapping of ransomware using the CiteSpace Visualization tool. The foundation for this research is the body of scientific literature on ransomware that was extracted between 2013 and 2023 from the Web of Science Core Collection database. The study methodology is a five-step procedure that follows a systematic approach from data collection to bibliometric analysis using Citespace, identifying hot research topics, analyzing the results, and offering future research paths. It provides cited reference analysis, author and country cooperation networks and co-citation networks, institution collaboration networks, annual publishing patterns, and journal co-citation analysis. Using keyword co-occurrence analysis, it also provides information about hot research topics and emerging trends that pertain to ransomware. The findings identify the most influential authors, countries, and institutions that have actively contributed to ransomware research. The current study also identifies emerging trends and hot research topics, thus providing future research directions in these areas. Ransomware is a highly sophisticated type of malware that has become a dynamic threat to the cyberspace during the last ten years. Technological and IT infrastructure developments have increased the attack surface. The capability of ransomware to wreak havoc in cyberspace has made it a persistent danger. Therefore, it is essential to carry out comprehensive research using scientometric analysis to address the worldwide landscape of ransomware and provide insights into its current situation.

Keywords: *Cite space, Cyber security, Deep learning, Machine learning, Ransomware, Scientometric analysis.*

DOI: 10.55284/smr.v9i1.1259
Citation | Mohanty, K., Goswami, V., Qurashi, S. N., & Barik, R. K. (2024). Ransomware and its future trends: A scientometric analysis. *Scientific Modelling and Research*, *9*(1), 15–43.

Funding: This study received no specific financial support.
Institutional Review Board Statement: Not applicable.
Transparency: The authors confirm that the manuscript is an honest, accurate, and transparent account of the study; that no vital features of the study have been omitted; and that any discrepancies from the study as planned have been explained. This study followed all ethical practices during writing.
Competing Interests: The authors declare that they have no competing interests.
Authors' Contributions: Conceived the study, performed the analysis, and wrote the manuscript, K.M. and V.G.; helped in the analysis, and simulation of drawing the right structures, S.N.Q. and R.K.B.; reviewed the manuscript and refined the tables, K.M. All authors have read and agreed to the published version of the manuscript.
Acknowledgement: The authors gratefully credit the Research Lab, School of Computer Applications, KIIT Deemed to be University, Bhubaneswar, India for providing computational resources.
History: Received: 18 October 2024/ Revised: 25 November 2024/ Accepted: 3 December 2024/ Published: 17 December 2024
Publisher: Online Science Publishing

Highlights of this paper

- The article presents a detailed overview of the research on ransomware that has been conducted over the past ten years (2013–2023).
- The present article describes a method for identifying relevant research areas in the growing field of research.
- The present study tries to identify the most influential researcher, actively participating countries, annual publication trends, and most influential journals in the field of ransomware research.

1. INTRODUCTION

Decades ago, there was a time when the IT industry had not spread its tentacles over the world. Ransom demands were still on. The difference was it was a phone call for your child or colleague, not your data or privacy. It's just the fact that things got digitized. This is the age of information and Technology. The number of cyber attackers has increased due to the IT industry's explosive growth. Ransomware attacks, or RaaS attacks, are becoming the primary weapon and revenue model cybercriminals use. Since ransomware is one of the most dangerous forms of malware and poses a high threat to cyberspace, academicians, research scholars, and cyber security experts have focused their attention on conducting various studies and investigations in this area [1].

Malware known for its ability to extort money or services through ransom demands is termed as ransomware. Ransomware is the most notorious and highly evolving malware of the time. Ransomware bypasses firewalls and other conventional defenses like the IDPS and DMZs, among others, to gain unauthorized access to the target's system. From there, it locks the system or encrypts important files in anticipation of a ransom, typically paid in cryptocurrency like bitcoins. Ransomware works in five steps in chronological order, commencing with the "infection phase," "communication with C & C phase," "destruction phase," followed by the "order and extortion phase," and lastly, the "concluding phase". Figure 1 describes the overall working phase of Ransomware. Ransomware must accomplish major duties in each phase for a successful attack [2, 3]. Joseph Popp created the AIDS Trojan ransomware strain in 1989, and it was sent to victims via email in the form of floppy discs. This ransomware, which employed symmetric cryptography for encryption, was regarded as the weakest ever created since it lacked several crucial features, including practical and quick encryption techniques, anonymous ways to pay the ransom, a broad range of infection channels, and a robust IT infrastructure. It was kept quiet until the early 2000s, but as IT infrastructure advanced, this shortcoming progressively became a significant concern and a big advantage for cybercriminals. The idea of crypto virology served as the foundation for the following generation of ransomware strains, which employed robust and effective encryption [4]. Wannacry was a prominent example supporting this claim [5]. By 2005, the introduction of GP CODE had once again shown the full force of this buried failure. With the growing popularity of the World Wide Web, new strains such as prevention, crypto locker, song, and others began to emerge gradually, and phishing attempts became one of the most often used attack vectors. As time went on, new strains began to appear. In 2017, "WannaCry" one of the most popular strains, caused an estimated $4 billion in losses [6].

Figure 1. Phases in ransomware attack.

It was the worst strain ever! The most recent worst-case scenario was the" Corona" ransomware, which increased during the shutdown and began encrypting folders, including patient information and more hospital data.

The prevalence of ransomware has increased significantly across various industries, negatively affecting the digital world. As per Adam [2] report by Sophos, approximately 60% of ransomware attacks target the healthcare industry, making it one of the most vulnerable sectors [7]. The healthcare sector is particularly susceptible to ransomware attacks due to the financial benefits that may be obtained by exploiting vulnerabilities in complex IT systems and legacy applications inside the business [8]. Still, ransomware does not primarily affect the healthcare sector. The industries most impacted by ransomware are displayed in Figure 2. The top 5 industries that are most severely impacted by ransomware, according to the Sophos analysis, are state and local governments, the media industry, lower and higher education, construction and real estate, and the federal or central government.

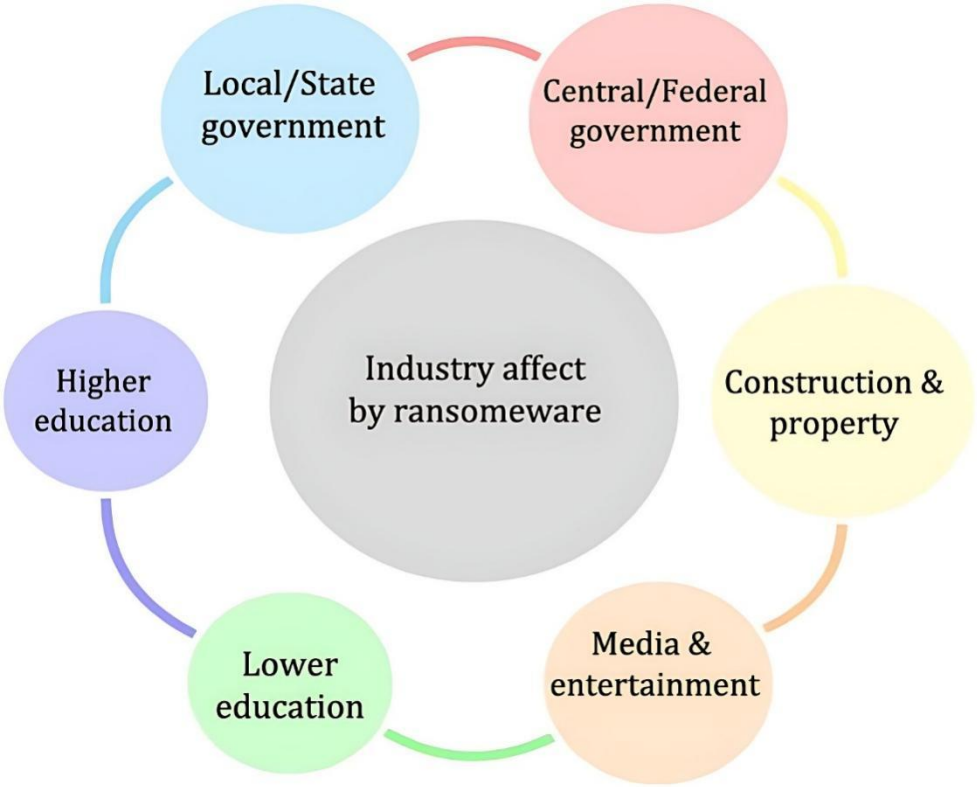


Figure 2. Industries affected by ransomware.

However, ransomware is not limited to any one sector of the economy. Nowadays, practically every business employs IoT-based technological advances that gather and process data stored on cloud-based storage systems. Furthermore, ransomware does not spare a detail when infiltrating networks that employ IoT devices essentially indicating that a ransomware attack is visible in every sector and that it has its tendrils extended over all of them. This necessitates a wide range of ransomware research [9, 10].

Numerous cybersecurity professionals and academics have previously conducted extensive studies on ransomware and have attained remarkable success. While some academics concentrate on the development of ransomware from failed malware to a massive cyber threat, others work on its detection and prevention [4, 11-13]. In contrast, still, others study its potential implications in the future. Even with the abundance of research, organizing, summarizing, and conducting a quantitative analysis of ransomware over a specific period remains challenging, impeding the pace of progress in the sector. There is a growing number of ransomware incidents, which calls for

increased attention. To progress further in their work, researchers should equally prioritize having a cursory comprehension of the subject matter. A bibliographic analysis of ransomware is thus necessary for this in-depth investigation [3, 14, 15].

A detailed review or analysis of numerous research publications, such as book chapters, articles, journals, and many more, is known as a bibliometric analysis. For assessing the effect and influence of research on a particular topic, in this example, ransomware, biblio-metric analysis has gained popularity as an analytical technique. In a specific field, it is a statistical measure to assess research institutes and researchers' productivity and performance. Numerous kinds of analysis and units are offered, such as bibliographic coupling, co-authorship analysis, citation analysis, etc. It gives academic researchers a foundation for investigating topics with limited comprehension yet a thorough grasp of the literature. Bibliometric analysis provides a broad spectrum of scholarly research that may be understood from a micro to a macro viewpoint [16]. Therefore, to conduct a thorough investigation and systematic study of ransomware, this research uses network visualization maps created by VOSviewer. The study also uses additional bibliometric analytics tools, such as Gephi to produce Eigenvector centrality values from the GML file produced from VOSviewer for various analyses, and Citespace to build visualization maps for co-citation journal analysis and references with high citation burst for cited reference analyses. The Eigenvector centrality score represents the effect and influence variables in a particular network.

In addition to seeking cooperation and collaboration with other research scholars and research institutes through the systematic collection and analysis of research papers relating to ransomware, this study offers a helpful pathway for cybersecurity experts and research scholars interested in researching various aspects of ransomware. Additionally, the analysis and projections based on the findings of the bibliometric study might offer future directions for this field's research. The following research questions must be answered to get the desired outcome.

1.1. Research Questions

The purpose of this article is to provide a map of global ransomware research in the field of computer science and engineering. The following research questions will describe the scope of this study.

RQ1: How has the field evolved? For example, what publishing trends have been generally, and how does ransomware research demonstrate its multidisciplinary nature?

By answering this question, the researchers will be better able to assess the state of ransomware research today and its trajectory of progress. Additionally, the multidisciplinary approach offers insights into the most prevalent subject areas that have drawn attention to ransomware research.

RQ2: Which countries, authors, and organizations have significantly contributed to ransomware research?

By answering this question, researchers will discover global connections and cooperation between various research institutes and researchers. Moreover, this would make comprehending how research publications are distributed geographically more accessible.

RQ3: Which cited journals have significantly advanced the knowledge area of ransomware research?

To publish a research paper, this question must be answered. This will assist the researchers in finding a fast-publishing journal with an adequate volume of comparable papers.

RQ4: Which references are most frequently cited throughout research on ransomware? By answering this question, the researchers will be better able to pinpoint important papers and the driving force behind ransomware research.

RQ5: Which cutting-edge research topics and developing patterns are covered in ransomware papers?

Answering this question will assist the researchers in determining the future directions for their ransomware research.

1.2. Contributions

To establish future research road maps, this article presents a detailed overview of the research on ransomware that has been conducted over the past ten years (2013–2023). The following are some noteworthy contributions that have been made to the realm of analysis of ransomware. The present article describes a method for identifying relevant research areas in the growing field of research.

The present study tries to identify the most influential researcher, actively participating countries, annual publication trends, and most influential journals in the field of ransomware research. It contributes an approach to identifying influential topics and most active research topics through keyword co-occurrence analysis and a bag of keywords generated using Python libraries. The present article also investigates the industries affected by ransomware and its root causes. The overall proposed work contributes to Ransomware research by providing a comprehensive study of ransomware research, its worldwide state, and its future research directions.

1.3. Organizations

The overall structure of this research paper is as follows: First, the research methodology that incorporates methods and data collection is described in section 2. Subsequently, the findings of the bibliometric analysis are presented in section 3. This section includes bibliometric analysis of annual publication outputs, countries, and organization collaboration networks, Subject categories analysis, Journal co-citation analysis, and analysis of cited references. Co-occurrence analysis is also used to analyze and showcase popular research topics and new trends. Furthermore, section 4 analyzes the hot research topics and emerging trends in ransomware research. In addition, section 5 discusses the key aspects of research findings and provides future research direction for the researchers. Finally, the paper concludes with section 6.

2. RESEARCH METHODOLOGY

A visual representation and analysis of the ransomware literature is provided, which presents the methodology that was utilized for this paper. Figure 3 outlines the overall research methodology for the present article. The authors broke down the research approach into five distinct steps that were all very crucial.

Stage 1: The first stage in our research was data collection where a good amount of data were collected from the WoS databases' core collection with basic retrieval as 'ransomware' spanning over years between 2013 and 2023.

Stage 2: This stage involves bibliometric analysis of the scientific literature on the ransomware dataset. The results were examined in terms of annual publication trend, subject categories, author collaboration and co-citation, countries and institutional collaboration, Journal Co-citation, Cited references, and keyword co-occurrence analysis.

Stage 3: The keyword co-occurrence analysis gave insights into the hot research topics and emerging trends in ransomware research. Moreover, an author's keyword search was implemented to investigate the top 5 authors in the most frequent keyword.

Stage 4: The results were analyzed and discussed in this section. It also provides insights into the research limitations, and implications, and provides future research directions.

Stage 5: Finally, the research paper is summarized in the conclusion section.

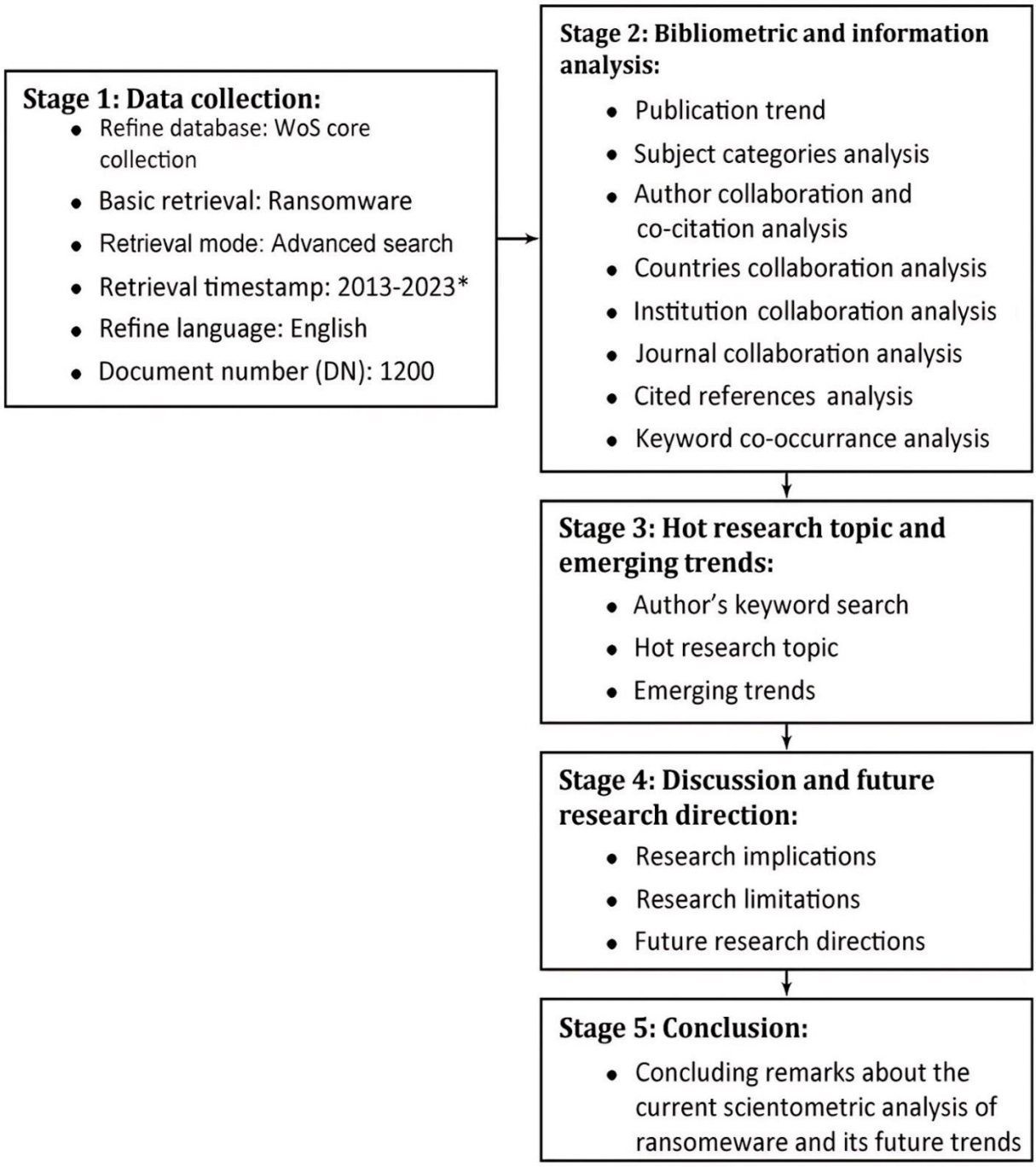


Figure 3. Research methodology.

2.1. Data Acquisition

We have collected abundant data and thoroughly covered global ransomware research using the WoS database. The data that was exported from WoS comprises detailed information (including complete records and cited references in the form of plain text) on the author, institution, year of publication, and source journals. The search type used was advanced search, which searched for the "ransomware" key across all fields and accounted for 1200 records between 2013 and 2023.

2.2. Methods

We leveraged Citespace, a well-known Java application developed by Chaomei, to generate network visualization maps that include countries, authors, journals, institutions, references, citation bursts, and keyword clustering. Emerging research trends in the field are revealed by co-occurrence network analysis and clustering map generated by CiteSpace. Gephi, developed by Mathieu Jacomy, was used to compute the Eigenvector Centrality. Additionally, a node represents an item (such as a country, institution, keyword, author, or journal). The term "ransomware" was chosen by pre-analysis and comparison, and the retrieval period began in 2013 to 2023.

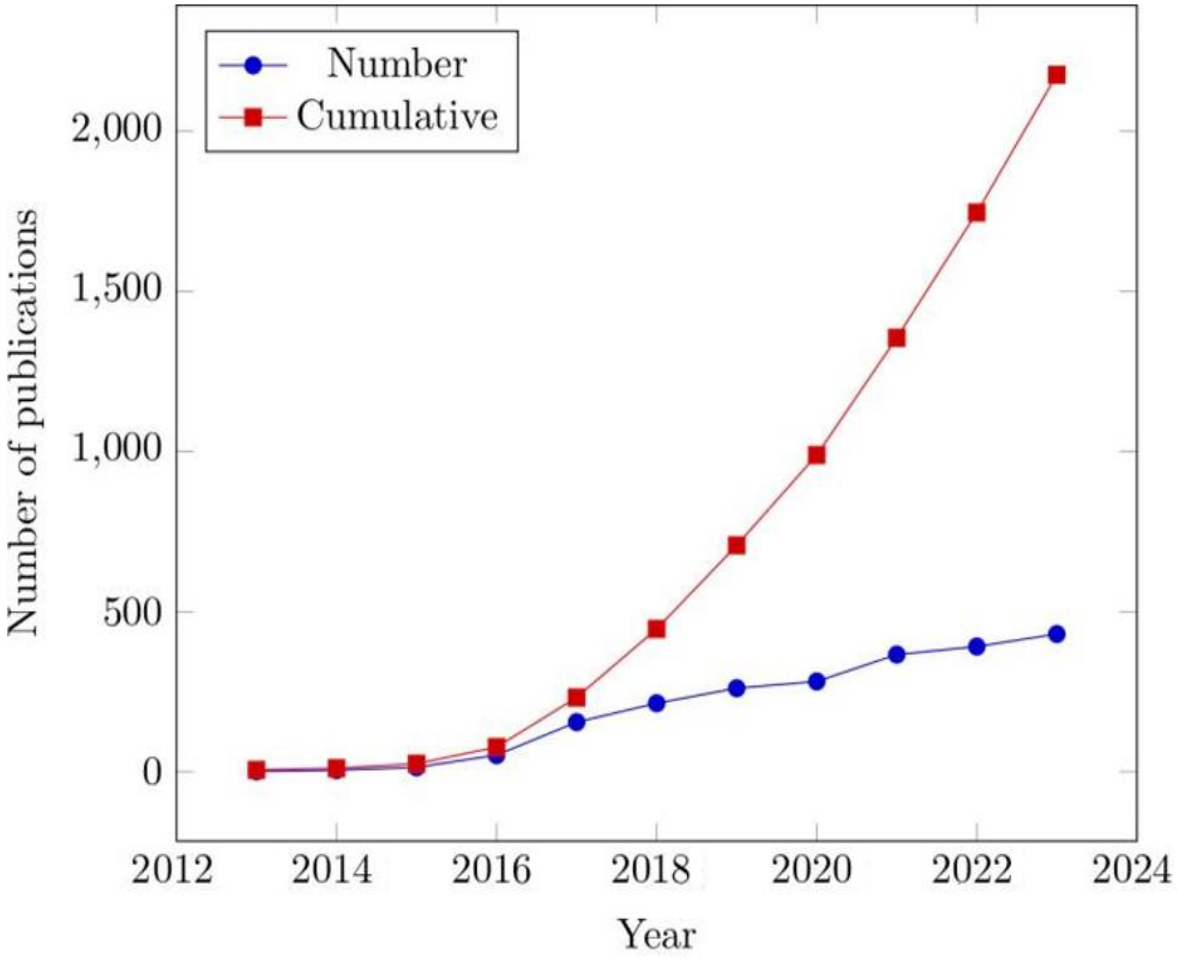


Figure 4. Annual publication trend.

3. RESULTS

3.1. Analysis of Publication Outputs

The annual trend in publication output can be attributed to the growing intensity of ransomware attacks and the investigation of various measures to mitigate them. Figure 4 shows the research publications about ransomware that were published between 2013 and 2023. Two phases may be distinguished in the history of publishing: the first, from 2013 to 2016, had a cumulative publication output of fewer than 100 research publications. However, since 2017, a significant shift has occurred, with an increase in articles on ransomware. The publications were observed to be reaching ever-higher levels throughout the second research period, which ran from 2017 to 2023. There were 143 articles published in 2018 compared to just 93 published in 2017. After reaching 152 articles in 2019, 159 articles in 2020, 180 pieces in 2021, 233 articles in 2022, and 197 articles published in 2023.

Table 1. Top 5 subject categories based on publication.

Subject category	2013 - 2023	% papers	Subgroups in different period	
			2013 - 2016	2017 - 2023
Computer science information systems	523	43.58	40	483
Computer science theory methods	389	32.42	57	332
Engineering electrical electronic	309	25.75	42	267
Telecommunications	207	17.25	25	182
Computer science artificial intelligence	162	13.50	19	143

The cumulative publications increased from just 136 in 2017 to 1200 in 2023, which is still increasing. This upward trend in yearly publications may be partially explained by the rise in the frequency and intensity of ransomware attacks, which is drawing the attention of academics and specialists in cybersecurity [17].

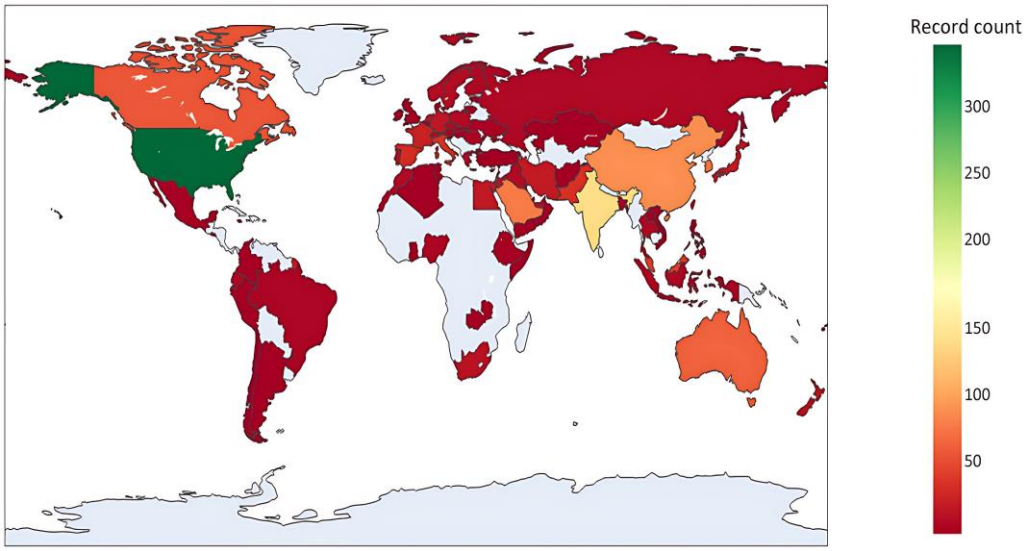


Figure 5. Demonstration of country-wise publications.

3.2. Analysis of Categories

In the category analysis, the top subjects that have drawn the most interest and study on ransomware are reviewed from the WoS database analytics, along with a representation of the state of research development. There were 25 subject categories for ransomware literature in the Web of Science, and Table 1 lists the top 5 subjects for each category. With 523 total research publications between 2013 and 2023, computer science Information Systems leads the field. Computer Science Theory Methods comes in second with 389 articles, Engineering Electrical Electronic with 309 papers, Telecommunications with 207 articles, and Computer Science Artificial Intelligence with 162 articles. There was a marked shift in the number of publications in nearly every topic category from the first evolution time frame, which ran from 2013 to 2016, to the second evolution period, from 2017 to 2023. However, the Information Theory fields of computer science were the most noteworthy of the five, where the overall number of research articles increased significantly.

In just six years, there was an increase from just 40 articles to 523 (or 483 more) on computer science information systems and from just 57 papers to 389 (or 332 more) on Computer Science Theory Methods. With time, the number of research articles in Electrical Electronic Engineering, Telecommunications, and Computer Science Artificial Intelligence gradually increased.

3.3. Analysis of Countries and Institutions

Figure 5 depicts the country-wise distribution of research papers according to their frequencies, with the United States of America leading with 332 publications, followed by India with 136 publications, England with 121 publications, China with 82 publications, and South Korea with 73 publications.

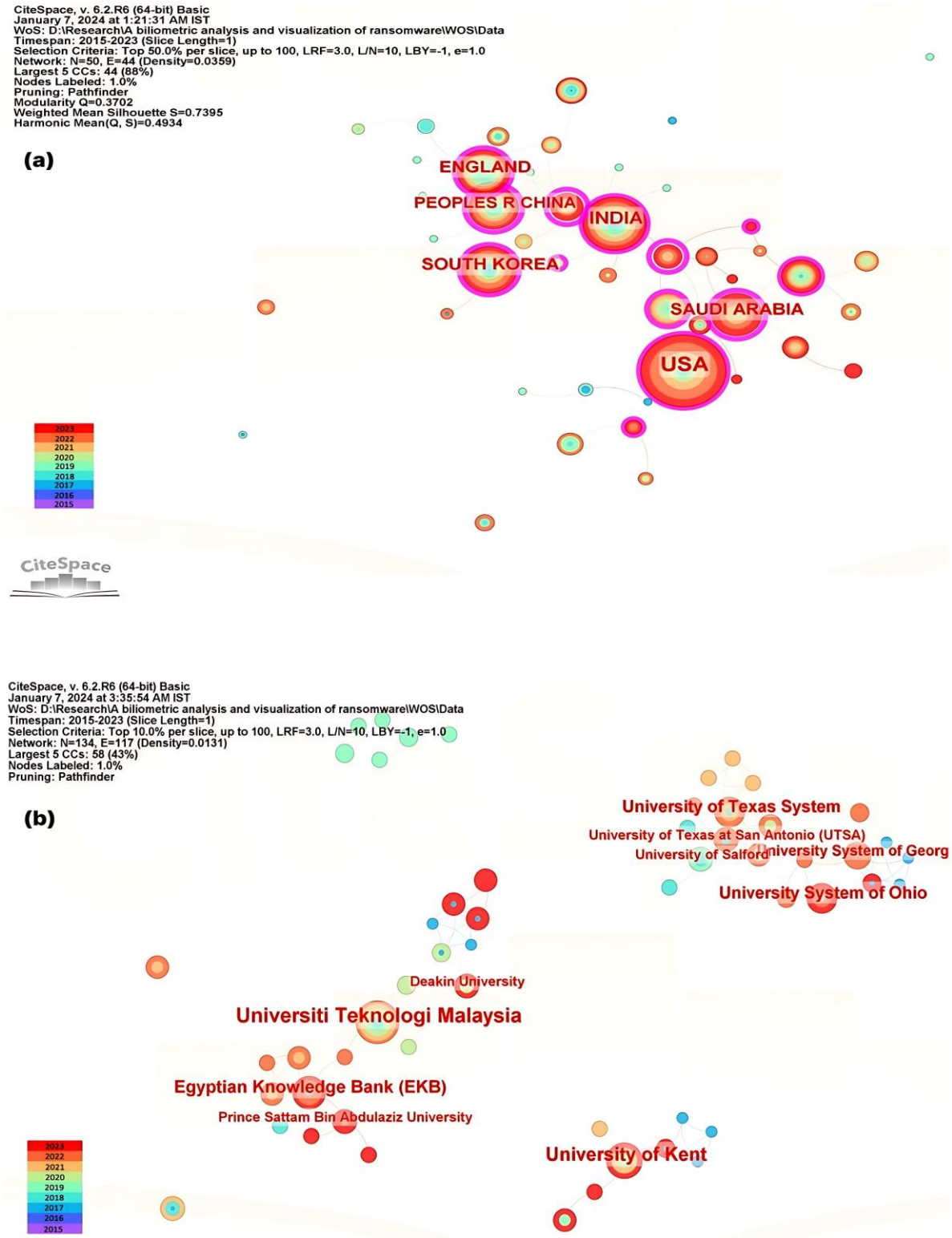


Figure 6. The visualization map of countries and institutions participating in ransomware research: (a) Mapping of main countries participating in ransomware research, (b) Mapping of main institutions participating in ransomware research.

The regional and geographic dispersion of research papers on ransomware may be found by analyzing the research levels conducted in various countries. The country cooperation network could be well understood from Figure 6a (parameter settings: year(s) per slice:1; node type: Country; pruning: pathfinder and pruning the merged network; top N per slice:100; top N%:50%). The node's size, in direct proportionality, represents the amount of research articles produced in a particular nation. The node's size increases when the quantity of research publications in a specific country rises. Notably, Canada is a bridge connecting other nations in the co-authorship network, as evidenced by its highest betweenness centrality of 0.71. Italy and Saudi Arabia have the lowest betweenness centralities, at 0.14 and 0.00, respectively, indicating they are less connected to other nations. Again, Canada has the highest Eigenvector centrality of 1.0, suggesting it is central to a network of countries with high publication counts. England has the second-highest eigenvector centrality of 0.85. However, with the highest h-index of 34, the United States of America is the country with the most frequently cited publications. It is followed by England, which has an h-index of 22; India, which has an h-index of 18; China, which has an h-index of 16; and South Korea, which has an h-index of 13. The top 10 countries are listed in Table 2.

Figure 6 b (parameter settings: year(s) per slice:1; node type: Institution; pruning: pathfinder and pruning the merged network; top N per slice:100; top N%:10%) depicts the leading producers of ransomware research at an institutional level. The University of Texas System (including UTSA), University of London, University System of Georgia, Egyptian Knowledge Bank (EKB), University of Kent, and the Universiti Teknologi Malaysia lead in publication count, suggesting active research efforts. The top 10 most productive institutions for ransomware research are listed in Table 3.

Table 2. Top 10 countries based on publication.

Rank	Country	Publications	% of papers	Total citations (TC)	TC/P	Eigenvector centrality	Betweenness centrality	H-index
1	USA	332	27.67	4218	12.70	0.20	0.17	34
2	India	136	11.33	1192	8.76	0.15	0.18	18
3	England	121	10.08	1664	13.75	0.85	0.25	22
4	China	82	6.83	1227	14.96	0.58	0.42	16
5	South Korea	73	6.08	914	12.52	0.27	0.32	13
6	Saudi Arabia	72	6.00	826	11.47	0.43	0.43	17
7	Australia	59	4.917	725	12.29	0.34	0.14	14
8	Canada	54	4.50	773	14.31	1.0	0.71	15
9	Malaysia	44	3.67	801	18.21	0.36	0.37	14
10	Italy	36	3.00	951	26.42	0.18	0.00	13

The University of Texas leads with a total publication of 31 research papers, followed by the University of London and the University of Texas at San Antonio with 22 research papers each. The University System of Georgia comes next with 20 articles, followed by the Egyptian Knowledge Bank (EKB) and the University of Kent with 18 research papers.

The Universiti Teknologi Malaysia has 16 research papers, followed by the State University of Florida and the University System of Ohio with 13 articles each.

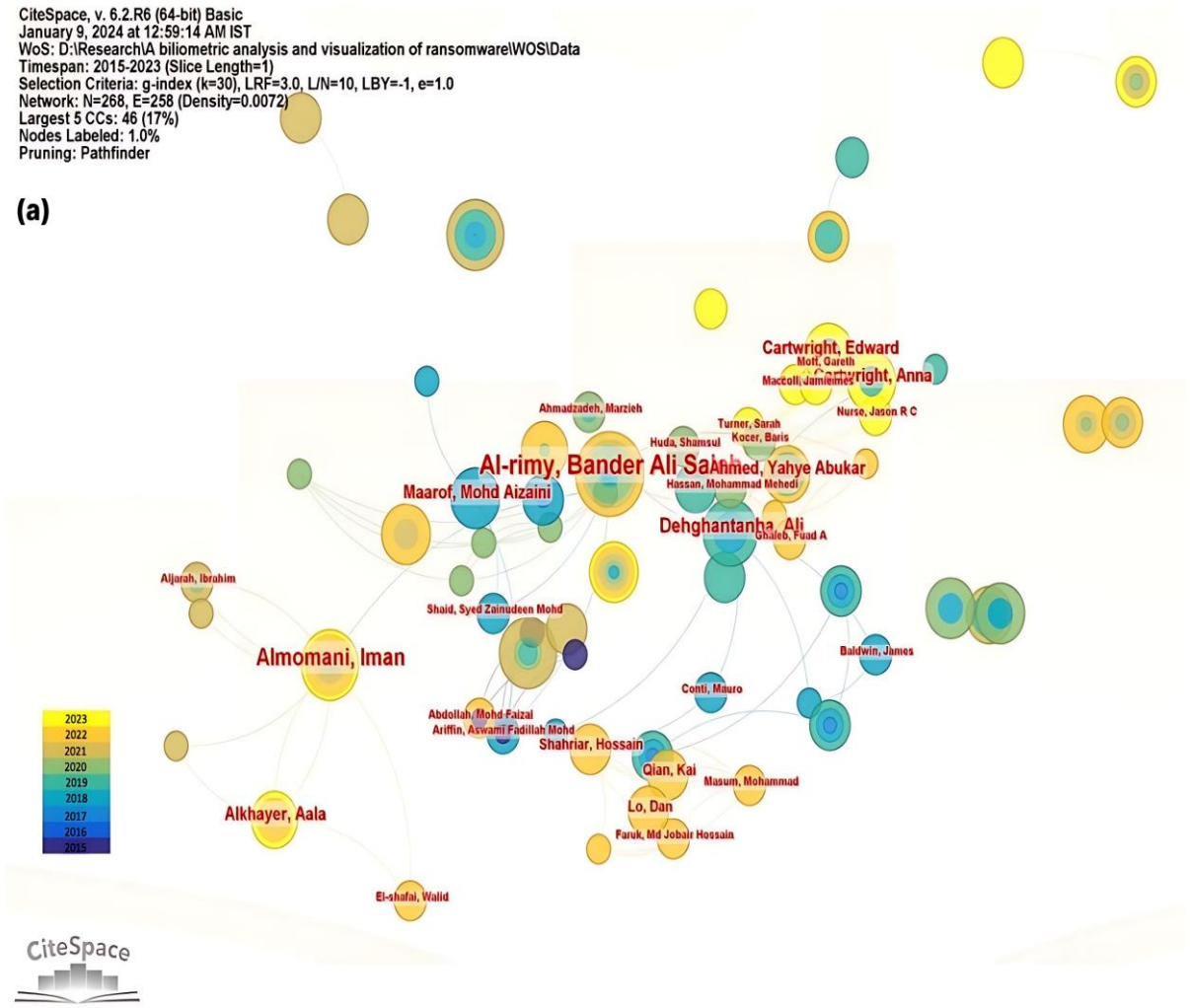
Finally, the Indian Institute of Technology System has 12 articles. Institutions span the US, UK, Egypt, Malaysia, and India, highlighting a global interest in ransomware research. Notably, five out of the top ten positions are occupied by American institutions, suggesting a significant contribution from the US—science Theory Methods. With time, the number of research articles in Electrical Electronic Engineering, Telecommunications, and Computer Science Artificial Intelligence gradually increased.

Table 3. Top 10 institutes based on publication.

Rank	Institution	Publications	% of total	Centrality	Country
1	University of Texas	31	2.58	0.02	USA
2	University of London	22	1.83	0.00	England
3	University of Texas at San An-Tonio UTSA	22	1.83	0.01	USA
4	University system of Georgia	20	1.67	0.02	USA
5	Egyptian knowledge bank Ekb	18	1.50	0.02	
6	University of Kent	18	1.50	0.00	England
7	University of technology Malaysia	16	1.33	0.02	Malaysia
8	State university system of Florida	13	1.08	0.00	USA
9	University system of Ohio	13	1.08	0.00	USA
10	Indian institute of technology system iit system	12	1.00	0.00	India

3.4 Analysis of Author Collaboration Network and Author Co-Citation Network

Analyzing the author's collaboration network and the most productive author reveals valuable insights into ransomware research. Based on 1200 publications, the author collaboration network was mapped in Figure 7a (parameter settings: year(s) per slice:1; node type: Author; pruning: pathfinder and pruning the merged network; g-index:10).



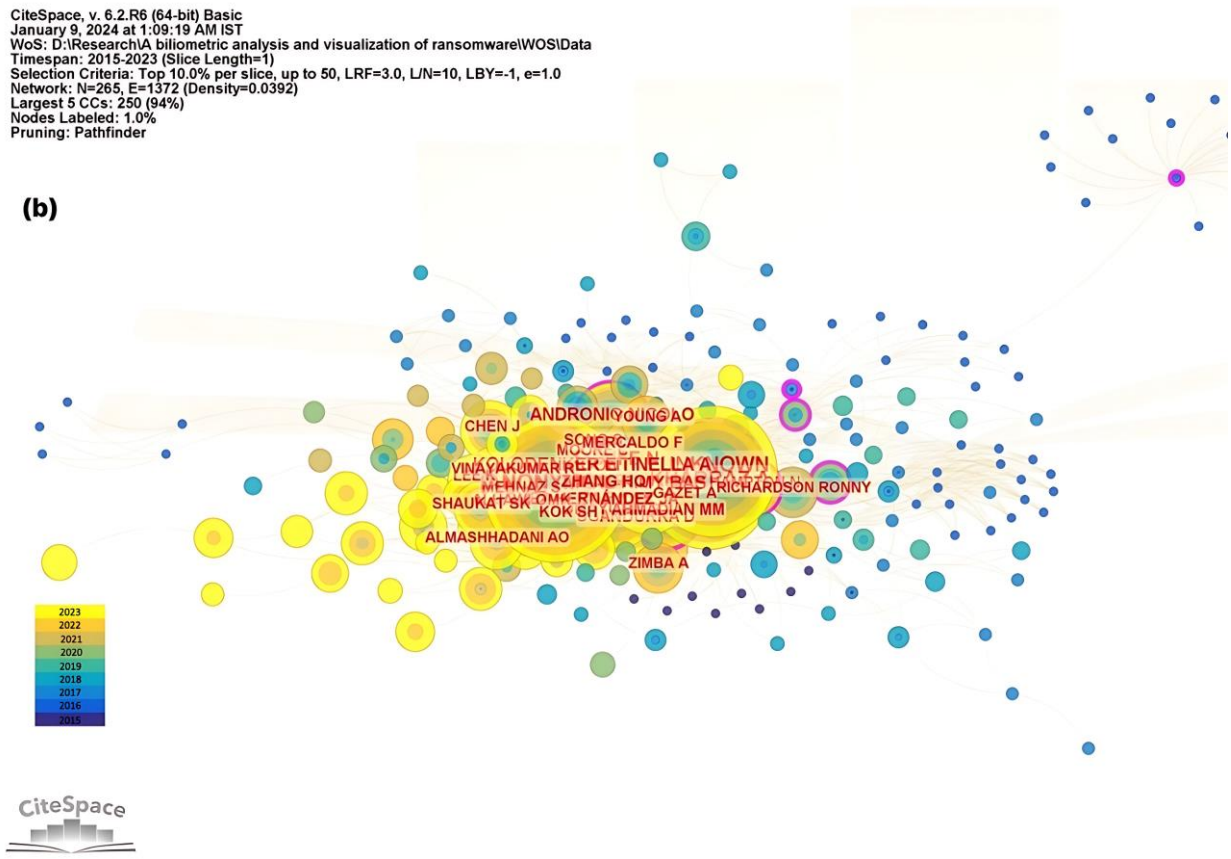


Figure 7. The visualization map of authors participating in ransomware research: (a) mapping of main authors participating in ransomware research. (b) mapping of co-cited authors participating in ransomware research.

One may note that Bander Ali Saleh Al-rimy stands out as a central figure, with two distinct clusters of collaborators, likely reflecting his affiliations with multiple institutions. Julio Hernandez-Castro, Francesco Mercaldo, and Josephine Wolff also hold prominent positions in the network.

The top ten most productive authors from 2013 to 2023 were listed in Table 4. Suhyeony Lee tops the list with 15 publications. Other authors with commendable publication performance were Bander Ali Saleh Al-rimy and Julio Hernandez-Castro with 12 publications each, followed by Francesco Mercaldo and Josephine Wolff, with 11 publications, and so on.

The co-authorship analysis is depicted in Figure 7b (parameter settings: year(s) per slice:1; node type: Cited Authors; pruning: pathfinder and pruning the merged network; top N per slice:50; top N%:10%). The network visualization map had 265 nodes representing authors and 1372 edges representing co-citation links. Amin Kharraz, Nolen Scaife, Krzysztof Cabaj, and Bander Ali Saleh Al-rimy were identified as the most influential authors, having high citation frequencies and a good centrality. Chen J, Androniououng AO, Somercaldo F, Moore, Vinaykumar, Reretinello Alow, Mehnat Zhang Hoay Bas Richardson Ronny, Shaukat SK Omkernadez Gazeta, Kok Shammadian MM, Almashhadani AO, Zimba A, were some of the central authors. The Table 5 lists the top 10 cited authors and their top-cited articles in ransomware research between 2013 and 2023. Kharaz, et al. [18] was the top cited author, with a frequency of 211 and a centrality of 0.00, who proposed UNVEIL, a large-scale automated approach to detect ransomware [18]. Scaife et al. came in second, with a frequency of 147 and a centrality of 0.00, and proposed Cryptodrop. This model uses a set of behavior indicators to halt a process of tampering with a large amount of data [19]. Subsequently, Krzysztof Cabaj recorded a frequency of 111 and a centrality of 0.19 and proposed a novel SDN-based approach that uses the HTTP traffic characteristics to detect ransomware, while Al-rimy et al. trailed behind

with a frequency of 107 and a centrality of 0.11. Al-rimy et al. surveyed ransomware: its threat success factors, taxonomy, and countermeasures [20, 21]. Andrea Continella was next, with a frequency of 106 and a centrality of 0.07, proposed ShieldFS, an add-on driver that makes the Windows native filesystem immune to ransomware attacks [22]. Then the following authors were Eugene Kolodenker, Andronio Nicolo, Daniele Sgandurra, Sajad Homayoun, Hanqi Zhang.

Table 4. Top 10 most productive authors in ransomware research: 2013 – 2023.

Rank	Publications	Authors	Institutions
1	15	Suhyeony Lee	Daegu Gyeongbuk institute of science & technology
2	12	Bander Ali Saleh Al-rimy	Universiti of technology Malaysia
3	12	Julio Hernandez-Castro	University of Kent
4	11	Francesco Mercaldo	University of Molise, Cam- pobasso, Italy
5	11	Josephine Wolff	Tufts university
6	10	Ali Dehghantanha	University of Guelph
7	10	Nitin Naik	Aston University
8	9	Paul Jenkins	Cardiff metropolitan university school of technologies
9	9	Jean-Louis Lanet	Universite de Rennes
10	8	Gabriele Lenzini Lanet	University of Luxembourg

Table 5. Top 10 cited authors and their citation in ransomware research: 2013-2023.

Rank	Frequency	Centrality	Year	Author	Highly cited reference
1	211	0.00	2016	Amin Kharraz	UNVEIL: A large-scale, auto- mated approach to detecting ransomware
2	147	0.00	2016	Nolen Scaife	CryptoLock (And drop it): Stop-ping ransomware attacks on user data
3	111	0.19	2018	Krzysztof Cabaj	Software-defined networking-based crypto-ransomware detection using HTTP traffic characteristics
4	107	0.11	2018	Bander Ali Saleh Al-rimy	Ransomware threat success factors, taxonomy, and counter- measures: A survey and research directions
5	106	0.07	2016	Andrea Continella	ShieldFS: A self-healing, Ransomware-aware filesystem
6	92	0.01	2016	Eugene Kolo-denker	PAYBREAK: Defense against cryptographic ransomware
7	82	0.03	2015	Andronio Nicolo	HELDROID: Dissecting and detecting mobile ransomware
8	68	0.22	2019	Daniele Sgandurra	On deception-based protection against cryptographic ransomware
9	66	0.39	2020	Sajad Homayoun	Know abnormal, find evil: Frequent pattern mining for ransomware threat hunting and intelligence
10	64	0.04	2019	Hanqi Zhang	Classification of ransomware families with machine learning based on N-gram of opcodes

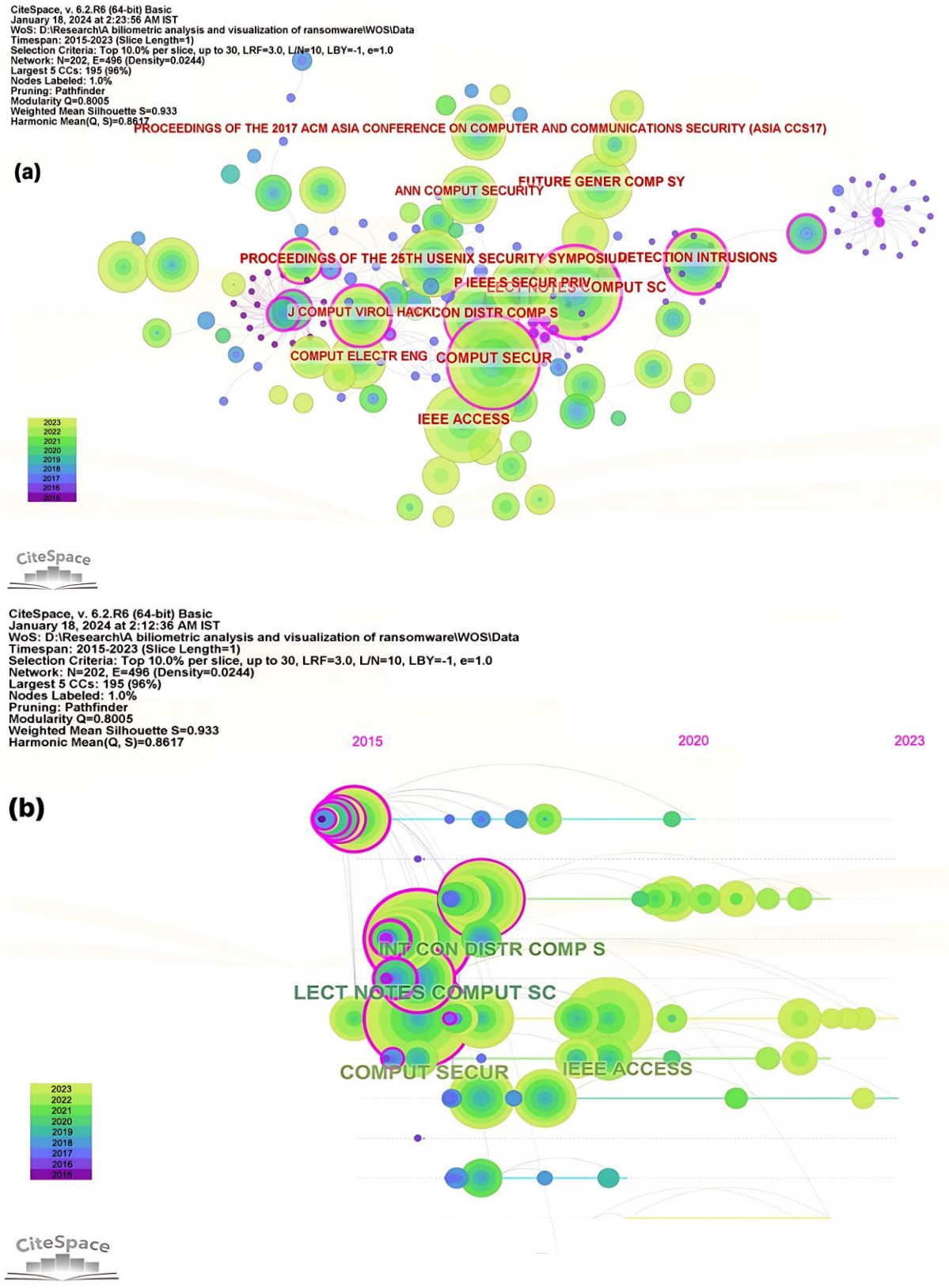


Figure 8. The visualization map of co-citation journal network in ransomware research: (a) normal view (b) timeline view.

3.5. Analysis of Co-Citation Journals

The journal co-citation analysis is an integral part of bibliometrics. It helps to recognize the core journals in a particular field. The Table 6 lists the top 5 productive journals in ransomware research between 2013 and 2023. With

58 publications overall, or 4.84% of the total papers, IEEE Access was the most effective journal. It was followed by Lecture Notes in Computer Science, with 52 publications overall, or 4.33%; Computers Security, with 38 publications overall, or 3.17%; International Journal of Advanced Computer Science and Applications, with 18 publications overall, or 1.5%, and Sensors, with 17 publications overall, or 1.4% of the total papers. When comparing total citations and average citations per paper, IEEE Access again tops the list with 1010 citations overall and 17.41 average citations per paper. The next highest number of citations is 662 for Computers Security, with an average of 17.42 per paper; Lecture Notes in Computer Science, with 477 total citations and an average of 13.1 per paper; Sensors, with 99 total citations and an average of 6.06 per paper; and, lastly, the International Journal of Advanced Computer Science and Applications, with 18 total citations and an average of 1.33 per paper.

Table 6. Top 5 productive journals in ransomware research:2013-2023.

Rank	Journal	P_a	% of papers	TC_b	TC/P_c	H-index	JCI
1	IEEE access	58	4.84	1010	17.41	19	0.89
2	Lecture notes in computer science	52	4.33	477	13.1	13	NA
3	Computers security	38	3.17	662	17.42	12	1.34
4	International journal of advanced computer science and applications	18	1.5	24	1.33	3	0.17
5	Sensors	17	1.41	99	6.06	5	0.89

Note: P_a : The total publication of the journal in ransomware; TC_b : The total citations of a journal, based on WoS* analytics; TC/P_C: The average number of citations per paper for a journal, based on WoS* analytics; JCI: A three-year citation indicator, based on the data from the 2022 edition of clarivate master journal list re- ports; H-index: Sourced from WoS* analytics; Note:- Fields with missing data value are represented with NA.

When the top 5 listed journals are compared based on their H-index, IEEE Access ranks highest with an H-index of 19, followed by Lecture Notes in Computer Science (H-index: 13), Computers Security (H- index: 12), Sensors (H-index: 5), and International Journal of Advanced Computer Science and Applications (H-index: 3). The Journal Citation Indicator (JCI) is an important metrics that provides a fair, stable, and interpretable measure of journal impact across different research disciplines.

The Computer Security Journal has the highest JCI score of 1.34, suggesting its strong influence and indicating that its articles are cited 34% more often than the average within its field of computer security. IEEE Access and Sensors both have the same JCI score of 0.89, which is quite close to average, suggesting that its impact is on par with other journals in the domain.

While the Lecture Notes in Computer Science lacks a JCI score, the International Journal of Advanced Computer Science and Applications ranks last among the top five with a low JCI score of 0.17, indicating its articles are cited less frequently than the average in its field. Figure 8a represents the journal co-citation analysis network visualization map, whereas Figure 8b represents the timeline view of the network. The network visualization map indicating the occupancy of central positions in the network suggests several journals' influences and potential role as knowledge hubs by journals like Computers & Security.

Meanwhile, several other journals are located on the network's periphery, including the International Journal of Network Security, the Journal of Computer Virology and Hacking Techniques, and the Future Generation Computer Systems Journal. Additionally, the map reveals a strong connection between computer security and other central journals, suggesting its role as a major hub for knowledge exchange and collaboration in ransomware research. visualization map.

3.6. Analysis of Cited References

Cited Reference Analysis or Citation Analysis is an integral part of bibliometric analysis that helps to identify the top cited works in literature within a specific domain. Burst detection is one of the most effective tools for identifying significant information within a certain timeframe. The top 25 strongest references generated by CiteSpace from 2013- 2023 (parameter settings: year(s) per slice:1; node type: reference; pruning: pathfinder and pruning the merged network; top N per Slice:30; top N%:10%) have been depicted in Figure 9. The time interval and the period of reference occurrence were defined by the blue and red parts, respectively. The burst detection is divided into two periods corresponding to the publication output. The first period, which spanned between 2013 and 2016, accounted for only one reference with a citation burst strength of 5.12, which came from a paper written by Yang et al. that discusses automated detection and analysis of android ransomware [23]. The burst began in 2016 and ended in 2019. During the second period, which spanned between 2017 and 2023, it accounted for the largest citation burst of 15.22, which came from a paper written by Kharraz, et al. [24]. The burst began in 2017 and ended in 2020. Kharraz, et al. [24] conducted a long-term study on ransomware, unfolding its evolution and proposing various methods of defense against ransomware, one of which is analyzing abnormal file system activity [24]. The second-highest reference with the strongest citation burst during this period accounted for a strength of 9.84 and was written by Andronio, et al. [25]. The burst began in 2017 and ended in 2020. Andronio et al. proposed a framework, HELROID - to detect mobile ransomware [25].

Figure 9. Top 10 references with the strongest citation burst.

Table 7. Top 10 cited references.

Citation	Year	Author(s)	Title	Source	Country
291	2017	Yanfang Ye	A survey on malware detection using data mining techniques	ACM computing survey	USA
218	2019	TaeGuen Kim	A multimodal deep learning method for android malware detection using various features	IEEE transactions on Information forensics & security	South Korea
214	2016	Nolen Scaife	CryptoLock (and drop it): Stopping ransomware Attacks on user data	Proceedings 2016 IEEE 36th International conference on distributed computing systems ICDCS 2016	USA
213	2016	Amin Kharraz	UNVEIL: A large-scale, automated approach to detecting ransomware	Proceedings of the 25th USENIX Security Symposium	USA
179	2020	Abdullah Alsaedi	TON_IoT telemetry dataset: A new generation dataset of IoT and IIoT for data-driven intrusion detection systems	IEEE access	Australia
161	2018	Bander Ali Saleh Al-rimy	Ransomware threat success factors, taxonomy, and countermeasures: A survey and research directions	Computers & security	Malaysia
145	2014	Spagnuolo Michele	BitIodine: Extracting intelligence from the Bitcoin network	Lecture notes in computer science	Italy
142	2017	Ibrar Yaqoob	The rise of ransomware and emerging security challenges in the Internet of things	Computer networks	Malaysia
140	2018	Coventry Lynne	Cybersecurity in healthcare: A narrative review of trends, threats and ways forward	maturitas	Ireland
136	2017	Clemens Scott Kruise	Cybersecurity in healthcare: A systematic review of modern threats and trend-	technology and healthcare	USA

Table 7 lists the top 10 references, according to their citation frequency, for ransomware research during the span of the study. The most cited reference in ransomware research was that of Yanfang Ye, "A Survey on Malware Detection using Data Mining Techniques" [26] Table 5 and Table 7 display overlapped or redundant authors, indicating that highly influential authors were nearly always cited authors based on publication and citation statistics. The authors were Kharaz, et al. [18] and Scaife, et al. [19]. However, it was noteworthy that there was no convergence of authors in table 4 and Table 7. Regarding geographical distribution, it was found that Yanfang Ye, Nolen Scaife, Amin Kharraz, and Clemens Scott Kruise were from the USA, demonstrating that the United States of America was among the first to research ransomware.

4. HOT RESEARCH TOPICS AND EMERGING TRENDS

4.1. Author's Keyword Search

This section implements an author's keyword search to investigate the author's keywords that pertain to ransomware and the top 5 authors as per their publication count in each keyword. The Table 8 analyzes the authors and keywords related to ransomware research. Suhyeon Lee, Bander Ali Saleh Al-Rimy, Francesco Mercaldo, Ali Deghantanha, and Nitin Naik take the lead in journals associated with ransomware research. However, several other authors work on areas related to ransomware like Malware, Security, CyberSecurity, Machine Learning, Analysis, Bitcoin, Ransomware Detection, Deep Learning, Data, and many other fields. Meanwhile, ransomware also has a good correlation with fields and industries like the healthcare industry and the emerging field of blockchain. The keywords were obtained from WoS databases' keyword search within the basic retrieval of ransomware.

Table 8. Top 10 keyword search along with Author name.

S.no.	Keywords	Top 5 author names (By publication count)	No. of articles	S.No.	Keywords	Top 5 author names (by publication count)	No. of articles
1	Ransomware	Suhyeon Lee	15	6	Analysis	Julio Hernandez-castro	8
		Bander Ali Saleh Al-Rimy	12			Paul Jenkins	8
		Francesco Mercaldo	12			Nitin Naik	8
		Ali Dehghantanha	10			Jean-Louis Lanet	7
		Nitin Naik	10			Bander Ali Saleh AlRimy	6
		Francesco Mercaldo	12			Stephen McCombie	4
		Bander Ali Saleh Al-Rimy	10			Adam Brian Turner	4
2	Malware	Ali Dehghantanha	9	7	Bitcoin	Allon J. Uhlmann	4
		Nitin Naik	9			Bander Ali Saleh AlRimy	3
		Paul Jenkins	8			Kim-Kwang Raymond Choo	3
		Francesco Mercaldo	8			Harith Al-Sahaf	6
		Julio Hernandez-castro	7			Soojin Lee	6
3	Security	Ali Dehghantanha	6	8	Ransomware detection	Francesco Mercaldo	6
		Bander Ali Saleh Al-Rimy	5			Ian Welch	6
		Sanggu Lee	5			Muhammad Shabbir Abbasi	5
		Suhyeon Lee	15			Muna Al-Hawawreh	4
		Julio Hernandez-castro	13			Mario Antunes	4
4	Cybersecurity	Bander Ali Saleh Al-Rimy	12	9	Deep learning	Frederick Sheldon	4
		Francesco Mercaldo	12			Elena Sitnikova	4
		Josephine Wolff	12			Abdullah Qunayfith Alqah Tani	3
		Mario Antunes	5			Soojin Lee	9
		Ali Dehghantanha	5			Bander Ali Saleh AlRimy	8
5	Machine learning	Nir Nissim	5	10	Data	Budi Arief	7
		Iman Almomani	4			Iman Almomani	6
		Aviad Cohen	4			Simon Davies	6

The keyword search was refined to 'must include' and searched accordingly. Figure 10 displays a bag of keywords generated by a Python "wordcloud" library from given texts. The texts were the authors' keywords that were taken from the experiment dataset, which served as the foundation for all of the current study's experimentation.

After preprocessing the texts, word frequencies were produced, and word clouds were used for visualization. It must be ensured that the required libraries which include pandas, numpy, matplotlib, and wordcloud is installed in the system before the experiment.

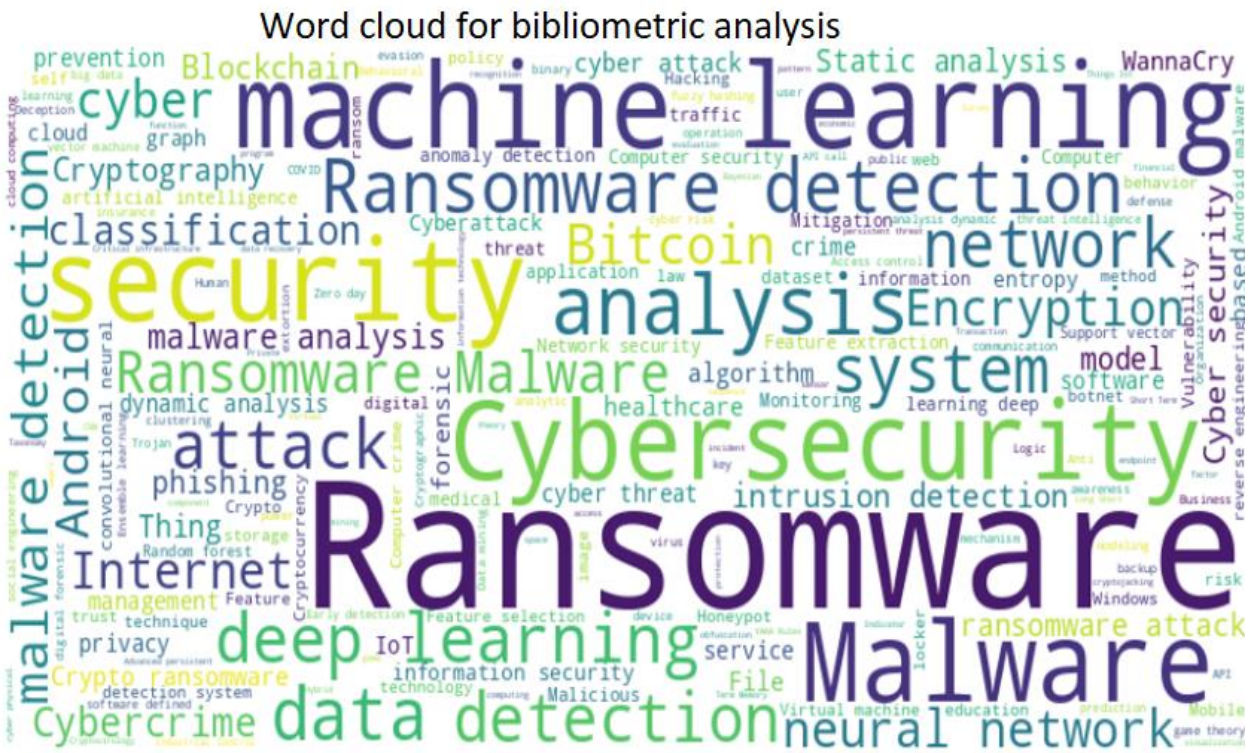


Figure 10. Bag of keyword.

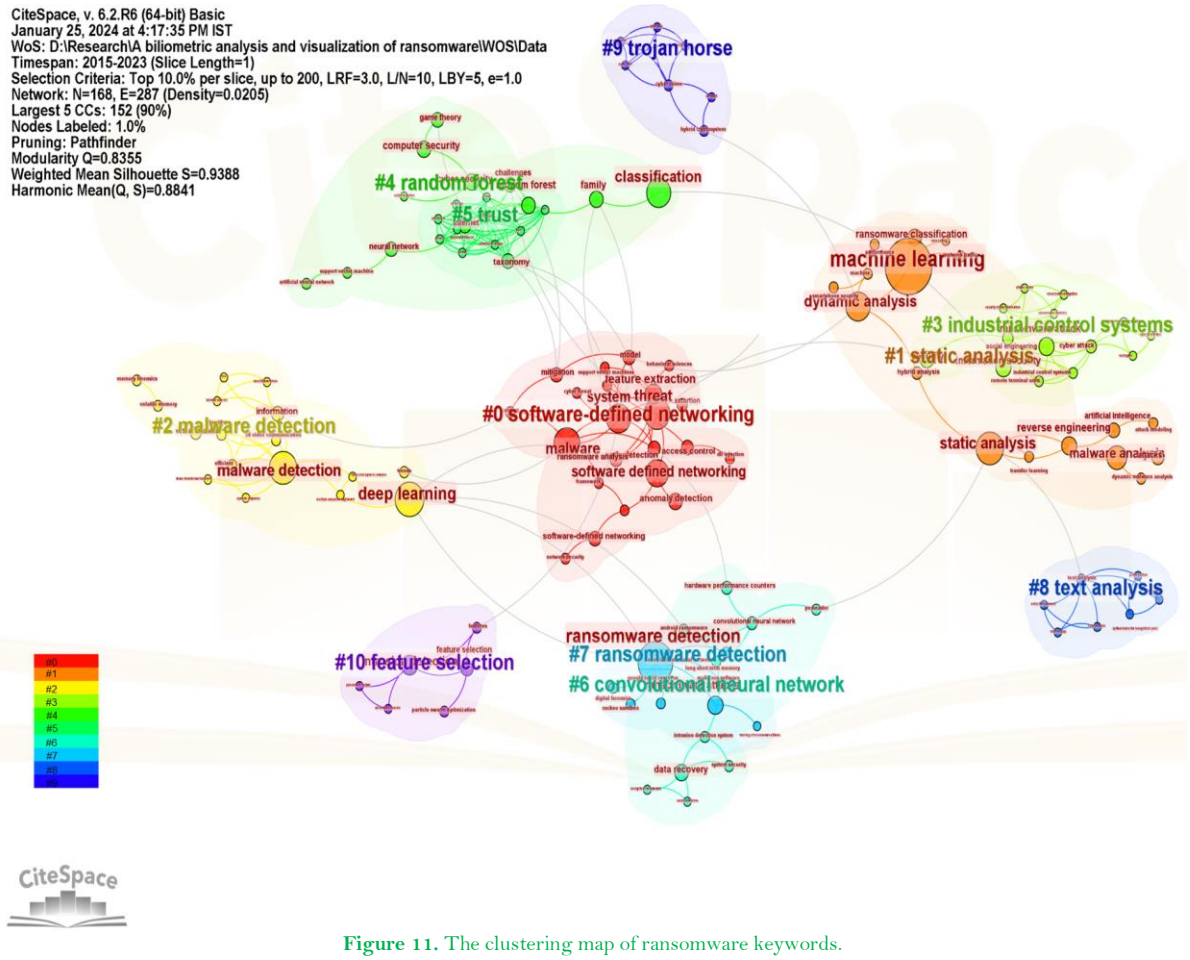


Figure 11. The clustering map of ransomware keywords.

4.2. Hot Research Topics

We have constructed a keyword co-occurrence map using citespace, thus identifying the emerging research domains in ransomware. Figure 11 presents the group of 16 different clusters. Figure 12 depicts a knowledge domain map of keyword co-occurrence. The top five largest clusters are listed below.

Cluster #0 (Feedback-based annotated TF-IDF Technique): The largest cluster (#0) has 22 members and a silhouette value of 0.833. It is labeled as feedback-based annotated tf-idf technique by LLR, dynamic crypto-ransomware pre-encryption boundary delineation by LSI, and windows-based ransomware taxonomy (1.09) by MI. The major citing article of the cluster is: The age of Ransomware: a survey on the Evolution, taxonomy, and research directions [27].

Cluster #1 (Hybrid Analysis) : The second largest cluster (#1) has 17 members and a silhouette value of 0.952. It is labeled as hybrid analysis by LLR, ransomware classification by LSI, and industrial control system (3.07) by MI. The major citing article of the cluster is: Rwarmor: a static-informed dynamic analysis approach for early detection of cryptographic windows ransomware [28].

Cluster #2 (Android IoT Device): The third largest cluster (#2) has 17 members and a silhouette value of 0.951. It is labeled as android iot device by LLR, ransomware detection by LSI, and third generation cerber ransomware (0.63) by MI. The major citing article of the cluster is: A crypto-steganography approach for hiding ransomware within hevc streams in android iot devices [29].

Cluster #3 (Critical Infrastructure) : The 4th largest cluster (#3) has 14 members and a silhouette value of 0.935. It is labeled as critical infrastructure by LLR, social engineering by LSI, and multistage game (0.03) by MI. The major citing article of the cluster is: Ransomware impact to scada systems and its scope to critical infrastructure [30].

Cluster #4 (detecting ransomware attack): The 5th largest cluster (#4) has 12 members and a silhouette value of 0.883. It is labeled as detecting ransomware attack by LLR, ransomware detection by LSI, and ensemble model ransomware classification (0.41) by MI. The major citing

article of the cluster is: Ransomware detection using random forest technique [3]. As shown in Table 9, Machine Learning was the hottest topic in the field of ransomware, with the highest frequency of 94, followed by ransomware detection with a frequency of 50, followed by deep learning with a frequency of 34, followed by static analysis with a frequency of 29, followed by malware detection with a frequency of 28 and so on.

4.2.1. Machine Learning

Arthur Samuel first proposed the concept of Machine Learning in the 1950s, and it later became a multidisciplinary field of research. Machine Learning is about developing an algorithm where a computer or a program can learn from the surroundings, on datasets, and make decisions without human interventions [31]. Much research has already been done on the multidisciplinary ransomware and machine learning field. One such work was 'RansomWall' by Shaukat and Ribeiro [32]. The paper introduces a layered defense system that follows a static and dynamic analysis of particular factors categorizing ransomware behavior. Another work involved classifying ransomware families by leveraging machine learning with N-grams of Opcode [26]. Machine learning could be leveraged as a wall of defense against ransomware, thus seeking more attention and research.

4.2.2. Ransomware Detection

Ransomware detection has become a significant challenge in the realm of ransomware. There has been ample research conducted on the detection of ransomware, but ransomware still stands among the top malware in cyberspace. One of the reasons could be the highly evolving nature of ransomware and the fact that machine learning

has been widely used as a defense against it. Machine Learning detects ransomware on file system behaviors and analyses binaries of ransomware files, but it may not be able to classify zero-day exploits [33]. While the Machine Learning approach forms the majority of detection approaches, there are other approaches. One such approach is Whitelist and Blacklist-based ransomware detection [34]. Some other detection methods include Rule-based detection, static, and dynamic analysis- based detection approaches [35, 36].

Table 9. Top 30 keywords of ransomware research based on the frequency.

Rank	Frequency	Year	Keyword	Rank	Frequency	Year	Keyword
1	94	2017	Machine learning	16	9	2017	Ransomware attacks
2	50	2018	Ransomware detection	17	8	2017	Information security
3	34	2017	Deep learning	18	8	2019	Ransomware attack
4	29	2017	Static analysis	19	7	2019	Family
5	28	2017	Malware detection	20	7	2020	Random forest
6	27	2019	Threat	21	7	2016	Computer security
7	26	2018	Malware	22	7	2019	Ransomware classificati
8	24	2019	Classification	23	6	2019	Security
9	23	2017	Dynamic analysis	24	6	2017	Data recovery
10	22	2019	System	25	6	2016	Cyber security
11	20	2018	Software-defined networking	26	6	2018	Anomaly detection
12	14	2019	Malware analysis	27	5	2017	Taxonomy
13	11	2020	Feature extraction	28	5	2017	Information
14	11	2017	Intrusion detection	29	5	2020	Early detection
15	9	2018	Reverse engineering	30	5	2017	Crypto-ransomware

4.2.3. Deep Learning

The term Deep Learning was first coined by Rina Dechter in 1986. Unlike Machine Learning, Deep Learning leverages Artificial neural networks to mimic the learning process of a human brain and make decisions without any human interventions [37]. After Machine Learning, there has been various research on Deep Learning techniques to detect ransomware. One such work is "DeepWare", proposed by Ganfure, et al. [38]. DeepWare offers a promising approach to ransomware detection by combining deep learning with HPC data analysis. Another work by Maniath et al. suggests a deep learning-based LSTM-based detection [39]. This approach using LSTM networks for API call sequence analysis offers a promising avenue for automated ransomware detection.

4.2.4. Static Analysis

Static analysis, in simpler words, refers to the analysis of a code file without its execution [40]. Unlike Dynamic analysis, code files are not executed in a virtual environment and are analyzed manually; its structural information, including the malicious string that represents the malicious software, is extracted [20, 41]. Static analysis can be used to extract information and build datasets to train a Machine Learning Model on ransomware detection and predict the nature of a particular code/file, i.e. whether it is a malicious file or not. This indicates a direct linkage to tools like Machine Learning, which can be leveraged to form an additional layer of defense against ransomware. Numerous studies on the field of ransomware static analysis and the multidisciplinary area of static analysis and machine learning in ransomware have previously been conducted by different scholars. PEFile was proposed by Poudyal et al., which provides a valuable approach for statically studying ransomware behavior [33]. Extracting features and building databases lays the groundwork for developing advanced detection methods using machine learning and other techniques While some scholars contributed by proposing different frameworks and models, many others focused on achieving high accuracy, i.e. f1 score and other metrics using static analysis. Chanajitt et al. combined the static and dynamic analysis techniques to achieve a high F1 score [42]. RWArmor was proposed by

Ayub, et al. [28]. RWArmor presents a promising approach for tackling the challenging task of ransomware detection. Combining the strengths of static and dynamic analysis allows for faster and more accurate identification of known and unknown ransomware threats, potentially mitigating the risks and consequences of these attacks. Although static analysis provides a promising solution for ransomware, it also has some limits by which it abides. One such limit is dealing with packed families, i.e. the ransomware families that use packers for compression and encryption of their payloads. Moreover, this approach fails to deal with highly evolving malware strains using obfuscation techniques [43].

Table 10. Top 5 key labels as ransomware research topics.

Cluster ID	Label (LLR)	Label (Keyword)	Size	Silhouette	Mean	Representative article
0	Feedback-based annotated tf-idf technique	Software-defined networking	22	0.833	2020	[27]
1	Hybrid analysis	Static analysis	17	0.952	2019	[40]
2	Android iot device	Malware detection	17	0.951	2018	[29]
3	Critical infrastructure	Industrial control system	14	0.935	2017	[30]
4	Detecting ransomware attack	Random forest	12	0.833	2019	[44]

4.3. Emerging Trends

The cluster analysis of keywords, generated using a keyword co-occurrence analysis in CiteSpace, illustrated in Figure 11 was taken into account to obtain the emerging trends in ransomware research. The modularity = 0.8355 > 0.3, and the weighted mean silhouette = 0.9388 > 0.5, which indicates good clustering. The Table 10 lists the cluster labels by LLR and Keyword (of WoS) along with cluster-ID, size, silhouette, and mean year.

The cluster #0 was labeled as "feedback-based annotated tf-idf technique" by LLR and "Software-Defined Networking" by Keyword (WoS). The most actively cited article by Razaulla et al. provided a comprehensive report on the evolution of ransomware, its taxonomy, and its state-of-the-art research contributions. The study also revealed the lacuna in ransomware research relating to zero-day exploits, certain issues with machine learning models, and real-time protection against ransomware [27].

The cluster #1 was labeled as "Hybrid Analysis" by LLR and "static analysis" by Keyword (WoS). Md wrote the most cited article. Ayub et al. proposed a new framework, "RWArmor", that combines static and dynamic analysis for ransomware detection [28]. By leveraging both existing knowledge (static features) and real-time behavior (dynamic analysis), RWArmor can identify even never-seen-before ransomware within a critical window of 30-120 seconds. Their tests on 215 actual ransomware samples yielded impressive accuracy rates, ranging from 97.67% to 86.42%, depending on the time frame for analysis. This promising approach offers faster and more accurate detection than traditional methods, potentially saving organizations from data loss and financial harm.

The cluster #2 was classified as "Malware Detection" by Keyword (WoS) and "Android IoT Devices" by LLR. Almomani et al. is credited with writing the most cited article [29]. They experimented with hybrid crypto-steganography, which conceals dangerous Android malware behind high-resolution films. They used two methods: first, they used a robust AES algorithm to encrypt the ransomware data, guaranteeing its integrity; second, they used LSB steganography, which alters the least perceptible portions of the video pixels, to conceal the encrypted data inside video frames.

The fourth largest cluster, i.e., cluster #3, was labeled as "Critical Infrastructure" by LLR and "Industrial Control Systems" by WoS Keyword. The most cited article was written by Ibarra et al., which discussed the impact

of ransomware in SCADA systems and ICS and its scope on Critical Infrastructure [30]. While cluster # 2 focused on the detection of ransomware, cluster #3 focused on the spread of ransomware to different networks and systems considered crucial by the Government, i.e., critical infrastructure.

The cluster #4, the fifth largest cluster, was classified as "Detecting ransomware attack" by LLR and "Random Forest" by WoS Keyword. The cluster focused on leveraging the random forest principle to detect ransomware. The most cited article was written by Khammas, which discussed the detection of ransomware using a random forest classifier [44]. The methodology involved the usage of static analysis to detect ransomware attacks efficiently. The proposed framework uses frequent data mining on raw byte data to surpass the slow disassembly process. The feature selection successfully identified 1000 features and got an accuracy of 97.74% in predicting the ransomware.

5. DISCUSSIONS AND FUTURE RESEARCH DIRECTIONS

This research paper provided a visual and systematic review of the current topics on ransomware. This bibliometric research analyzed 1200 papers on ransomware from the WoS Core Collection database from 2013 to 2023. The number of publications on ransomware remained less than 100 between 2013 and 2016. However, 2017 marked a surge in papers relating to ransomware, and the count will gradually increase to as high as 1200 by 2023. Research in ransomware was extensive and involved various subjects, including computer science information systems, computer science theory methods, engineering electrical electronics, telecommunications, and computer science artificial intelligence.

The Country and Institutional Cooperation Analysis has resulted in a deeper understanding of the top countries and institutions that play a crucial role in ransomware research. The United States of America (USA) was identified as the leading country in ransomware research. Meanwhile, in chronological order, India, England, the People's Republic of China, and South Korea were the other countries after the USA. When institutional cooperation analysis was performed, The University of Texas topped the list, followed by the University of London, the University of Texas at San Antonio UTSA, the University System of Georgia, and the Egyptian Knowledge Bank. Among the top 10 institutions, 5 out of 10 institutions came from the USA. It indicates the USA's influence on the research field.

The analysis of the author collaboration network and author co-citation network resulted in an understanding of the top authors and the authors most cited in the ransomware field. Suhyeon Lee was the principal author of ransomware as per the publication count, followed by Bander Ali Saleh Al-rimy, Julio Hernandez-Castro, Francesco Mercaldo, Josephine Wolff and so on. Meanwhile, Kharraz, et al. [24] was found to be the most cited authors with a total number of citations of 211, followed by Scaife, et al. [19]; Cabaj, et al. [21]; Al-Rimy, et al. [20] and Continella, et al. [22] and so on. It is noteworthy that Al-rimy was among both the top authors and top-cited authors. It indicates that Al-rimy remained consistent in ransomware, producing high-quality research.

The Journal Co-citation analysis gave a brief overview of the top journals in the ransomware field. IEEE Access was found to be the most influential and productive journal in terms of its total citations, number of publications, H-index, and JCI score. Lecture Notes in Computer Science came in second after IEEE Access, followed by Computers Security, International Journal of Advanced Computer Science and Applications, and finally, Sensors.

The Keyword Co-occurrence network generated by CiteSpace provided a knowledge domain map and helped to analyze the hot research topics and emerging trends in ransomware research. The top 5 keywords were: (a) Machine Learning, (b) Ransomware Detection, (c) Deep Learning, (d) Static Analysis, and (e) Malware Detection. The hot research topics need a better understanding in the future.

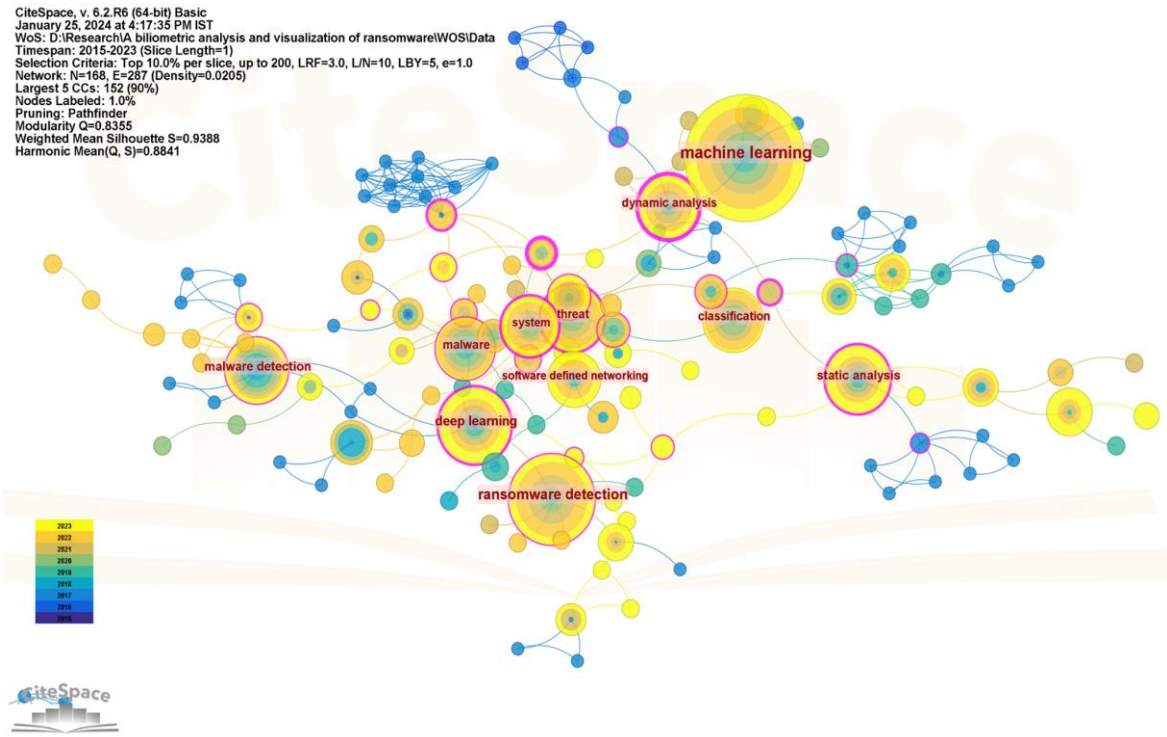


Figure 12. Network visualization map of a keyword co-occurrence network related to ransomware.

The Clustering of the Keyword Co-occurrence network generated by Citespace provided various clusters within the network, which assisted in identifying the emerging trends. Sixteen of the most significant clusters were formed and labeled per the WoS keywords. The top 5 Keywords were: (1) Software-Defined Networking (or Feedback-Based Annotated TF-IDF Technique - by LLR), (b) Static Analysis (or Hybrid Analysis - by LLR), (c) Malware Detection (or Android IoT Device - by LLR), (d)Industrial Control Systems (or Critical infrastructure -by LLR), and (e) Random Forest (or detecting ransomware attack -by LLR).

Directions for further investigations are suggested in these fields. Moreover, this paper utilized CiteSpace and Gephi to get a clear and more thorough understanding of ransomware. Figure 13 summarizes hot research topics and emerging trends in ransomware research.

Research Implications: The study's implications provide an impetus for the progress of ransomware research. The following lists a few implications.

- The current study will help the emerging researchers willing to investigate ransomware, to connect their interests and build a foundation on ransomware and its current state.
- The study provides hot research topics and emerging trends in ransomware research that will guide researchers to investigate these topics in depth.
- The research also provides insights on future research directions to guide researchers to the most challenging areas that require innovative solutions.
- Furthermore, the article could help cybersecurity engineers identify the key challenges in ransomware, and develop models and frameworks to solve them.

Research Limitations: However, this paper has several limitations. The research was limited to papers with the medium of study as English only, from the Web of Science Core Collection database. Furthermore, the utility of a single database makes the scope of research confined to a specific limit. Other databases, like Scopus, Dimensions, PubMed, etc., could have been used in addition to the Web of Science database. However, the WoS forms the largest

used database for analytics of scientific publications. Extensive research needs to be conducted on each identified hot research topic and emerging trends. Lastly, the CiteSpace Basic version had some limitations, too, abiding by the rules by which the experiment was conducted.

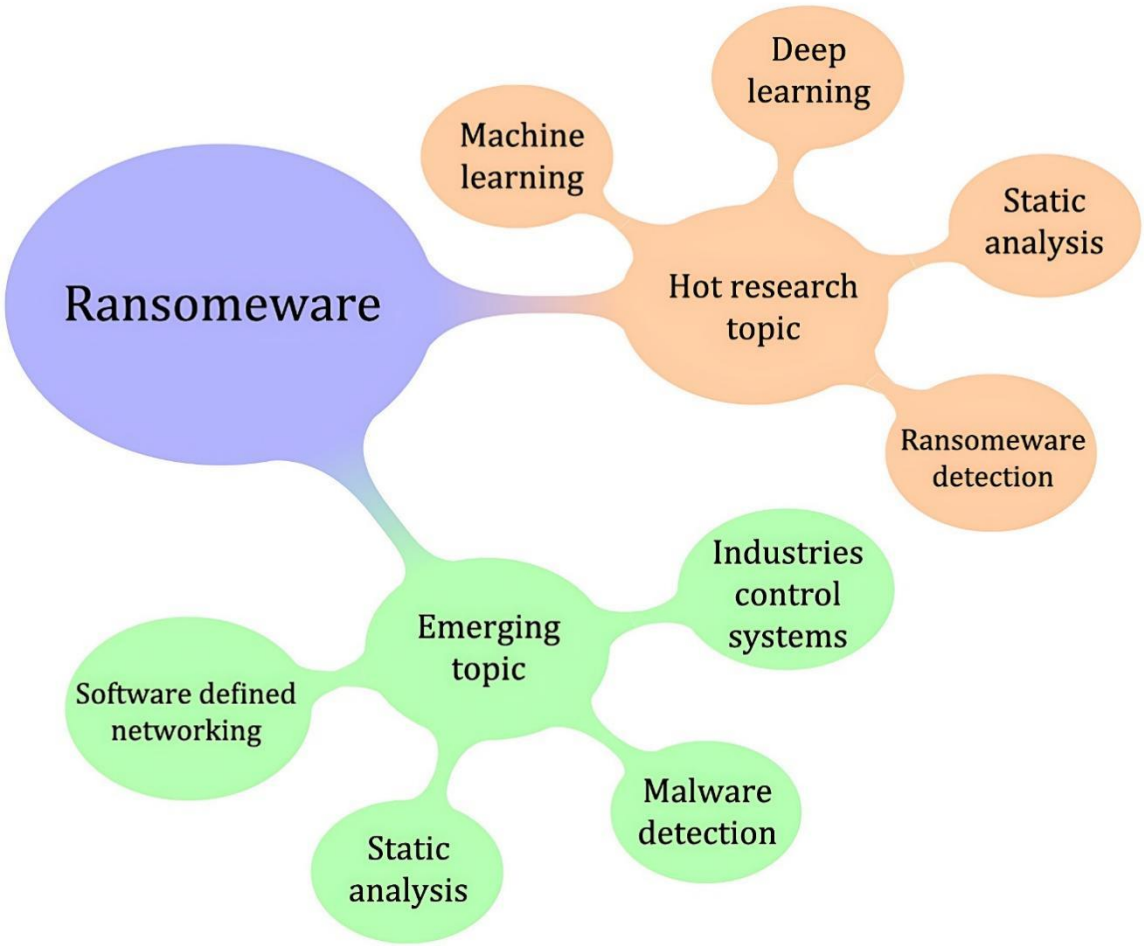


Figure 13. Hot research topics and emerging trends.

5.1. Future Research Directions

The current research paper conducts a comprehensive study on ransomware leading to several future research directions to assist cybersecurity experts and researchers in investigating further in the directions provided

- The hot topics were Machine Learning, Ransomware Detection, Deep Learning, Static Analysis and Malware Detection. More research needs to be done in ransomware detection than others due to their limitations to cyberspace like cyber attribution.
- The cluster analysis upon keyword co-occurrence network identified the emerging areas that demand more research in the context of ransomware. The fields were Software Defined Networking, Static Analysis, Malware Detection, Industrial Control Systems, and Random Forest. More research in the field of Software Defined Networking is advisable, due to security threats it pulls towards it due to centralized control planes and increased attack surface due to more communicational interfaces.
- However, Industrial Control Systems also draw attention to the need for more research into it to improve systems security while maintaining a degree of reliability and cost-efficiency and, come out of the pool of security threats and vulnerabilities that are being created due to the heavy usage of the legacy software system.

- The comprehensive study of ransomware offers a thorough analysis that will assist researchers and cybersecurity experts in bridging the knowledge gap on ransomware and investigating other areas of interest and current research topics.

6. CONCLUSION

The Current study provides the scientometrics analysis of Ransomware literature, based on 1200 publications over ten years between 2013 and 2023. Bibliometric analysis was used to evaluate contributed works in terms of annual publishing trend, country collaboration, author collaboration, author co-citation, and journal co-citation analysis. The United States of America, India, England, China, and South Korea were the top contributors in terms of publications count, however, the United States of America, England, India, Saudi Arabia, and China were among the top as per the H-index. Furthermore, the current research provides insights into the most influential journals that published prolific research on ransomware. They were identified as IEEE Access, Lecture Notes in Computer Science, Computers Security, International Journal of Advanced Computer Science and Applications, and Sensors. Moreover, the study focused on identifying the hot research topics. Five major research areas that pertain to ransomware research were noted, i.e. machine learning, ransomware detection, deep learning, static analysis, and malware detection. In addition, the research identified notable institutions that have actively contributed to ransomware research. Five institutions from the United States of America out of the top ten institutions contributed actively to ransomware research i.e. 50% of the top ten institutes were from the United States. The current study also identified the emerging trend in ransomware research through cluster analysis of the keyword co-occurrence network and provided future research directions.

LIST OF ABBREVIATIONS

All of the acronyms used in this study are listed in Table 11.

Table 11. List of abbreviations used.

Acronyms	Description	Acronyms	Description
RaaS	Ransomware-as-a-service	LSI	Latent semantic indexing
GML	Graph modeling language	MI	Mutual information
IoT	Internet of things	LSTM	Long-short term memory
WoS	Web of science	HPC	High performance computing
SDN	Software defined networking	SCADA	Supervisory control and data acquisition
HTTP	Hyper text transfer protocol	ICS	Industrial control system
LLR	Log-likelihood ratio	IDPS	Intrusion detection and prevention system
DMZ	Demilitarized zone	C&C	Command and control
TF-IDF	Term frequency - inverse document frequency	AES	Advanced encryption standard
JCI	Journal citation indicator	LSB	Least significant bit

REFERENCES

[1] H. Kettani and P. Wainwright, "On the top threats to cyber systems," presented at the 2019 IEEE 2nd International Conference on Information and Computer Technologies (ICICT), 2019.

[2] S. Adam, "The State of Ransomware, Sophos white paper," Retrieved: https://news.sophos.com/en-us/2022/04/27/the-state-of-ransomware-2022/2023.

[3] K. Mohanty, G. S. Bopche, S. Brahnam, and S. R. Dash, "Ransomware-as-a-weapon (raaw): A futuristic approach for understanding malware as a social weapon. In Contemporary Challenges for Cyber Security and Data Privacy," IGI Global, 2023, pp. 247-266.

[4] P. O'Kane, S. Sezer, and D. Carlin, "Evolution of ransomware," *Let Networks*, vol. 7, no. 5, pp. 321-327, 2018.

[5] S. Mohurle and M. Patil, "A brief study of wannacry threat: Ransomware attack 2017," *International Journal of Advanced Research in Computer Science*, vol. 8, no. 5, pp. 1938-1940, 2017.

[6] J. Berr, "WannaCry ransomware attack losses could reach $4 Billion, CBS News," Retrieved: https://www.cbsnews.com/news/wannacry-ransomware-attacks-wannacry-virus-losses/. 2017.

[7] C. P. Gibson and S. M. Banik, "Analyzing the effect of ransomware attacks on different industries," in *2017 International Conference on Computational Science and Computational Intelligence (CSCI)*, 2017: IEEE, pp. 121-126.

[8] N. G. Samy, R. Ahmad, and Z. Ismail, "Security threats categories in healthcare information systems," *Health Informatics Journal*, vol. 16, no. 3, pp. 201-209, 2010. https://doi.org/10.1177/1460458210377468

[9] M. Humayun, N. Jhanjhi, A. Alsayat, and V. Ponnusamy, "Internet of things and ransomware: Evolution, mitigation and prevention," *Egyptian Informatics Journal*, vol. 22, no. 1, pp. 105-117, 2021. https://doi.org/10.1016/j.eij.2020.05.003

[10] I. Yaqoob *et al.*, "The rise of ransomware and emerging security challenges in the Internet of Things," *Computer Networks*, vol. 129, pp. 444-458, 2017. https://doi.org/10.1016/j.comnet.2017.09.003

[11] S. I. Bae, G. B. Lee, and E. G. Im, "Ransomware detection using machine learning algorithms," *Concurrency and Computation: Practice and Experience*, vol. 32, no. 18, p. e5422, 2020.

[12] R. Brewer, "Ransomware attacks: Detection, prevention and cure," *Network Security*, vol. 2016, no. 9, pp. 5-9, 2016. https://doi.org/10.1016/s1353-4858(16)30086-1

[13] H. Oz, A. Aris, A. Levi, and A. S. Uluagac, "A survey on ransomware: Evolution, taxonomy, and defense solutions," *ACM Computing Surveys*, vol. 54, no. 11s, pp. 1-37, 2022. https://doi.org/10.1145/3514229

[14] C. Beaman, A. Barkworth, T. D. Akande, S. Hakak, and M. K. Khan, "Ransomware: Recent advances, analysis, challenges and future research directions," *Computers & Security*, vol. 111, p. 102490, 2021. https://doi.org/10.1016/j.cose.2021.102490

[15] D. Garg, A. Thakral, T. Nalwa, and T. Choudhury, "A past examination and future expectation: Ransomware," presented at the 2018 International Conference on Advances in Computing and Communication Engineering (ICACCE), 2018.

[16] A. F. Van Raan, "For your citations only? Hot topics in bibliometric analysis," *Measurement: Interdisciplinary Research and Perspectives*, vol. 3, no. 1, pp. 50-62, 2005. https://doi.org/10.1207/s15366359mea0301_7

[17] M. Ryan, *Ransomware case studies. In: Ransomware Revolution: The Rise of a Prodigious Cyber Threat," Advances in Information Security*. Cham: Springer. https://doi.org/10.1007/978-3-030-66583-8_5, 2021, pp. 65-91.

[18] A. Kharaz, S. Arshad, C. Mulliner, W. Robertson, and E. Kirda, "{UNVEIL}: A {Large-Scale}, automated approach to detecting ransomware," presented at the 25th USENIX Security Symposium (USENIX Security 16), 2016.

[19] N. Scaife, H. Carter, P. Traynor, and K. R. Butler, "Cryptolock (and drop it): Stopping ransomware attacks on user data," presented at the 2016 IEEE 36th International Conference on Distributed Computing Systems (ICDCS), 2016.

[20] B. A. S. Al-Rimy, M. A. Maarof, and S. Z. M. Shaid, "Ransomware threat success factors, taxonomy, and countermeasures: A survey and research directions," *Computers & Security*, vol. 74, pp. 144-166, 2018. https://doi.org/10.1016/j.cose.2018.01.001

[21] K. Cabaj, M. Gregorczyk, and W. Mazurczyk, "Software-defined networking-based crypto ransomware detection using HTTP traffic characteristics," *Computers & Electrical Engineering*, vol. 66, pp. 353-368, 2018. https://doi.org/10.1016/j.compeleceng.2017.10.012

[22] A. Continella *et al.*, "Shieldfs: A self-healing, ransomware-aware filesystem," presented at the Proceedings of the 32nd Annual Conference on Computer Security Applications, 2016.

[23] T. Yang, Y. Yang, K. Qian, D. C.-T. Lo, Y. Qian, and L. Tao, "Automated detection and analysis for android ransomware," presented at the 2015 IEEE 17th International Conference on High Performance Computing and Communications, 2015 IEEE 7th International Symposium on Cyberspace Safety and Security, and 2015 IEEE 12th International Conference on Embedded Software and Systems, 2015.

[24] A. Kharraz, W. Robertson, D. Balzarotti, L. Bilge, and E. Kirda, "Cutting the gordian knot: A look under the hood of ransomware attacks," presented at the Detection of Intrusions and Malware, and Vulnerability Assessment: 12th International Conference, DIMVA 2015, Milan, Italy, July 9-10, 2015, Proceedings 12, 2015.

[25] N. Andronio, S. Zanero, and F. Maggi, "Heldroid: Dissecting and detecting mobile ransomware," presented at the Research in Attacks, Intrusions, and Defenses: 18th International Symposium, RAID 2015, Kyoto, Japan, November 2-4, 2015. Proceedings 18, 2015.

[26] Y. Ye, T. Li, D. Adjeroh, and S. S. Iyengar, "A survey on malware detection using data mining techniques," *ACM Computing Surveys*, vol. 50, no. 3, pp. 1-40, 2017.

[27] S. Razaulla *et al.*, "The age of ransomware: A survey on the evolution, taxonomy, and research directions," *IEEE Access*, vol. 11, pp. 40698-40723, 2023. https://doi.org/10.1109/access.2023.3268535

[28] M. A. Ayub, A. Siraj, B. Filar, and M. Gupta, "RWArmor: A static-informed dynamic analysis approach for early detection of cryptographic windows ransomware," *International Journal of Information Security*, vol. 23, no. 1, pp. 533-556, 2024. https://doi.org/10.1007/s10207-023-00758-z

[29] I. Almomani, A. Alkhayer, and W. El-Shafai, "A crypto-steganography approach for hiding ransomware within HEVC streams in android IoT devices," *Sensors*, vol. 22, no. 6, p. 2281, 2022. https://doi.org/10.3390/s22062281

[30] J. Ibarra, U. J. Butt, A. Do, H. Jahankhani, and A. Jamal, "Ransomware impact to SCADA systems and its scope to critical infrastructure," presented at the 2019 IEEE 12th International Conference on Global Security, Safety and Sustainability (ICGS3), 2019.

[31] I. El Naqa and M. J. Murphy, *What is machine learning? Machine Learning in Radiation Oncology*. Cham: Springer. https://doi.org/10.1007/978-3-319-18305-3_1, 2015.

[32] S. K. Shaukat and V. J. Ribeiro, "RansomWall: A layered defense system against cryptographic ransomware attacks using machine learning," presented at the 2018 10th International Conference on Communication Systems & Networks (COMSNETS), 2018.

[33] S. Poudyal, K. P. Subedi, and D. Dasgupta, "A framework for analyzing ransomware using machine learning," presented at the 2018 IEEE Symposium Series on Computational Intelligence (SSCI), 2018.

[34] D. Kim and J. Lee, "Blacklist vs. whitelist-based ransomware solutions," *IEEE Consumer Electronics Magazine*, vol. 9, no. 3, pp. 22-28, 2020. https://doi.org/10.1109/mce.2019.2956192

[35] S. Bhardwaj and M. Dave, "Integrating a rule-based approach to malware de- tection with an lstm-based feature selection technique," *SN Computer Science*, vol. 4, no. 6, p. 737, 2023. https://doi.org/10.1007/s42979-023-02177-2

[36] P. Shijo and A. Salim, "Integrated static and dynamic analysis for malware detection," *Procedia Computer Science*, vol. 46, pp. 804-811, 2015. https://doi.org/10.1016/j.procs.2015.02.149

[37] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436-444, 2015.

[38] G. O. Ganfure, C.-F. Wu, Y.-H. Chang, and W.-K. Shih, "Deepware: Imaging performance counters with deep learning to detect ransomware," *IEEE Transactions on Computers*, vol. 72, no. 3, pp. 600-613, 2022. https://doi.org/10.1109/tc.2022.3173149

[39] S. Maniath, A. Ashok, P. Poornachandran, V. Sujadevi, P. S. AU, and S. Jan, "Deep learning LSTM based ransomware detection," presented at the 2017 Recent Developments in Control, Automation & Power Engineering (RDCAPE), 2017.

[40] B. Chess and G. McGraw, "Static analysis for security," *IEEE Security & Privacy*, vol. 2, no. 6, pp. 76-79, 2004.

[41] T. Ball, "The concept of dynamic analysis," *ACM SIGSOFT Software Engineering Notes*, vol. 24, no. 6, pp. 216-234, 1999.

[42] R. Chanajitt, B. Pfahringer, and H. M. Gomes, "Combining static and dynamic analysis to improve machine learning-based malware classification," presented at the 2021 IEEE 8th International Conference on Data Science and Advanced Analytics (DSAA), 2021.

[43] S. Banescu, T. Wuchner, A. Salem, M. Guggenmos, M. Ochoa, and A. Pretschner, "A framework for empirical evaluation of malware detection resilience against behavior obfuscation," in *2015 10th International Conference on Malicious and Unwanted Software (MALWARE)*, 2015: IEEE, pp. 40-47.

[44] B. M. Khammas, "Ransomware detection using random forest technique," *ICT Express*, vol. 6, no. 4, pp. 325-331, 2020. https://doi.org/10.1016/j.icte.2020.11.001