Winter 11-2019

Global Research Trend on Cyber Security: A Scientometric Analysis

Somesh Rai
Department of Library and Information Science, Banaras Hindu University, somesh.rai15@bhu.ac.in

Kunwar Singh Dr
Department of Library and Information Science, Banaras Hindu University, singhdlisbhu2015@gmail.com

Akhilesh Kumar Varma
Department of Library and Information Science, Banaras Hindu University

Global Research Trend on Cyber Security: A Scientometric Analysis

Somesh Rai[1], Dr. K. Singh[2] & A. K. Varma[3]

ABSTRACT

Scientometrics is a quantitative analysis of scholarly literature related to a particular subject or area (well defined by some limits, scope and coverage), which helps to understand different aspects about the scholarly literature's growth in various dimensions of knowledge. Similarly, this study is a quantitative analysis of the Global research trends in cyber security. Some works related to scientometrics of 'deception, counter-deception in cyberspace' had been published in 2011, but we have focused on 'cyber security' as the topic of research. For analysis we have utilised the published data available in Scopus database, which is directly related to 'cyber security'. The data of 2720 documents published on 'cyber security' for the period of 2001-2018, have been retrieved from the Scopus database. We have calculated and tabulated various quantitative indicators from the data; then, we have plotted different graphs for these indicators to understand the growth pattern, collaborations, citations, authorship, countrywise contributions, funding, affiliations etc. We found that the United States of America has the highest number of publications in the area of cyber security research followed by the United Kingdom, China, and India. It can be seen as a result of the fact that among the top 15 funding agencies eight funding agencies are from the USA only. Contrary to this, India is the fourth highest contributor in cyber security research with comparatively meagre funding. Further we have also discussed how collaboration has grown over the years with the publication growth and citation growth. We have also found that Cybersecurity has a growing trend of collaboration in research. This study concludes that Indian defence need to focus on Cybersecurity and plan strategically for the future with effective collaborations. Best defence strategy in the cyberspace is preventive measures supported by continuous research.

Keywords: *Cyber security, Scientometric Analysis, Bibliometrics, Collaboration, Authorship pattern, Author Productivity, Research Funding.*

1. Introduction

At present we live in a world which is continuously being highly embedded through cyberspace. The term 'Cyber Security' has gained initial popularity over the span of 18 years(1990-2008) ("Google Ngram Viewer," n.d.). Afterwards, according to Google Trends worldwide interest ("Google Trends," n.d.) has been increasing for the term 'Cyber Security', and a spike appears in september 2009. It is coincident with increased cyber attacks worldwide. It is intuitive to assume it as a cause to increase research activity.

Cyber security can be termed simply as security of cyber infrastructures against adverse events. Craigen et al (2014) have provided a detailed study with exhaustive literature review to define the term 'cyber security'.We are at a time where individuals as well as organisations are vulnerable in terms of 'Privacy Intrusion', 'Identity Theft', 'Data Theft' etc.. Cyber Security covers all of the security related concern throughout cyberspace. Because of the increased internet penetration, and use of computing infrastructure in almost every personal and business work, we operate in a highly interconnected computer environment. In this process, Sharing and Storing our data is a necessity of present times, so it can not be avoided but a default risk is associated with these processes and involved systems. Moreover, when individuals or organisations process their information through the web without taking proper

[1] Junior Research Fellow, Dept. of Library & Information Science, Banaras Hindu University, Varanasi - 221005
E-mail: somesh.rai15@bhu.ac.in
[2] Assistant Professor, Dept. of Library & Information Science, Banaras Hindu University, Varanasi - 221005
E-mail: singhdlisbhu2015@gmail.com
[3] Junior Research Fellow, Dept. of Library & Information Science, Banaras Hindu University, Varanasi - 221005

updated security measures, they end up powerless against digital assaults and cyberattacks in terms of control over their own information and cyber-systems. Besides, digital assaults can occur easily against any type of organisation to any ICT infrastructure which operates through interconnected networks. These assaults can be over an outside confronting DNS server or an inward firewall, which thusly impacts the information and framework inside the undertaking that intrinsically & critically harms the matter of the related association also. For example Power Grids and Banking Systems are two sectors, which had been most concerned since the dawn of computers application with interconnected infrastructure ("AIG Study," 2017). Both of these types of systems have seen assaults internally as well as externally. Besides these manufacturing sector (Wells et al, 2014) and the health sector (Martin et al, 2017) is also vulnerable in terms of cyber security; the futuristic smart cities incorporates all these establishments, which leaves their cyberspaces also vulnerable (Elmaghraby & Losavio, 2014). So, it is very much evident that the topic 'Cyber Security' bears a high degree of importance in our present world; and that is why computer engineers and scientists have been consistently working in this area over the last two decades. Choo (2011) has discussed the challenges and need of future technical research in the cyber security domain. To understand the impact and magnitude of the world wide research on cyber security, we turn to the scientometric analytics. Scientometrics is a quantitative method to study and estimate scholarly research output in terms of publications and their existence in the world of knowledge. In our paper we have assessed some quantitative indicators to derive patterns of the research growth and interpret that growth with other factors.

2. Literature Review

Scientometrics is a quantitative analytics, which is used to estimate and analyse the amount of research in any particular subject area; In Science community scholars rely on different publication mediums to communicate their ideas and discovery. There are some scientometric studies related to cyber security, most of them have focused on information security which is a related topic. In the very first scientometric study on information security domain, Lee (2008) presented a way to identification of new emerging fields by performing a scientometric assessment in information security domain. Olijnyk (2014) submitted a dissertation to the Faculty of Long Island University, which aimed at exploring and describing the profile, dynamics, and structure of the information security domain. Olijnyk (2015) further studied long term development of information security domain in the paper titled "A quantitative examination of the intellectual profile and evolution of information security from 1965 to 2015". Some scientometric assessment has been performed in ambit of cyber-Security also with specific purpose of topic modelling e.g. Chang (2016) attempted to model topics and development directions for cyber security and information security research.

Whereas, In this paper we are assessing literature growth, collaboration, subject wise distribution, affiliations and funding pattern along with country wise contributions in cyber security domain specifically. To accomplish these we resort to standard methods used by various scientometricians. Elango & Rajendran (2012) performed an assessment of Indian Journal of Marine Sciences published from 2001 to 2010. They utilised collaboration index and collaboration coefficient measures to estimate collaboration between authors. The Average Collaboration rate of 0.57 was acknowledged as the good collaboration among the authors. Poornima et al (2011) reported an analysis of 1060 articles published during 1998 to 2010 and indexed in Web of Science Database. It was observed from the study that *multi authored* publication pattern is popular among the authors in Food Science and Technology research as 95.94% of publications made by multiple authors. Multi Authorship pattern shows collaboratory practice among the authors of a particular research area. Velmurugan & Radhakrishnan (2016) had investigated for research productivity of Malaysian Journal of Library and Information Science for timespan of 2008 to 2014. It was found that the average author per paper was 2.36 and average productivity per author was

0.42. The highest author productivity was 74 (2.64) reported for the year 2011. Author productivity can be calculated for various groups (created by imposing different criteria e.g. source, subject, keyword, scope, coverage etc.) of documents; if we consider documents published on a particular topic, it gives the information about productivity on that topic of research.

3. Objectives

The prime objective of this study is to understand the growth pattern of the scholarly literature on the term 'cyber security'. This pattern constitutes of the information related to author collaboration, funding, author productivity etc. Our objective of this scientometric analysis is to:

- To identify the growth pattern of the scholarly literature directly related to cyber security'.
- To understand the collaborative pattern, subject wise distribution of the publications.
- To assess the author productivity, author preferred documents and subjects in cyber security related research.
- To know the citation pattern over the years.
- To identify the institutions, agencies, affiliations and countries active in the field of cyber security.

4. Scope and Coverage of this Study

We have performed this study for the coverage period of 18 years (2001 to 2018). We have focused on the scholarly literature directly related to the term 'cyber security', which is indexed in Scopus database.

5. Methodology and Data Collection

Scopus database provides comprehensive bibliographic datasets covering various aspects of scholarly communications, which have high utility for a Scientometric Analysis. So, We collected the Data from the Scopus database for the coverage of 2001 to 2018. The string 'Cyber' AND 'Security' was searched in the 'title' field of the Scopus database to retrieve the bibliographic data of publications directly related to cyber security. Formulated Search query string is:

TITLE (cyber AND security) AND (EXCLUDE (PUBYEAR , 2020) OR EXCLUDE (PUBYEAR , 2019) OR EXCLUDE (PUBYEAR , 1999) OR EXCLUDE (PUBYEAR , 2000)) AND (LIMIT-TO (LANGUAGE , "English")) AND (EXCLUDE (PUBSTAGE , "aip"))

A total 2720 records were exported from the database in May 2019. Further calculation for different key indicators have been performed. These indicators have been tabulated and presented in this paper, and It has been used for graph-plotting and analysis. To accomplish our objectives we have utilised these measures (Indicators) which are explained in detail under corresponding sub-headings. These are briefed below:

1. Average annual growth (AGR), Relative annual growth (RGR) and distribution of publications over the years have been tabulated, plotted and analyzed to understand the growth pattern of the .
2. To assess collaborations we have calculated measures like collaboration index (CI), collaboration coefficient (CC), degree of collaboration (DC). Further we have found subject wise contributions.
3. We have tabulated the Author productivity and types of documents.
4. We have calculated average citations and Distributed CiteScores(DCS) to asses Citation growth over the years.
5. In this subsection we have analysed Country wise contributions, Sources which published research on cyber security, Funding Agencies and Affiliations.

6. Data Tabulation and Plotting

6.1 Year wise distribution and growth of publications

Table 1 shows, a total number of 2720 articles are published during 2001-2018. From the table we can say that no. of publications have been increasing monotonically over the taken period, and about 69.18% of the total publications are contributed in the last 4 years only. There were only 3 articles in 2001 and since then a continuous growth of publications is observed during 2001-2018, only a dip appears in the year 2008.

Table 1 : Year wise distribution of article

Year	Documents	%
2001	3	0.11
2002	6	0.22
2003	12	0.44
2004	19	0.70
2005	29	1.07
2006	42	1.54
2007	50	1.84
2008	38	1.40
2009	72	2.65
2010	88	3.24
2011	126	4.63
2012	157	5.77
2013	226	8.31
2014	242	8.90
2015	320	11.76
2016	387	14.23
2017	416	15.29
2018	487	17.90
Total	**2720**	**100.00**

6.1.1 Annual growth rate RGR

Scientometricians frequently use two measures to assess growth rate of literature in any field. First one is Annual Growth Rate (AGR) and the Second one is Relative Growth Rate (RGR). These growth rates are often a measure of the annual increase or decrease in the number of publications in a particular discipline. AGR is determined as per the formula given below. The formula is given by:

$$AGR = \frac{N_p - N_t}{N_p} \times 100 \quad \textbf{......................(1)}$$

Where, N_p = Number of total Publications in *Previous Year*

N_t = Number of total Publications in *Present Year*

The second measure of growth RGR is calculated as the difference in natural logarithms of total number of publications at two points of time divided by the time interval. Which is:

$$RGR = (1-2^r) = \frac{ln(W2) - ln(W1)}{T_2 - T_1} \quad \text{........................(2)}$$

Where,

W1 = Total Number of Publications at *Initial time.*

W2 = Total Number of Publications at *Final time.*

$T_2 - T_1$ = Difference between the initial year and the final year.

We can find Doubling Time of the published literature, and it is equal to the natural logarithm of 2, divided by RGR.

$$Doubling\ Time = D(t) = \frac{0.693}{RGR} \quad \text{........................(3)}$$

Table 2 shows the AGR and RGR of the number of publications in the period between 2001 and 2018.

Table 2 : Annual Growth Rate (AGR) and Relative Growth Rate (RGR)

Year	Documents	Cumulative total	AGR	W1	W2	RGR	D(t)
2001	3	3	-	-	1.10	-	
2002	6	9	100.00	1.10	2.20	1.0987	0.63
2003	12	21	100.00	2.20	3.04	0.8473	0.82
2004	19	40	58.33	3.04	3.69	0.6444	1.08
2005	29	69	52.63	3.69	4.23	0.5453	1.27
2006	42	111	44.83	4.23	4.71	0.4755	1.46
2007	50	161	19.05	4.71	5.08	0.3719	1.86
2008	38	199	-24.00	5.08	5.29	0.212	3.27
2009	72	271	89.47	5.29	5.60	0.3089	2.24
2010	88	359	22.22	5.60	5.88	0.2813	2.46
2011	126	485	43.18	5.88	6.18	0.3009	2.30
2012	157	642	24.60	6.18	6.46	0.2805	2.47
2013	226	868	43.95	6.46	6.77	0.3017	2.30
2014	242	1110	7.08	6.77	7.01	0.246	2.82
2015	320	1430	32.23	7.01	7.27	0.2534	2.73
2016	387	1817	20.94	7.27	7.50	0.2396	2.89
2017	416	2233	7.49	7.50	7.71	0.2062	3.36
2018	487	2720	17.07	7.71	7.91	0.1973	3.51
Total	**2720**	**CAGR= 0.46**			**Mean RGR= 0.4006**		**1.73**

W1 = Natural logarithms of no. of paper published until previous year

W2 = Natural logarithms of no. of paper published until present year

D(t)= Doubling Time,

AGR= Annual Growth Rate, RGR= Relative Growth Rate

Figure 1.1 shows the plot of annual growth rate for the corresponding Year. It can be seen that the AGR has been monotonously decreasing since 2003 and reaches a minima in the year 2008 with -24% AGR and next year in 2009 it peaked to 89.47%, with a total change of 113.47%. Then 2010 onwards AGR has frequently varied between 15% to 45% only.

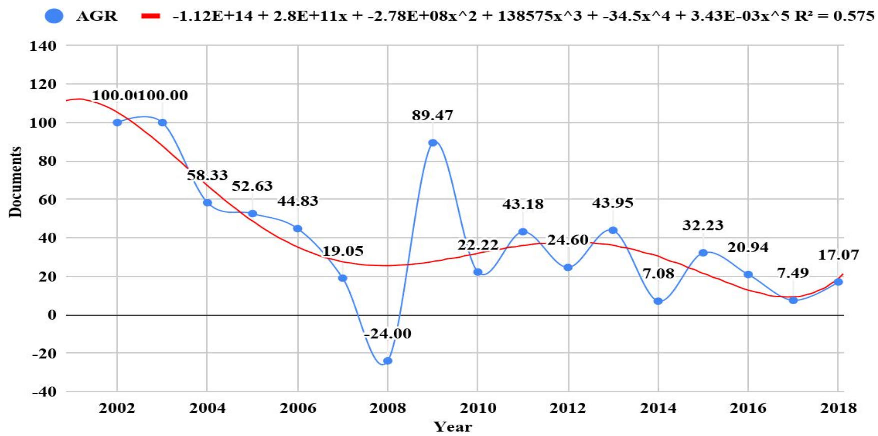

Fig. 1.1 (Documents, AGR vs Year)

We attempted to find a polynomial regression and we found that, although a fifth degree polynomial fits the data with a regression coefficient (R^2) of 0.575, but it can not be reliably extrapolated to get the AGR for upcoming years. If plotted, we get a steeply rising change in AGR, which is contradictory to the fact that publications in cyber security domain are influenced by the contemporary events in the cyberspace and world wide interest. Worldwide interest can change overtime because of different cyber security related activities e.g. cyber attacks, data breaches, cyber war etc. and it is intuitive from the previous patterns also. It can be concluded that the growth of the literature does not have a predictable trend by means of AGR.

In the figure 1.2 year wise variation of RGR has been presented, and a good trend is visible with a strong regression coefficient of 0.825. RGR is exponentially decreasing over the years. So, we conclude that RGR gives a better view of assessing the literature growth.

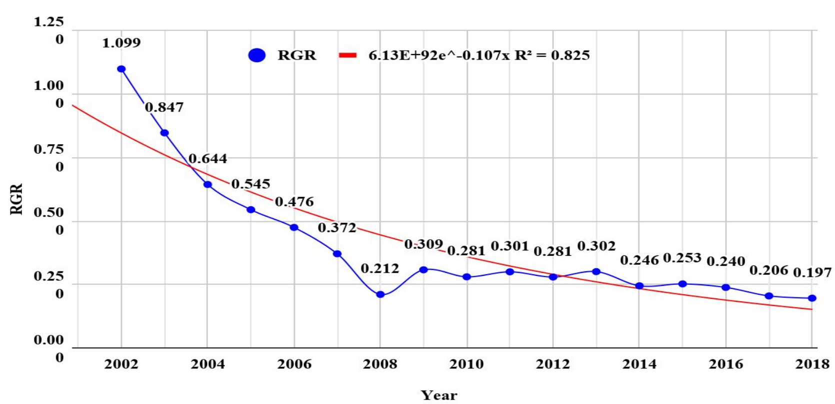

Fig. 1.2 (RGR vs Year)

6.2 Collaboration and Subject Wise Distribution

6.2.1 Degree of Collaboration

The DC is defined as the ratio of the number of collaborative research papers to the total number of research papers in the discipline during a certain period of time. The formula suggested by Subramanyam (1983) is used. It is expressed as:

$$DC = \frac{N_m}{N_m+N_s} \text{ ……………………(4)}$$

Where, *DC* – is the degree of collaboration in a discipline,

*N*m - Is the number of multi-authored research papers in the discipline published during the year.

*N*s - Is the number of single-authored papers in the discipline published during the same year.

Table 3 Degree of Collaboration and Collaboration Index

Year	Ns	Nm	Ns+Nm	DC	CI
2001	2	1	3	0.33	2
2002	4	2	6	0.33	1.5
2003	7	5	12	0.42	2.42
2004	11	8	19	0.42	1.63
2005	13	16	29	0.55	1.9
2006	15	27	42	0.64	2.21
2007	25	25	50	0.5	2.52
2008	16	22	38	0.58	2.58
2009	31	41	72	0.57	2.19
2010	25	63	88	0.72	2.77
2011	39	87	126	0.69	2.44
2012	57	100	157	0.64	2.69
2013	76	150	226	0.66	2.52
2014	79	163	242	0.67	2.5
2015	103	217	320	0.68	2.74
2016	127	260	387	0.67	2.64
2017	93	323	416	0.78	2.94
2018	82	405	487	0.83	3.11
Total	**805**	**1915**	**2720**	**0.7**	**2.72**

Ns = Number of Single Authored Papers, Nm = Number of Multi-Authored Papers,

Ns+Nm = Total Number of Papers

Table 3 shows the degree of collaboration (DC) and collaboration index (CI), which are collaboration measures derived from number single authored and multiple authored publications only. Figure 2.1 shows the variation of DC over the years. It is observed that DC has been above 0.5 since 2005.

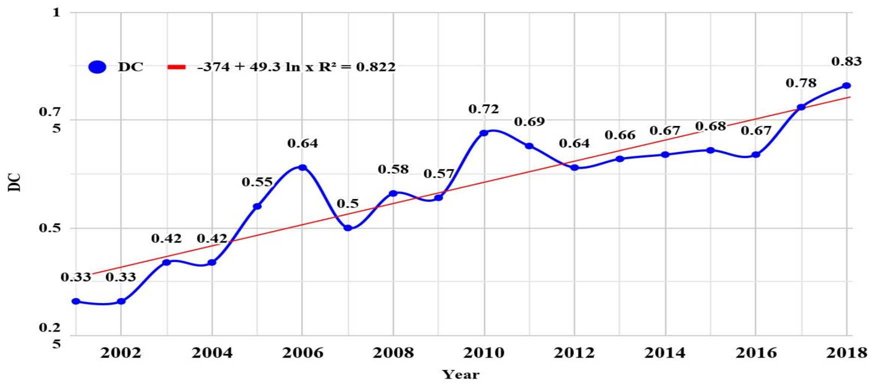

Fig. 2.1 DC vs Year

6.2.2 Collaboration Index (CI)

Collaboration index is the mean number of authors per joint paper. For this analysis, we have to omit the single authored papers which are equal to 1 always. To determine the mean number of authors per joint authored paper, the following formula has been used.

$$CI = \frac{Total\ No.\ of\ Authors}{Total\ no.\ of\ joint\ papers} \quad \textbf{......................(5)}$$

Table 3 provides the year wise mean number of authors per joint authored paper. CI ranges from 1.5 (2002) to 3.11 (2018) with an average of 2.72 per joint authored paper.

Figure 2.2 shows the variation of CI over the years. It is observed that CI plot is relatively smoother than the DC plot; which occurs because of the less complexity of the CI formulae.

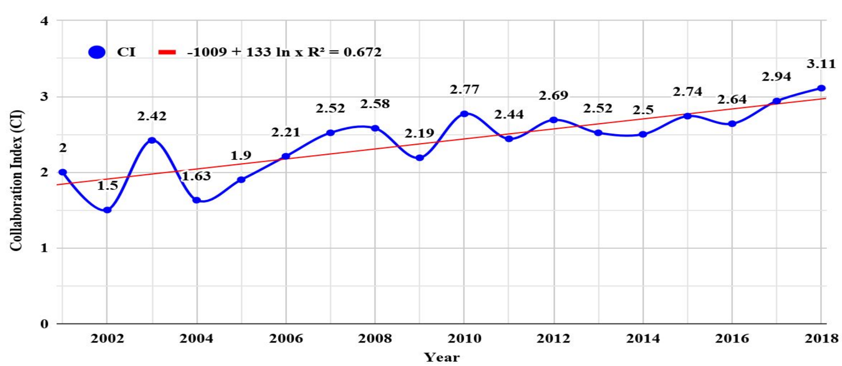

Fig. 2.2 CI vs Year

6.2.3 Collaboration Coefficient (CC)

To measure the strength of collaboration the following formula of collaboration coefficient as suggested by Ajiferuke, Burell, & Tague (1988) has been used.

$$CC = 1 - \frac{\sum_{j=1}^{k}(\frac{1}{j})f_j}{N} \quad ....(6)$$

Where; f_j = Total number of j authored research papers

N = Total number of research papers published in a year
k = The greatest number of authors per paper

Collaboration Coefficient is a numerical value between 0 and 1. The more it is bigger than 0.5 the better is the collaboration rate among the authors. When it is near 0, it means that authors have a weak collaboration rate.

Table 4 Shows year wise values of the collaboration coefficient, it is calculated by the formulae (6) which discreetly accounts various number of authors' contribution to a single publication. It implies that collaboration coefficient is a measure which takes a more detailed account of multiple authorships in comparison to DC and CI.

Table 4 Collaboration Coefficient (CC)

Year	Author										Total	CC
	1	2	3	4	5	6	7	8	9	10 +		
2001	2	0	0	1	0	0	0	0	0	0	3	0.25
2002	4	0	2	0	0	0	0	0	0	0	6	0.223
2003	7	2	1	1	0	0	0	0	1	0	12	0.276
2004	11	4	4	0	0	0	0	0	0	0	19	0.246
2005	13	11	1	3	1	0	0	0	0	0	29	0.318
2006	15	13	8	3	2	1	0	0	0	0	42	0.394
2007	25	9	7	4	3	0	0	1	0	1	50	0.327
2008	16	9	4	4	2	2	1	0	0	0	38	0.377
2009	31	16	15	5	3	0	1	1	0	0	72	0.36
2010	25	22	19	7	7	4	2	1	1	0	88	0.47
2011	39	34	24	23	3	1	1	1	0	0	126	0.439
2012	57	37	27	15	6	7	3	1	1	3	157	0.417
2013	76	51	46	30	12	7	2	1	1	0	226	0.432
2014	79	65	50	20	16	6	1	4	0	1	242	0.43
2015	103	64	60	43	20	17	7	4	2	0	320	0.456
2016	127	88	71	45	35	8	6	5	1	1	387	0.443
2017	93	92	98	74	27	17	9	2	3	1	416	0.519
2018	82	111	112	89	51	31	8	3	0	0	487	0.561
Total	**805**	**628**	**549**	**367**	**188**	**101**	**41**	**24**	**10**	**7**	**2720**	**0.464**
%	**29.60**	**23.09**	**20.18**	**13.49**	**6.91**	**3.71**	**1.51**	**0.88**	**0.37**	**0.26**	**100**	

Figure 2.3 shows variation of collaboration coefficient. It is found that the collaboration has an increasing trend of logarithmic nature with good regression coefficient of 0.866. It can also be seen from the Table 4 that since 2016 single authorship has declined and the joint authorship consisting of 2, 3, 4, 6 number of authors have increased in number of publications. It can be predicted from the trend that the cyber security will have more jointly authored publications than single authored. In 2018 CC is found to be 0.561 and overall collaboration until 2018 is 0.464.

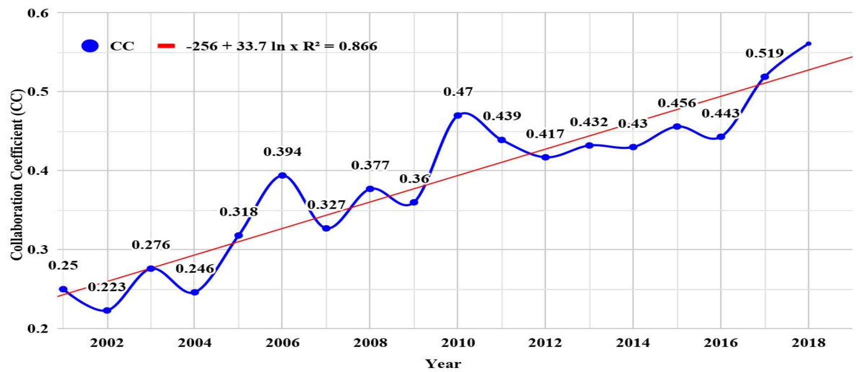

Fig 2.3 CC vs Year

6.2.4 Comparison of different Collaborative Measures

We were not sure which collaborative measure gives a better picture of the varying collaboration in the field; So, we have assessed three of them and after comparison we found that the collaboration coefficient(CC) is the best out of these three measures. It can be understood by comparing the plots in figure 2.4; the CC plot appears more frequently varying and it presents more details because of the fact that it utilizes detailed calculations for different type of joint authorship. Furthermore CC is found to be more informing than others as the values varies on a fixed scale of 0 to 1.

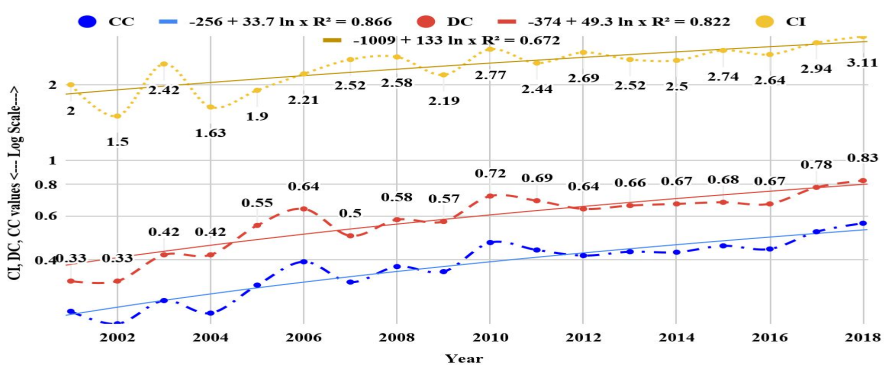

Fig. 2.4 CC, DC and CI Comparison

6.2.5 Single and Co-authorship Pattern

Table 3 shows that the details about the authorship pattern of single and joint contributions during the period of 2001 to 2018. The maximum number, 1915 (70.41%) contributions were by joint authors and the rest of 805 (29.59%) contributions were by a single author.

Figure 3.3 shows the growing number of jointly authored publications and number of single authored publications with dotted line and dashed line respectively. In this figure we have also plotted the regression functions corresponding to the Single authorship and joint authorship growth over time. We found that Joint authorship growth has an exponential regression function with R^2 (regression coefficient) equals to 0.992, whereas single authorship has a growth pattern of logarithmic function with regression

coefficient of 0.843. Here it is inferred that the growth of total documents over the year is mostly being contributed by the jointly authored papers, as the single authored papers are varying logarithmically.

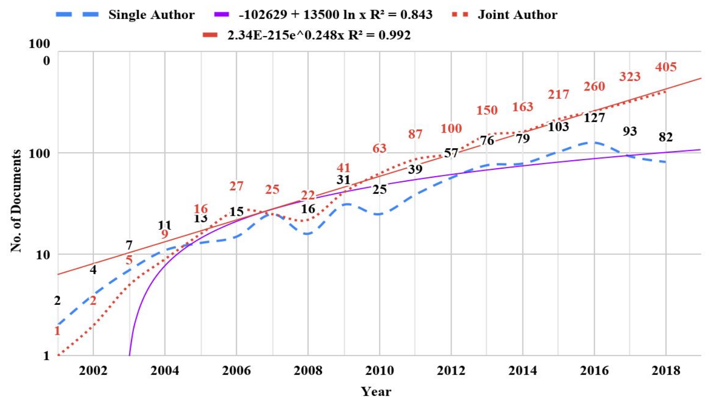

Fig. 2.5 Single Author and Joint Author (Log Scale Plot)

6.3 Productivity and Preferences

6.3.1 Author productivity

Author productivity is a measure of quantitative efficiency of the authors in a particular research domain; it shows that, to produce one publication on an average how many Scholars are involved. Quantitative author productivity is calculated as the total number of papers published divided by the total number of authors; which is nothing but the inverse of average number of authors per paper. Average author per paper in a year can be given by this formula.

$$AAP = \frac{Sum\ of\ Total\ Authors\ of\ each\ paper\ in\ a\ Year}{Total\ No\ of\ Papers\ published\ in\ a\ Year} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(7)$$

AAP = Average author per paper.

Figure 3.1 shows the variation of average number of authors per paper (AAP) over the years. We observed that the plot of AAP is very similar to the plot of CI; and a high negative correlation is found between the two with a correlation coefficient of 0.978. It is found that the average number of authors per paper has an increasing trend. Which is representative of the qualitative productivity of authors, if increase in collaboration is seen along with increase in citations over the years.

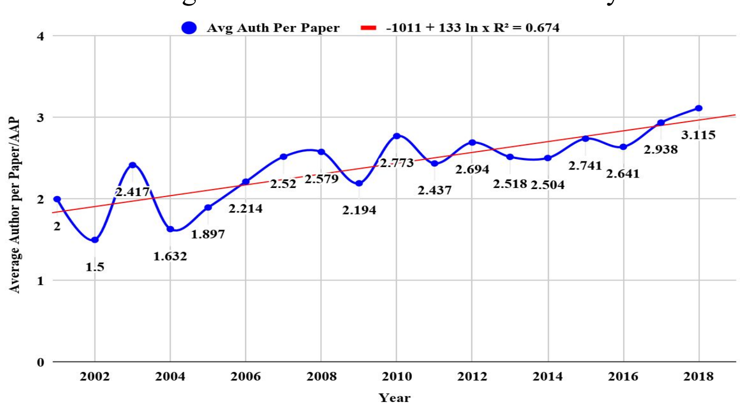

Fig. 3.1 (AAP vs. Year)

6.3.2 Type of Document

Figure 3.2 represents number of documents corresponding to two types of documents used as most preferred medium of publication. Among the total number of 2720 documents, 1561 documents are conference papers and 656 documents are published as articles. Conference papers and articles are the two most preferred medium of publication with a combined percentage of 81.51 percent, shown with blue and red pie-slice.

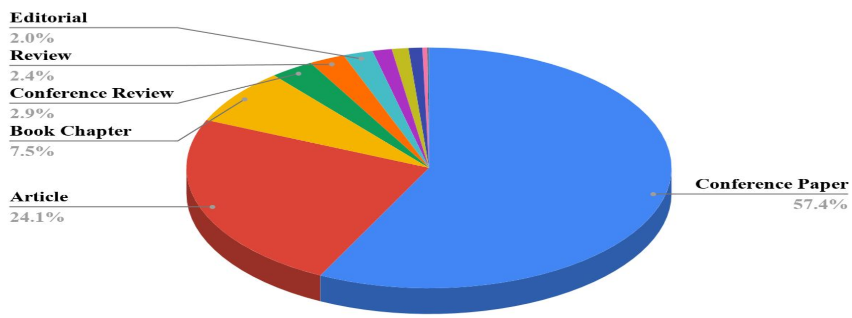

Fig. 3.2 Pie Distribution of Types of Documents

6.3.3 Subject wise distribution of articles

Figure 3.3 shows the subject wise distribution of publications which were produced during the stipulated period. This study identifies the authors' interest and involvement of subjects in terms of producing the publication in their respective specialization. The findings of the study reveal that the highest number 1324 (36.80%) of scientific scholarly publications were published in the subject of Computer science study due to the rapid growth of development in the area that the majority of authors are very much interested to do their research work and followed by 930 (25.85) of papers were from Engineering and 290(8.06%) of papers were from Social Sciences and 248(6.89) articles were published from Mathematics etc.

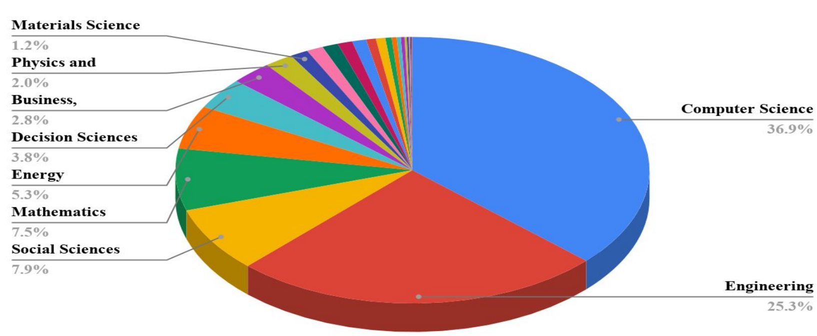

Fig. 3.3 Subject wise Pie-chart

6.4 Yearwise Citations

Table 5 represents year wise percentage of documents which has been cited at least once, total cites received for that year and Average Citations per Document (ACD) for each year; we have also calculated a measure Distributed CiteScore (DCS) for each year, distributed citescore is calculated by this formula given below:

$$DCS = \frac{Average\ no.\ citation\ per\ document}{Number\ of\ Years\ documents\ were\ in\ published\ domain} \quad \text{........................}\mathbf{(8)}$$

In calculating DCS we have incorporated the effect of obsoletion of a document over the years, if a document is published in very early stage but it is still getting citation, then only its DCS will remain constant otherwise it will tend to zero over the upcoming years. It is a good measure to evaluate the impact of an old paper relative to newly published papers.

There are total 15059 citations received by total 2720 articles published between 2001 and 2018, moreover we have taken citations till **May 2018.** Further we have also calculated a 'Distributed CiteScore' which incorporates the weight of timespan for which documents have been in published domain.

Table 5 Citation Analysis

Year	Documents	Cited	% Cited	Total Cites	ACD	DCS
2001	3	2	66.67	27	13.5	7.5
2002	6	2	33.34	115	57.5	33.83
2003	12	6	50	20	3.34	2.09
2004	19	10	52.64	51	5.1	3.4
2005	29	17	58.63	160	9.42	6.73
2006	42	22	52.39	281	12.78	9.83
2007	50	29	58	452	15.59	12.99
2008	38	21	55.27	284	13.53	12.3
2009	72	46	63.89	483	10.5	10.5
2010	88	64	72.73	1419	22.18	24.64
2011	126	97	76.99	994	10.25	12.81
2012	157	110	70.07	2720	24.73	35.33
2013	226	159	70.36	2299	14.46	24.1
2014	242	150	61.99	1152	7.68	15.36
2015	320	208	65	1363	6.56	16.39
2016	387	224	57.89	1630	7.28	24.26
2017	416	251	60.34	1127	4.5	22.46
2018	487	148	30.4	482	3.26	32.57
Total	**2720**	**1566**	**57.58**	**15059**	**9.62**	

*** DCS = Distributed CiteScore, ACD = Average Citations per Document**

Figure 4.1 depicts the year wise growth of number of cited documents, year wise average citations per document (ACD) and DCS. From table 5 it can be observed that the citable documents percentage is

between 50% to 70% for most of the years. Only 57.58 % of the all documents published in cyber security domain are citable until may 2018.

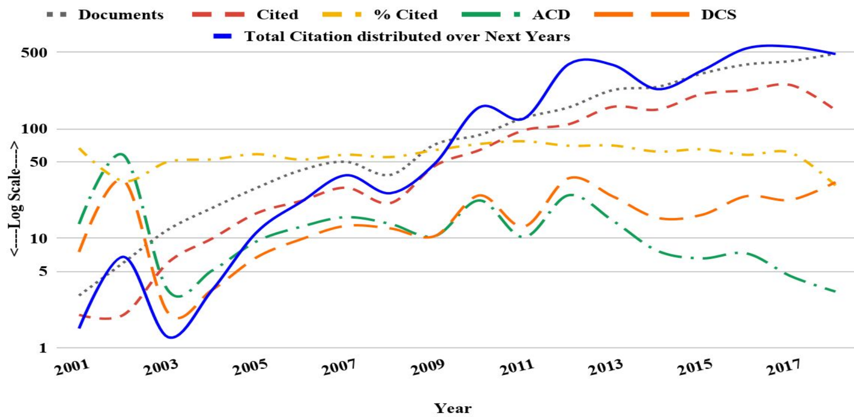

Fig. 4.1 Citation Analysis

6.5 Countrywise contributions, Sources, Affiliations, Funding Agencies

6.5.1 Country Wise distribution of articles

Table 6 shows that the majority of 1078 contributions out of 2720 are from United States of America. Which is around 39.63% of total publications and followed by UK, China and India with 8.75%, 5.22%, 4.74% respectively. These top four countries have a combined share of 58.35% of total publications. Other six countries out of top 10 contributors, have contributions ranging in between 2-3% of the total. So, it can be inferred that out of all the countries, the United States produced highest research compared to others.

Table 6 Country wise Publication share

Country	Documents	%
USA	1078	39.63
United Kingdom	238	8.75
China	142	5.22
India	129	4.74
Australia	80	2.94
South Korea	80	2.94
Italy	78	2.87
Canada	77	2.83
Germany	76	2.79
Japan	55	2.02

6.5.2 Sources of Publication

Table 7 represents top 20 sources publishing documents directly related to cyber security. Most contributory source is ‘Lecture notes in Computer Science including subseries lecture notes in Artificial

Intelligence And Lecture notes in Bioinformatics' with 108 documents followed by 'ACM International Conference Proceeding Series' with 70 documents. Although most of the sources in the top 20 are related to computer science and communications which is very obvious, but sources ranked 13,14,18 and 19 has also published a significant number of documents which are related to smart grid, infrastructure protection, gas pipeline infrastructure and electrical engineering; all these fields extensively use computing and networking. Even cyber security papers has been published in 'Economist United Kingdom' and 'automotive industries AI', which shows that this topic has very much impact on the automotive sector and economics also.

Table 7 Top 20 Sources in Cyber Security

	Source	No. of Documents
1	Lecture Notes In Computer Science Including Subseries Lecture Notes In Artificial Intelligence And Lecture Notes In Bioinformatics	108
2	ACM International Conference Proceeding Series	70
3	Communications In Computer And Information Science	32
4	Advances In Intelligent Systems And Computing	27
5	Computers And Security	24
6	CEUR Workshop Proceedings	21
7	Computer Fraud And Security	21
8	Control Engineering	21
9	Economist United Kingdom	20
10	IFIP Advances In Information And Communication Technology	20
11	Proceedings Of SPIE The International Society For Optical Engineering	19
12	International Journal Of Security And Its Applications	13
13	IEEE Transactions On Smart Grid	12
14	International Journal Of Critical Infrastructure Protection	12
15	Procedia Computer Science	11
16	Proceedings IEEE Military Communications Conference MILCOM	11
17	Automotive Industries AI	10
18	Lecture Notes In Electrical Engineering	10
19	Pipeline And Gas Journal	10
20	Proceedings Of The ACM Conference On Computer And Communications Security	10

6.5.3 Affiliations

Figure 5.3 shows Top 15 affiliations as reported by authors who have published documents in cyber security domain. cyber security Researchers affiliated to *University of Illinois Urbana-Champaign* have produced the highest number of publications followed by researchers affiliated with *Northwest National Laboratory Washington*, with 27 and 26 number of publications respectively. These are followed by *The Royal Institute of Technology KTH* with 22 publications and *University of Virginia* with 20 no of publications. Out of these top 15 affiliations, 13 are situated in the USA and other two *The Royal Institute of Technology KTH* and *University of Oxford* are in Sweden and the United Kingdom respectively.

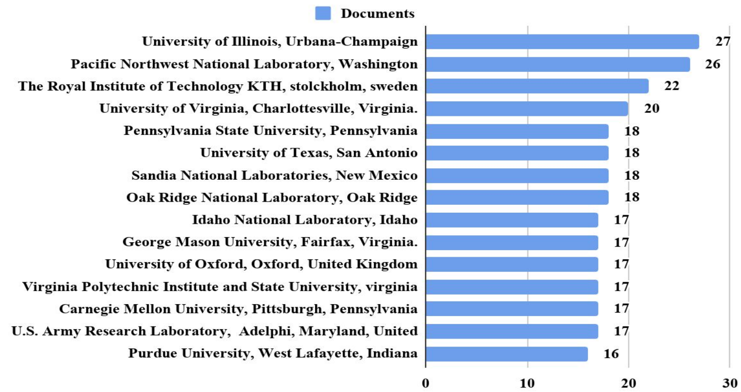

Fig. 5.3 Affiliations

6.5.4 Funding Agencies

Table 8 shows the top 15 Funding Agencies acknowledged by the published documents which were funded for cyber security research. Among these Agencies highest number (8 out of 15) of funding agencies are from the USA, 1 each from China, United Kingdom, European Union, Korea, Canada and Japan. USA is the country had a diverse no. of institution which funded the research directly related to cyber security. There are total 161 unique funding agencies worldwide which produced a total of 420 publications on cyber security.

Table 8 Top Funding Agencies in Cyber security Research

Funding Agency	Country	No. of Documents
National Science Foundation (NSF)	USA	58
National Natural Science Foundation of China (NSFC)	China	43
Engineering and Physical Sciences Research Council (EPSRC)	UK	18
Office of Naval Research (ONR), ***Defence***	USA	13
U.S. Department of Energy (DOE)	USA	12
European Commission (EC)	EU	11
U.S. Department of Homeland Security (DHS), ***Defence***	USA	10
The National Research Foundation of Korea (NRF)	Korea	8
Air Force Office of Scientific Research (AFOSR), ***Defence***	USA	8
Army Research Office (ARO), ***Defence***	USA	7
Army Research Laboratory (ARL), ***Defence***	USA	6
IEEE Foundation	World	6
Air Force Research Laboratory (AFRL), ***Defence***	USA	5
Natural Sciences and Engineering Research Council of Canada	Canada	5
Japan Society for the Promotion of Science	Japan	5

Figure 5.4(a) shows that funding by the National Science Foundation of the USA has produced the highest number of publications on cyber security, followed by the National Science Foundation of China.

Figure 5.4(b) shows how massively USA has funded cybersecurity research, a total of 119 publications are produced by 8 top funding agencies of USA listed as top 15 in the Table 8.

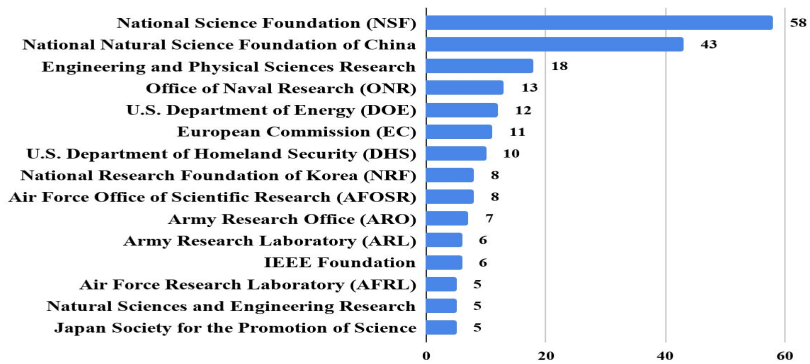

Fig. 5.4(a) Top funding agencies in Cyber security

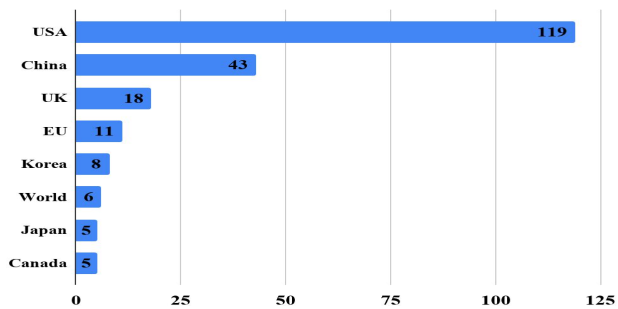

Fig. 5.4(b) Country wise aggregate publications by top 15 funding agencies

7. Results and Discussion

The major findings of the study are as follows:

- The highest 17.9% of papers were published in 2018 and the lowest 0.11% of research articles published in the year 2001.
- Year 2008 witnessed the lowest AGR of -24% and a drop in RGR to 0.212 which means high doubling time of 3.27, RGR has been decreasing exponentially and doubling time is highest with the value of 3.51 in 2018. The lowest AGR in 2008 is followed by the july 2009 cyber attacks ("Wikipedia", n.d.) and then worldwide interest in cybersecurity spiked ("Google Trends," n.d.).
- It is observed that the year 2008 dip in number of publications appears just before the large scale cyber attacks in the year 2008 and 2009 against United States of America.
- The sudden jump in the number of publications are because of high funding in cyber security research by America as a response to 2008 and 2009 cyber attacks.
- In 2018 CC is found to be 0.561. The trend shows that the cyber security domain will have more jointly authored publications than single authored.
- Out of total 2720 publications until year 2018, 1915(70.41%) were contributed by joint authors and the rest of 805(29.59%) contributions were by single authors, which is also on decline since last three years. Joint authorship has a strictly increasing trend since 2008.
- The average number of authors per paper has an increasing trend, which can be inferred as decreasing quantitative productivity of authors but increasing qualitative productivity of research as the DCS is increasing along with the increase in collaborations.

- The top contribution of 1078 publications are from the United States, followed by China on the second position with 238 number of publications, the United Kingdom is the third largest contributor with 142 publications and 129 contributions came from India.
- The maximum number of citations were 2720 in the year 2012 followed by 2299 in the year 2013; whereas, the highest percentage of citable papers was 76.99% in the year 2010 followed by 72.73% in the year 2010. Only the years from 2010 to 2014 has above 70% citable papers each year.
- United states of america has funded highest number of research publications and 49 out of total 119 funded research had received funding from Defence establishments.
- None of the Indian research on cyber security is found, which reported funding.

8. Conclusion

The study of cyber security research publication based on Scopus database and the following conclusions are made from the above analysis and discussions. It is also observed that the maximum number of articles were published in 2016. It is identified that the highest numbers of contributions were by joint authors and the lowest contributions were by single authors; In the starting years research was mostly done by single authors but later joint authorship has taken over in terms of number of publications. Finally, it was noticed that most of the researchers preferred publishing as conference paper which is a well known medium to get recognition and increase collaborations. It is followed by journal articles which are the premier medium of information dissemination. Furthermore, it seems most of the cyber security research produced by India are funded by UGC, CSIR and other funding agencies with very basic funding in the form of scholarships and small project grants, otherwise funding would have been acknowledged in publications. There is one more observation that Defence funded work has not been carried out in India as compared to other competitors in terms of number of publications in cybersecurity domain worldwide.

Wilshusen and Barkakati (2013) submitted report at United States Government Accountability Office on Cybersecurity which aimed at defining National Strategy, Roles, and Responsibilities for more effective implementations. In the report, the US government has recognised the active challenges related to cyberspace infrastructures and has been involved in creating awareness and developing workforce for technical research in the area. It has also been admitted that even though the federal government has been active for Cybersecurity concerns, it continued to face challenges in implementing Cybersecurity which needs to be and can be addressed only by an Effective Strategy.

In light of the above facts we suggest to work on larger fundings, public awareness, competent workforce development for active research and planned public-private collaborations in the domain of Cybersecurity, which is a necessity to survive in the cyberspace; and defence establishments should take responsibility by undertaking serious research projects to address the national cyber security concerns. Recent malware attack on system of the Kudankulam power plant is a live example to explain the need of research in this area ("NPCIL admits malware attack at Kudankulam Nuclear Power Plant," 2019). Best strategy for Cybersecurity is to be prepared in advance against any vulnerability, which requires continuous research and frequent assessment of cyber infrastructures.

Future Work: Cyber Security is a very important area of research which is attracting immense interest all over India and worldwide. It is very important to map Cybersecurity research conducted in India, it will give a view of the preparedness of india in the area of Cybersecurity proficiencies and blank spots. We will attempt to identify the research gaps and research hotspots from the point of view of cyber security of power grids, power generation plants, nationalised banking systems and Indian cyber space etc..

References

AIG Study: Systemic Cyber Attacks Likely in 2017; Financial Services, Power/Energy, International Cyber Conflicts Key Concerns. (2017, May 10). Retrieved September 4, 2019, from https://www.businesswire.com/news/home/20170510005781/en/AIG-Study-Systemic-Cyber-Attacks-2017-Financial

Ajiferuke, I., Burell, Q., & Tague, J. (1988). Collaborative coefficient: A single measure of the degree of collaboration in research. *Scientometrics*, *14*(5), 421–433. Retrieved from https://doi.org/10.1007/BF02017100

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. Retrieved from https://doi.org/10.1142/S0219649216500441

Choo, K.-K. R. (2011). The cyber threat landscape: Challenges and future research directions. *Computers & Security*, *30*(8), 719–731. Retrieved from https://doi.org/10.1016/j.cose.2011.08.004

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. Retrieved from http://www.indianjournals.com/ijor.aspx?target=ijor:ijidt&volume=2&issue=3&article=003

Elmaghraby, A. S., & Losavio, M. M. (2014). Cyber security challenges in Smart Cities: Safety, security and privacy. *Journal of Advanced Research*, *5*(4), 491–497. Retrieved from https://doi.org/10.1016/j.jare.2014.02.006

Google Ngram Viewer. (n.d.). Retrieved July 28, 2019, from https://books.google.com/ngrams/graph?content=cyber+security&case_insensitive=on&year_start=1970&year_end=2008&corpus=15&smoothing=7&share=&direct_url=t4%3B%2Ccyber%20security%3B%2Cc0%3B%2Cs0%3B%3Bcyber%20security%3B%2Cc0%3B%3BCyber%20Security%3B%2Cc0%3B%3BCyber%20security%3B%2Cc0%3B%3BCYBER%20SECURITY%3B%2Cc0

Google Trends. (n.d.). Retrieved August 30, 2019, from Google Trends website: https://trends.google.com/trends/explore?date=all&q=cyber%20security

July 2009 cyberattacks**.** (n.d.). In Wikipedia. Retrieved December 04, 2019, from https://en.wikipedia.org/wiki/July_2009_cyberattacks

Lee, W. (2008). How to identify emerging research fields using scientometrics: An example in the field of Information Security. *Scientometrics*, *76*(3), 503–525. Retrieved from https://doi.org/10.1007/s11192-007-1898-2

Martin, G., Martin, P., Hankin, C., Darzi, A., & Kinross, J. (2017). Cybersecurity and healthcare: How safe are we? *BMJ*, *358*. Retrieved from https://doi.org/10.1136/bmj.j3179

NPCIL admits malware attack at Kudankulam Nuclear Power Plant. (2019, October 30). *The Hindu*. Retrieved from https://www.thehindu.com/news/national/npcil-acknowledges-computer-breach-at-kudankulam-nuclear-power-plant/article29834644.ece

Olijnyk, N. (2014). Information Security: A Scientometric Study of the Profile, Structure, and Dynamics of an Emerging Scholarly Specialty. *Selected Dissertation Abstracts, 2012-*. Retrieved from https://digitalcommons.liu.edu/post_select_disabst/17

Olijnyk, N. V. (2015). A quantitative examination of the intellectual profile and evolution of information security from 1965 to 2015. *Scientometrics*, *105*(2), 883–904. https://doi.org/10.1007/s11192-015-1708-1

Poornima, A., Surulinathi, M., Amsaveni, N. & Vijayaragavan, M. (2011). Mapping the Indian research productivity of food science and technology: A scientometric analysis. *Food Biology*. Retrieved from https://updatepublishing.com/journal/index.php/fb/article/view/1678

Stevenson, C. L. (2014). Dan Craigen, Nadia Diakun-Thibault, and Randy Purse. *Technology Innovation Management Review*, 10. Retrieved from https://timreview.ca/article/835

Subramanyam, K. (1983). Bibliometric studies of research collaboration: A review. *Journal of Information Science*, *6*(1), 33 -38. Retrieved from https://doi.org/10.1177/016555158300600105

Velmurugan, C., Radhakrishnan, N., Radhakrishnan, N., & Radhakrishnan, N. (2016). Malaysian Journal of Library and Information Science: A Scientometric Profile. *Journal of Scientometric Research*, *5*(1), 62 -70. Retrieved from https://doi.org/10.5530/jscires.5.1.9

Wells, L. J., Camelio, J. A., Williams, C. B., & White, J. (2014). Cyber-physical security challenges in manufacturing systems. *Manufacturing Letters*, *2*(2), 74–77. Retrieved from https://doi.org/10.1016/j.mfglet.2014.01.005

Wilshusen, G. C., & Barkakati, N. (2013). Cybersecurity: National strategy, roles, and responsibilities need to be better defined and more effectively implemented. *GAO Reports*, 1-104. Retrieved from https://www.gao.gov/assets/660/652170.pdf