Summer 3-5-2021

ANALYSIS OF RECENT TRENDS IN MALWARE ATTACKS ON ANDROID PHONE: A SURVEY USING SCOPUS DATABASE

Sonali Kothari Tidke
Symbiosis Institute of Technology, sonali.tidke@sitpune.edu.in

Vijayshri Khedkar
Symbiosis institute of Technology

ANALYSIS OF RECENT TRENDS IN MALWARE ATTACKS ON ANDROID PHONE: A SURVEY USING SCOPUS DATABASE

Abstract

In past few years, smartphone use has shifted from professional access to personal need. Smartphone has now become an essential requirement to perform day to day activities. This has made smartphones unsecured and vulnerable to cyber threats and malware attacks. This study is also focused on finding intersection between malware attacks and Android OS considering Android as the most widely used mobile OS. A comprehensive search is conducted on Scopus Database for peer-reviewed articles. The study is carried out on bibliometric data of the considered articles to generate a highly useful concept map.

Keywords

Scopus database, bibliometric analysis, malware detection, malware analysis, Android

I. Introduction

Malware attacks are an increasingly critical aspect of smartphone industry. The rapid digitization of personal information with profession data, from personal bank records and internet banking to mobile banking [1] and all professional communication over e-mails to WFH concept considering current pandemic situation, introduces high risk related to cyber threat and cyberattacks. Malware attacks can target desktop machines to personal smartphones of end user. Malware attacks can perform variety of activities including steal personal information, gain access to contact list, gallery, send messages to premium-rate numbers etc. to mention amongst few [2,3]. Malware attack becomes easy for attacker when user falls prey for malicious applications available on genuine or 3rd party app store [4].

As Android is one of the most commonly used platform by smartphones, it is more interesting to analyze malware attacks, malware detection methodologies on Android OS. This will give an elaborated view about security issues of particular platform exploited by cyber attackers. Android is permission based platform and hence all applications needs some permissions to execute on Android platform. As permissions are categorized in groups and sometime giving permission to one application gives access to all permissions in particular group [5]. Also malware attacks are

disseminated as updates of existing widely used applications on Android platform [6], Angry Birds and Zombies are one such example of malware attacks on famous games. There are many such examples to highlight.

With increasing use of smartphones, lack of knowledge about cyber threats on smartphones and dependency on smartphones for day to day activities, it is more easy to target smartphones and exploit various vulnerabilities of the platform and application store [5]. In past 15 years, many researchers have started working on detection of malware attacks. But not sufficient importance is given for prevention of them. Cybercrimes are increasing day by day and hence more importance is needed to avoid malware attacks on smartphones. Mobile devices are dangerous for business too as employees frequently access corporate resources from their devices. If any devices are compromised by malware, sensitive information related to business can be captured. Security of information depends highly on how mobile environment is set up and methods used by employees to access sensitive data. Unless corporates have policies and controls regarding mobile devices or smartphone access, employees probably access sensitive information from their smartphones and information can be stolen by malware.

II. Methods

Bibliometric analysis is the analysis of publications. It is a process of evaluating, analyzing and visualizing research terms. It helps in identifying current trends in research domain of interest, find publication information and check impact of effectiveness of the researcher. Bibliometric is considered to be the oldest method of research in library and information science [7]. Use of bibliometric research can be divided into two parts: one is based on general instructions and other based on publication details [8]. First part explains searching of article using search engine to select relevant articles and reduce errors. While second part is used to analyze publications such as impact factor, citations, publisher, organization, country etc. Methodology suggested in [8] is used by many researchers to carry out their studies like [9,10,11]. In proposed study, manual search and software based analysis is used to retrieve, exclude and filter articles based on required subject area, keywords frequency, author and country.

Study of eligibility criteria: A comprehensive search is performed on Scopus database peer-reviewed published articles from Journals, conferences, books etc. Search keywords for the study are identified and adopted from [12,13]. For returned results, many filters are applied to remove

unrelated documents and analyze only documents of interest. For analysis of documents of interest, articles focused on malwares attacks and detection are studied together. As Android is one of the most popular operating system, it is considered for study in place of other mobile OS.

With query search of "android malware detection" returned 7128 documents by Scopus database [14]. This is further refined to reduce documents out of scope of study. So search with query "detection of malware on android smartphones" returned 3572 documents. To further reduce documents of interest, keywords and subject areas are selected. For example, keywords like malware and Android (operating system) were dominating results but were not returning target documents while many keywords are not of interest in this study like telephone sets, e-learning, codes, commerce, robots etc. are filtered out. Malware detection needs to be performed statically and dynamically and hence these keywords are focused. Subject areas considered are computer science, engineering, mathematics, decision sciences and social sciences as subject areas like arts and humanities, chemistry, chemical engineering etc. have articles less than 25 and are not related to the field of study. Considering this, subject areas with less relevance to topic, keywords which are not required to consider are excluded.

This has returned 399 documents which are studied and analyzed as final result of search process. Documents returned are published between 2009 to 2021 while if we observe number of publications, it can be said that Android OS attacks are increasing in last few years only and with increase in number of users, it will keep on increasing if smartphones using this platform are not secured properly. This also leads to the conclusion that with availability of cost efficient smartphones, it's uses are increased and with widespread use of smartphones, malware attacks are also increasing and cyber criminals are more focused now on smartphones with other cyber-attacks.

Figure 1 shows the process followed for limiting search results to the topic of study. It highlights methodology used for searching, shortlisting and finalizing documents for study. This analysis of documents is useful for identifying how cyber security researchers are moving towards malware identification and specifically on Android OS.

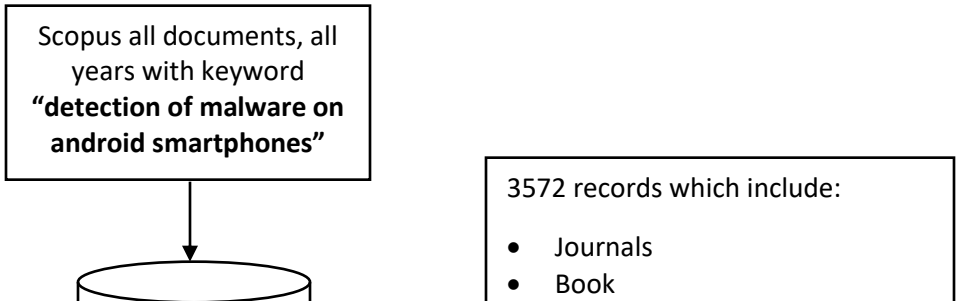

Fig. 1. Schematic of search method and results

III. Results and Discussion

To identify publication information on malware detection on android smartphones following query was used:

KEY: malware AND detection AND on AND android AND smartphones

Excluded subject areas considering scope of study: “PHYS”, “MATE”, “BUSI”, “MEDI”, “ENER”, “CENG”, “NEUR”, “ENVI”, “MULT”, “ARTS”, HEAL”, “AGRI”, “BIOC”, “CHEM”, ECON”, “PHAR”

Publication state : “final”

Keywords considered : "Smartphones" , "Android" , "Malware Detection" , "Mobile Security" , "Android Malware" , "Machine Learning" , "Android Applications" , "Static Analysis" , "Artificial Intelligence" , "Android Platforms" , "Mobile Devices" , "Security" , "Dynamic Analysis" , "Intrusion Detection" , "Malwares" , "Malicious Behavior" , "Smartphone Securities" , "Mobile Operating Systems" , "Android Smartphone" , "Android Apps" , "Malicious Activities" , "Malware Analysis" , "Malware Families" , "Smart-phone Applications" , "System Calls" , "Data Privacy" , "Static And Dynamic Analysis" , "Botnet" , "Botnet Detections" , "Detection Methods" , "Malicious Android Applications" , "Malware Classifications" , "Detection Accuracy" , "Malware Attacks" , "Permission" , "Smartphone Malware" , "Android Markets" , "Anti-malware" , "Mobile Phones" , "Ransomware" , "Smart Phones" , "Android Smart Phones"

A. Type of document

The result of initial query is filtered and limited to most relevant documents which in turn return 390 documents of English language while 9 documents of other languages: 283 conference papers comprises of 70.8% of total documents, followed by 102 documents of type Article which contributes to the 25.3 % of all documents in study whereas 2.5% that is 10 documents under study are book chapters and only 4 documents equivalent to only 1% is of type review.

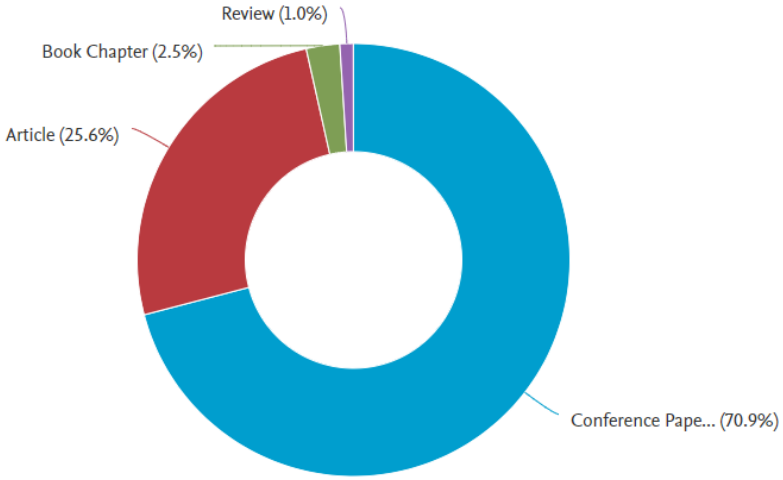

Fig. 2. Cumulative analysis of Scopus articles based on type

B. Publication Output

Documents returned by query ranges from 2009 to 2021. Year wise cumulative number of publications is shown in figure 3. Pearson correlation coefficient between year and the yearly published documents cumulative number is 0.595. In analyzed documents 390 are in English, 6 in Chinese, 1 in Turkish while 2 are in Spanish which suggests that 97.7% documents are in English while other languages contribute only to 2.3% of overall documents on the topic of study.

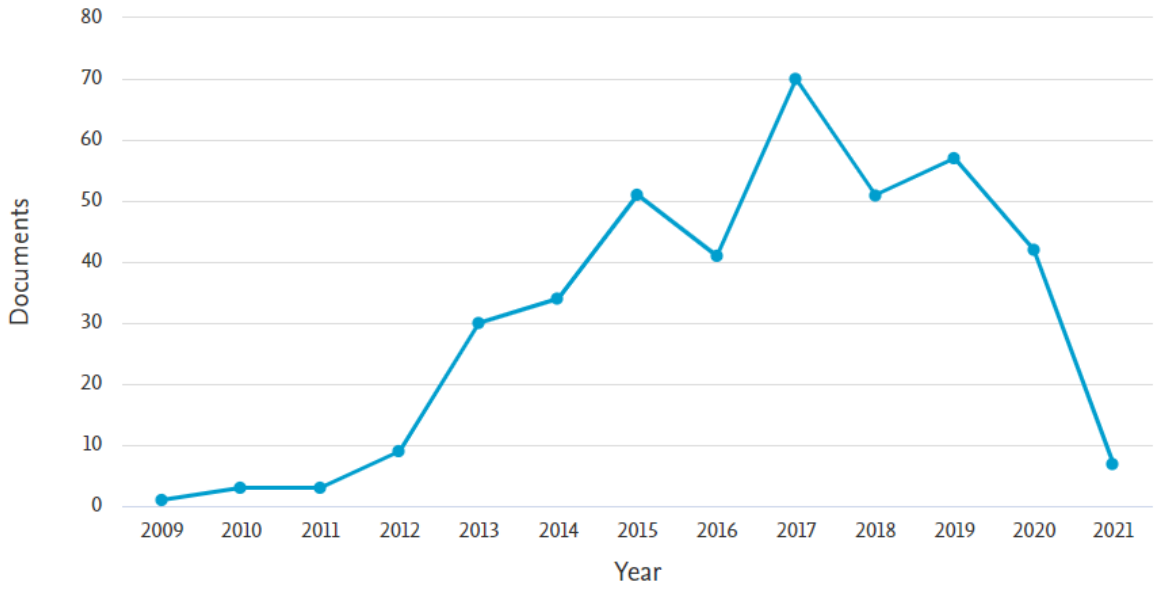

Fig. 3. Cumulative number of publications on malware detection on Android every year

To justify this paper, analysis shows country wise publication in figure 4 while cumulative publication details based on subject areas is depicted in figure 5.

There are many countries contributing in malware detection domain and it is increasing regularly. In figure top 10 such countries contributing in the area of malware detection is shown. Graph suggests that China is the leading player in malware detection on Android phones while India is the second country followed by US, Italy and Canada. Documents considered in study are from 54 countries but most of the countries are having very less contribution in research on malware detection. This needs attention of researchers that many countries are still less introduced and aware about risks of malware on Android phones and not contributing much in this field.

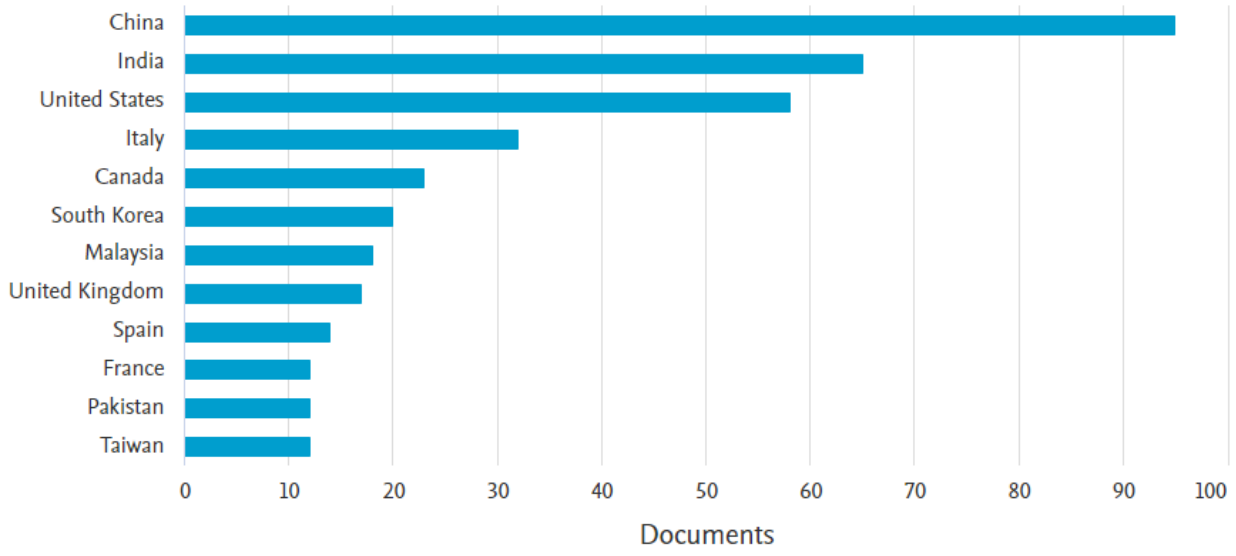

Fig. 4. Malware analysis documents published by top 12 countries

C. Distribution in subject areas

Figure 5 shows that Computer Science is the most contributing area for malware detection type with 58.6 % documents are published under this subject area while Decision Science and Social Sciences are contributing less than 10% that 5.2 % and 3.8 % respectively.

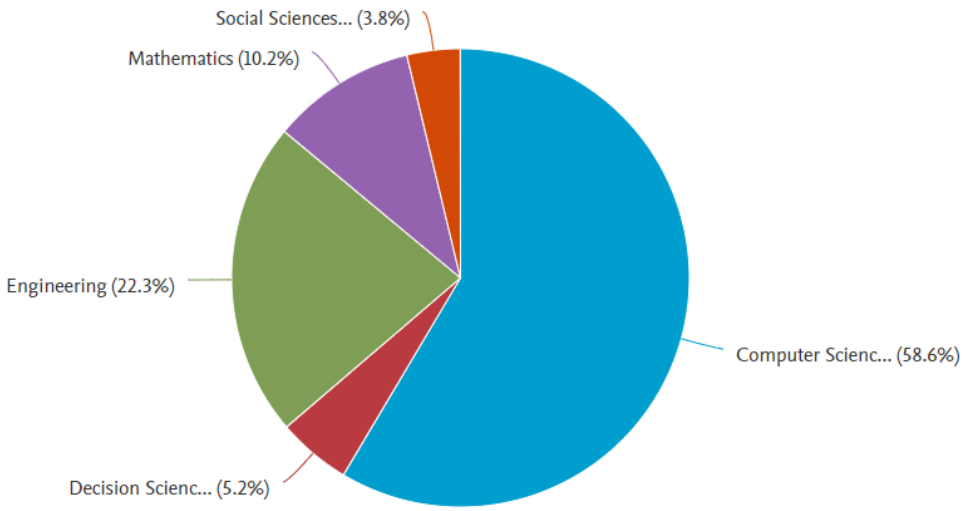

Fig. 5. Analysis of publications in various subject areas

Many publishers and sources are constantly thriving to make Android phone users aware about risks involved in unsecured application uses and downloads. Many sources are publishing documents based on malware detection regularly. More than 70 sources are available in the dataset

of analysis. Out of those 70 journals, figure 6 shows such top 5 sources who are contributing maximum in field of research. Table I shows top 9 journals who are publishing work at the intersection of malware detection and Android smartphone with the citescore of journal which is extracted from Scopus database. Overall, the 399 documents are published in 70 unique journals. Journals are ranked according to the number of articles published. For top journals, threshold of minimum 5 publications is considered.

Table I. Top journals with most documents published

Journal	Number of published documents	Journal CiteScore [15]
"Advances In Intelligent Systems And Computing"	21	0.9
"ACM International Conference Proceeding Series"	17	0.8
"Lecture Notes In Computer Science Including Subseries Lecture Notes In Artificial Intelligence And Lecture Notes In Bioinformatics"	17	1.9
"Communications In Computer And Information Science"	16	0.7
"Computers And Security"	11	7.5
"Security And Communication Networks"	11	4.2
"IEEE Transactions On Information Forensics And Security"	5	14.7
"Lecture Notes In Electrical Engineering"	5	0.5
"Proceedings Of The ACM Conference On Computer And Communications Security"	5	9.1

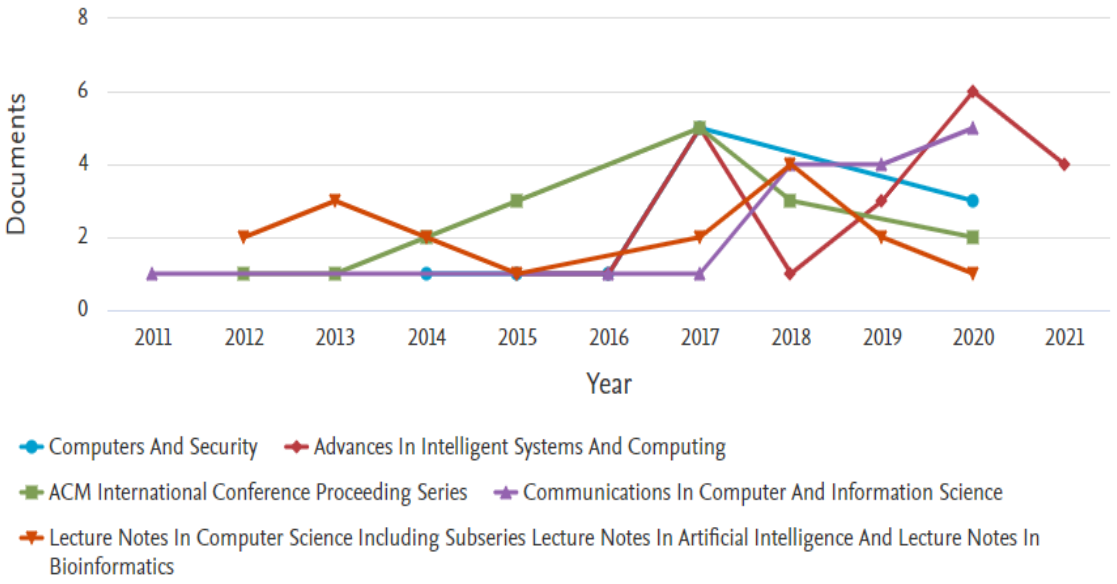

Fig. 6. Journals publishing research articles on malware detection

Analysis of dataset shows that there are 99 funding agencies which are sponsoring research in the domain of detection of malware on Android phones. Major funding is given by Chinese funding agencies to promote researchers and support their research work. National Natural Science Foundation of China is leading them with 31 documents published under their research funding. Table II. gives details of funding agency and documents published under their research grants. Figure 7 gives cumulative analysis of funding agencies and research publications promoted/funded by them.

Table II. Top 10 Funding sponsors and documents counts under their sponsorship

Funding Sponsor	Documents count
"National Natural Science Foundation of China"	31
"National Basic Research Program of China (973 Program)"	8
"National Science Foundation"	8
"Engineering and Physical Sciences Research Council"	5
"European Regional Development Fund"	4
"Ministry of Higher Education"	4
"European Commission"	3

“Fundamental Research Funds for the Central Universities”	3
“Institute for Information and Communications Technology Promotion”	3
“Ministry of Science, ICT and Future Planning”	3

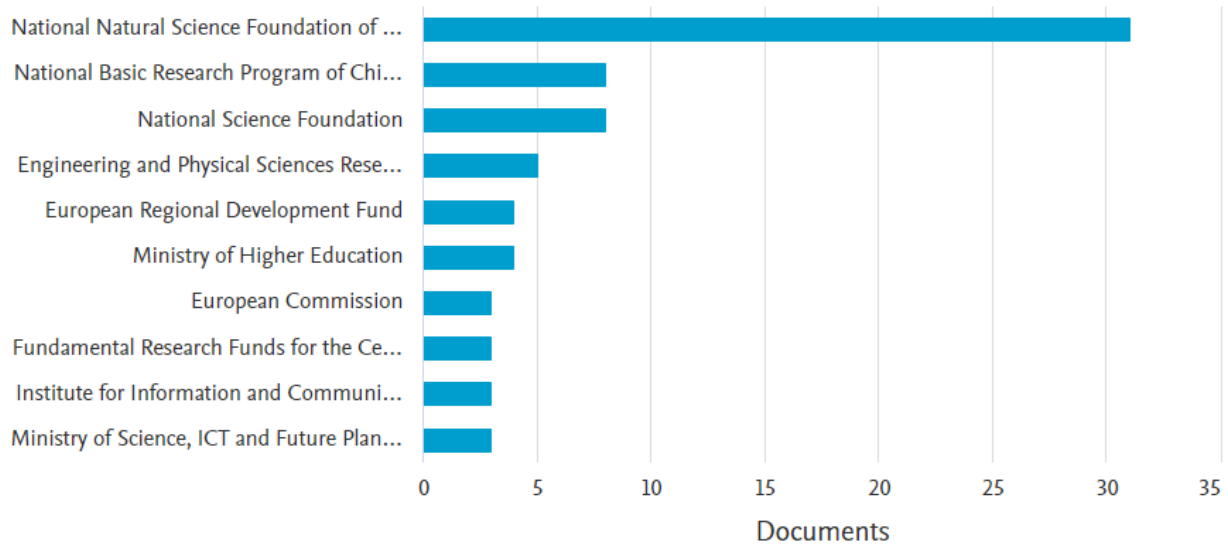

Fig. 7. Analysis of funding agencies supporting publication and research in malware detection on Android

Table III shows top 7 most-productive authors with at least 6 publications in malware detection on android smartphones areas, their affiliations [15], h-index [16] and overall citation of author [16] while figure 8 highlights documents published by them in the field.

Table III. Top productive Authors

Author	Affiliation	Documents published	Overall Citations	h-index
Mercaldo, Francesco	University of Molise, Campobasso, Italy	13	2198	28
Conti, Mauro	University of Padua, Italy	11	13320	58
Martinelli, Fabio	Rome, Italy	9	4699	34

Laxmi, Vijay	MNIT, Jaipur, India	8	3252	29
Gaur, Manoj Singh	IIT, Jammu, India	7	1740	25
Ghorbani, Ali Akbar	Canadian Institute for Cybersecurity, Canada	7	-	-
Mohd Saudi, Madihah	Universiti Sains Islam Malaysia, Nilai, Malaysia	7	381	11
Sakir Sezer, Sakir	Belfast, United Kingdom	6	5963	32

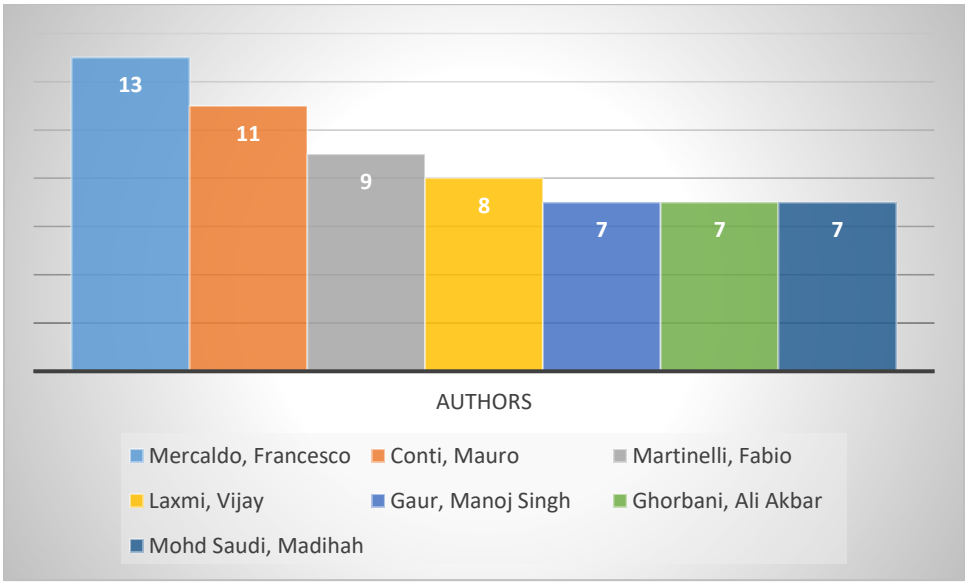

Fig. 8. Author wise publications from the dataset of study

D. Citation analysis of documents

Although number of citations does not indicate quality of any article, it is a measure of prestige and impact of author's work. Most frequently cited document in the dataset is "Dissecting Android malware: Characterization and evolution" by "Zhou Y., Jiang X." published in 2012 in IEEE Symposium on Security and Privacy with IF of 9.5 [17] with SJR 1.89 and is cited

1342 times from its publication in 2012. Article citations from its publication year is illustrated in Figure 9.

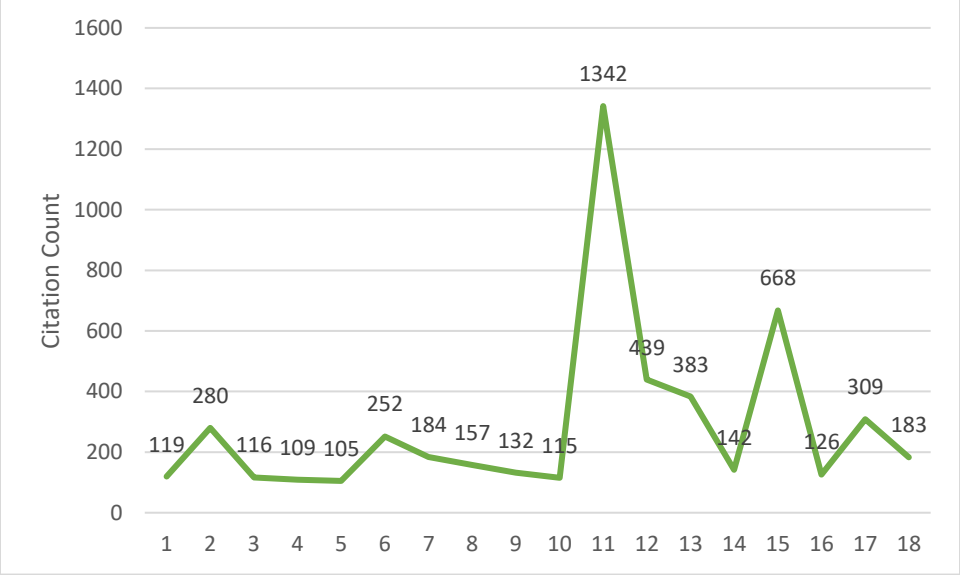

Fig. 9. Most cited document in its life after publication

Table IV shows details about publication year, author and title of the document with it citation score for documents having citations more than 100.

Table IV. Article wise citations

Author	Title	Citations	Year of publication
Zhou Y. et al.	"Dissecting Android malware: Characterization and evolution"	1342	2012
Burguera I. et al.	"Crowdroid: Behavior-based malware detection system for android"	668	2011
Grace M. et al.	"RiskRanker: Scalable and accurate zero-day android malware detection"	439	2012
Wu D.-J. et al.	"DroidMat: Android malware detection through manifest and API calls tracing"	383	2012

Bläsing T. et al.	“An android application sandbox system for suspicious software detection”	309	2010
Faruki P. et al.	“Android security: A survey of issues, malware penetration, and defenses”	280	2015
Rastogi V. et al.	“AppsPlayground: Automatic security analysis of smartphone applications”	252	2013
Peiravian N. et al.	“Machine learning for Android malware detection using permission and API calls”	184	2013
Schmidt A.-D. e tal.	“Static analysis of executables for collaborative malware detection on android”	183	2009
Suleiman Y. et al.	“New Android malware detection approach using Bayesian classification”	157	2013
Dini G. et al.	“MADAM: A multi-level anomaly detector for android malware”	142	2012
Zheng M. et al.	“Droid analytics: A signature based analytic system to collect, extract, analyze and associate android malware”	132	2013
Isohara T. et al.	“Kernel-based behavior analysis for android malware detection”	126	2011
Tam K. et al.	“The evolution of android malware and android analysis techniques”	119	2017
Yuan Z. et al.	“Droid-Sec: Deep learning in android malware detection”	116	2015
Amos B. et al.	“Applying machine learning classifiers to dynamic android malware detection at scale”	115	2013
Lindorfer M. et al.	“MARVIN: Efficient and Comprehensive Mobile App Classification through Static and Dynamic Analysis”	109	2015
Shabtai A. et al.	“Mobile malware detection through analysis of deviations in application network behavior”	105	2014

IV. Conclusion

In malware detection on Android smartphones research is determined to illustrate research work in the field of interest. The paper illustrated dominant countries, languages and authors. Paper also depicted most cited article, most prominent journals and numerous sponsoring agencies to carry forward research in the field of malware detection. Cumulative analysis of documents based on publication year, countries, subject areas is also illustrated. Bibliometric maps using vosviewer software are created to illustrate some more details about citations network. Figure 10 illustrates bibliometric coupling amongst documents while figure 11 maps countries and citations to documents published. In this mapping, out of 54 countries only 22 countries having published documents more than 5 while top 15 countries are depicted in network map.

As the study suggest, many countries are far behind in observing vulnerabilities of Android devices and more research is required in this domain. To rule out the possibility that few countries are not using Android smartphones, figure 12 shows region wise sales of Android smartphone to end users between 2018 to 2020 in million units [18]

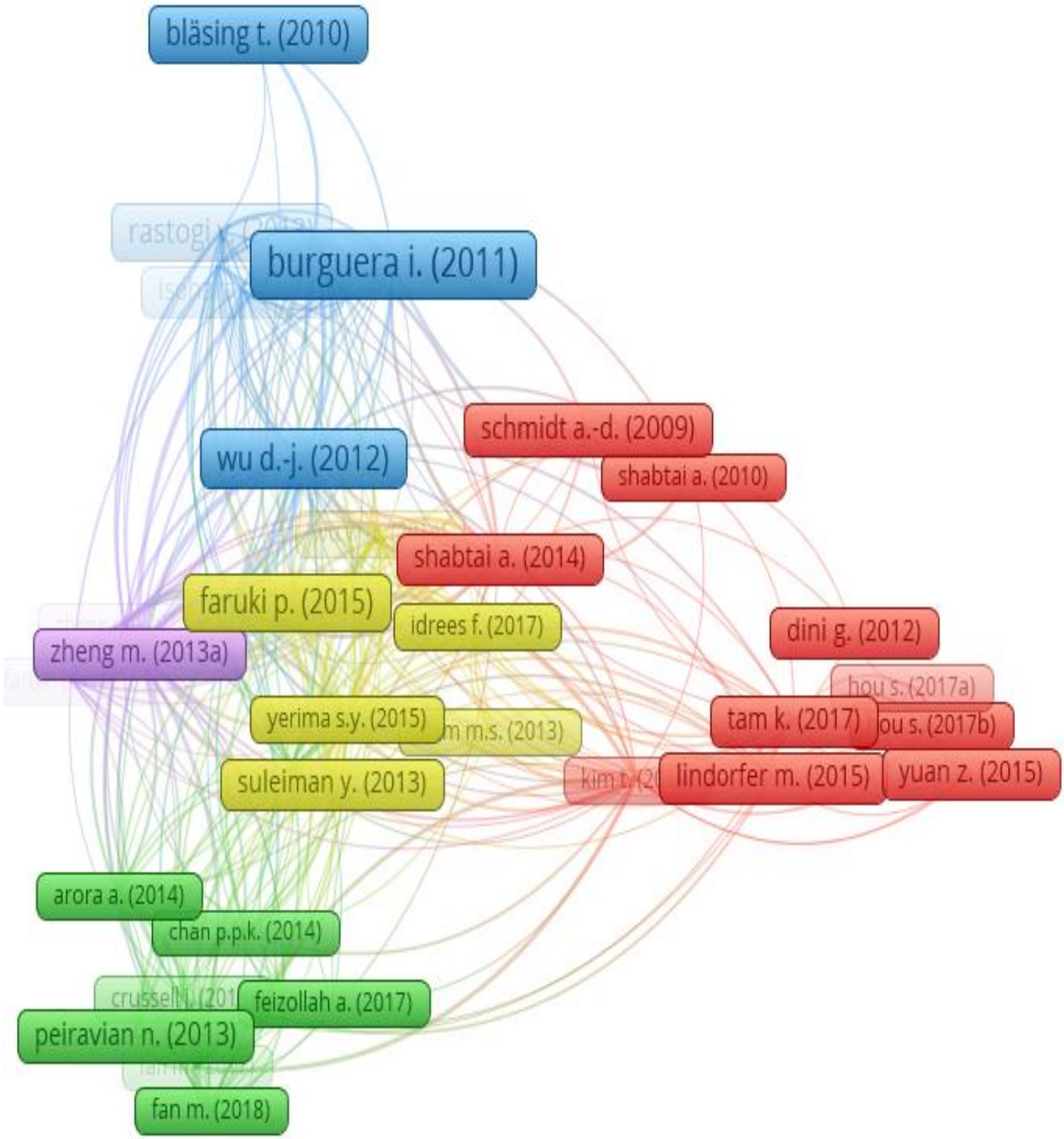

Fig. 10. Bibliometric coupling in documents

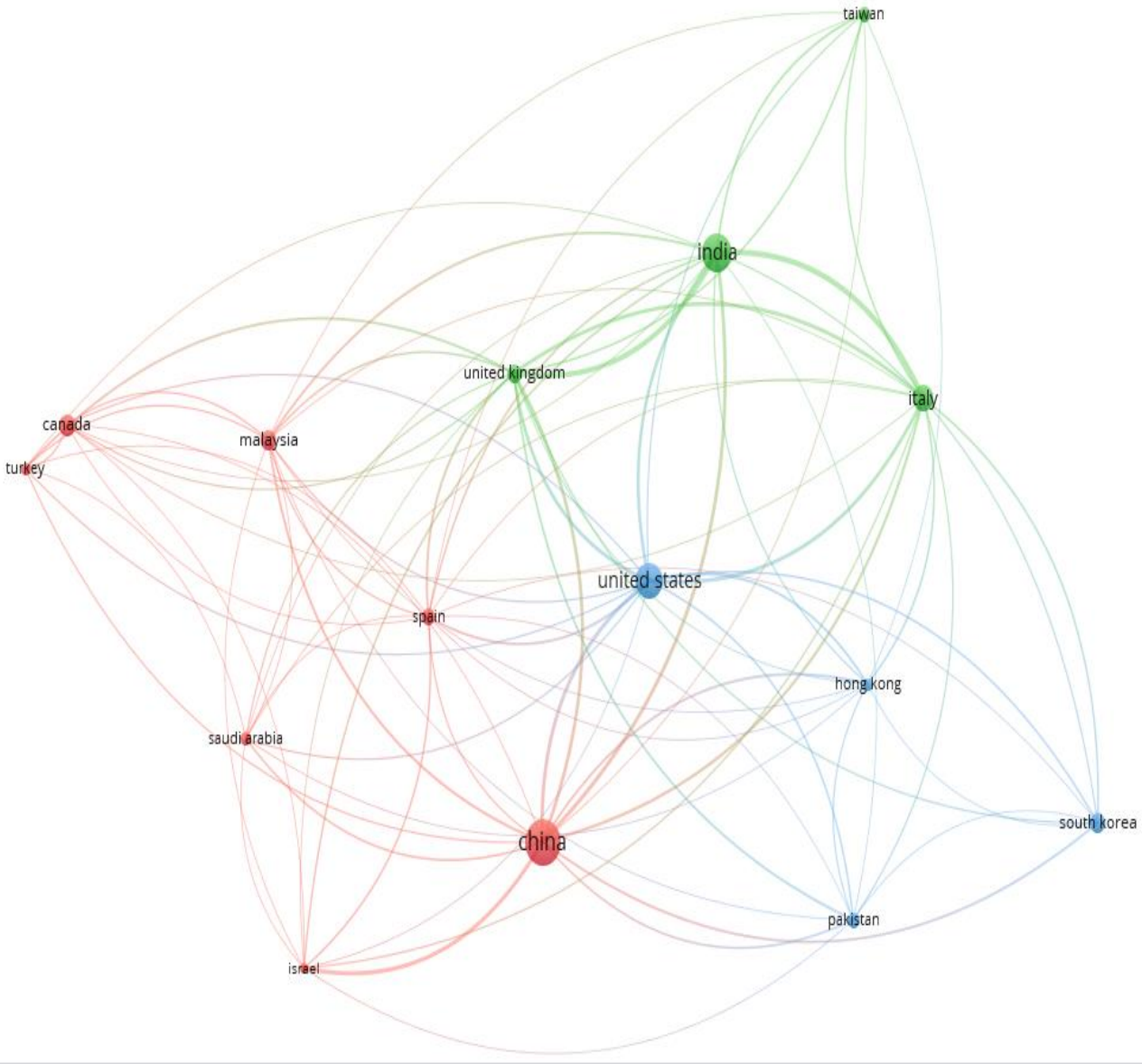

Fig. 11. Country wise citations map on bibliometric data

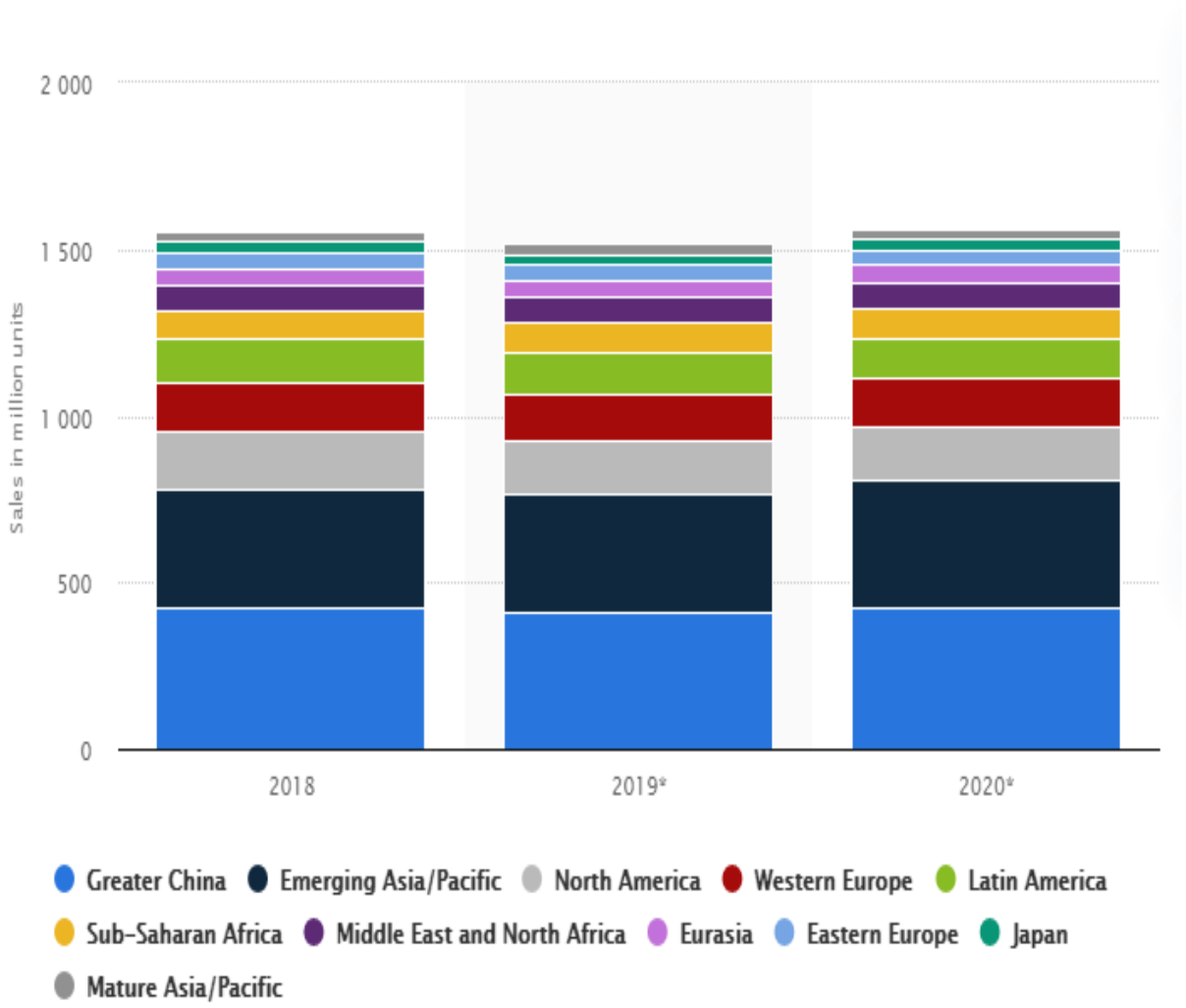

Fig. 12. Region wise smartphone sales between 2018-20 [18]

References

1) Mohammad S Jalali; Sabina Razak; William Gordon; Eric Perakslis; Stuart Madnick; "Health Care and Cybersecurity: Bibliometric Analysis of the Literature", Journal of Medical Internet Research, 15 Feb 20219. https://www.researchgate.net/publication/326879475

2) Mathur A., Podila L.M., Kulkarni K., Niyaz Q., Javaid A.Y., "NATICUSdroid: A malware detection framework for Android using native and custom permissions", Journal of Information Security and Applications, 2021.

3) O.S J.N., S M.S.B., “Detection of malicious Android applications using Ontology-based intelligent model in mobile cloud environment”, Journal of Information Security and Applications, 2021
4) Kothari S.,“Real Time Analysis of Android Applications by Calculating Risk Factor to Identify Botnet Attack”, Lecture Notes in Electrical Engineering,2020
5) Yang Y., Du X., Yang Z., Liu X. , “Android malware detection based on structural features of the function call graph”, Electronics (Switzerland), 2021
6) Kothari S., Joshi S., “Analysis of Android Applications to Detect Botnet Attacks”, Proceedings of the 2020 International Conference on Smart Innovations in Design, Environment, Management, Planning and Computing, ICSIDEMPC 2020, 2020
7) Wilson, V.,2016, .Evidence based library and information practice. Evid.BasedLibr. Inf. Pract.11,50–52.
8) Koskinen,J.,Isohanni,M.,Paajala,H.,Jääskeläinen,E.,Nieminen,P.,Koponen,H., Tienari, P.,Miettunen,J.“How to use bibliometric methods in evaluation of scientific research? An example from Finnish schizophrenia research”, Nord.J. Psychiatry,136–143. http://dx.doi.org/10.1080/08039480801961667.
9) Wu,X.,Chen,X.,Zhan,F.B.,Hong,S., “Global research trends in land slides during 1991 – 2014:a bibliometric analysis”. http: //dx.doi.org/10.1007/s10346-015-0624-z.
10) Dehdarirad, T.,Villarroya,A.,Barrios,M., “Research on women in science and higher education:a bibliometric analysis”. Scientometrics103,795–812. http: //dx.doi.org/10.1007/s11192-015-1574-x.
11) Loomes, D.E.,vanZanten,S.V., “Bibliometrics of the top 100 clinical articles in digestive disease”.. http://dx.doi.org/10.1053/j. gastro.2013.02.013.
12) National Initiative for Cybersecurity Careers and Studies. 2017 Nov 27. Explore Terms: A Glossary of Common Cybersecurity Terminology URL: https://niccs.us-cert.gov/glossary [accessed 2021-02-01]
13) BSI. Glossary of cyber security terms URL: https://www.bsigroup.com/en-GB/Cyber-Security/Cyber-security-for-SMEs/Glossary-of-cyber-security-terms/ [accessed 2021-02-01]
14) http://www.elsevier.com/online-tools/scopus [accessed 2021-02-01]
15) Scopus CiteScoreRank URL: https://www.scopus.com/ [accessed 2021-02-20]

16) Google Scholar citations of authors – searched individually https://scholar.google.com/citations?view_op=search_authors [accessed 2021-02-20]
17) https://www.resurchify.com/all_ranking_details_2.php?id=1525 [accessed 2021-02-20]
18) Statistica, https://www.statista.com/statistics/755388/global-smartphone-unit-sales-by-region/ [accessed 2021-02-20]