EURASIAN JOURNAL OF ACADEMIC RESEARCH
Innovative Academy Research Support Center
UIF = 8.1 | SJIF = 5.685 www.in-academy.uz

A SYSTEMATIC LITERATURE REVIEW ON MALWARE ANALYSIS

***Shoraimov Khusanboy Uktamboyevich,**
****Akhmadjonov Islomjon Kozimjon o'gli.**
* Teacher of the Department, "Systematic and Practical Programming", Tashkent University of Information Technologies named after Muhammad Al-Khwarizmi, UZBEKISTAN.
** Teacher of the Department, "Systematic and Practical Programming", Tashkent University of Information Technologies named after Muhammad Al-Khwarizmi, UZBEKISTAN.

https://doi.org/10.5281/zenodo.7471397

ARTICLE INFO

Received: 13th December 2022
Accepted: 21th December 2022
Online: 22th December 2022

KEY WORDS

Malware; Malware analytics; Malware code; Taxonomy; Signature-based; Anomaly-based; Malware system requirements.

ABSTRACT

Malware is a significant security danger on the Internet nowadays. Hostile to Virus organizations get a huge number of malwares tests each day. It is intended to harm PC frameworks without the information on the proprietor utilizing the framework and method headways are presenting enormous difficulties for scientists in both the scholarly world and the business. Malware tests are arranged and gathered for additional investigation. In this literature review, we did the manual research on the publications from the year 2014 to 2020. We selected about 27 articles out of 55 articles as primary studies and applied quality evaluation criteria and deducted research questions from them. The motivation behind this SLR is to inspect the accessible literary works on malware examination and to decide how exploration has developed and progressed regarding the amount, substance, and publication outlets. We also discussed the issues and challenges we are facing in malware analysis along with detection system requirements. Large numbers of the malicious programs are tremendous and confounded so it is difficult for researchers to fathom its subtleties. Scattering of malicious data beyond clients of the web and furthermore preparing them to effectively utilize against malicious items are critical to shielding clients from malicious attack. This review paper will give a comprehensive book index of techniques to help with battling malicious data.

INTRODUCTION. Malware is an overall term that incorporates infections, Trojans, Spywares, and other obtrusive code is far and wide today. Malware investigation is a multistep cycle giving knowledge into malware design and usefulness, encouraging the development of a cure.

EURASIAN JOURNAL OF ACADEMIC RESEARCH
Innovative Academy Research Support Center
UIF = 8.1 | SJIF = 5.685 www.in-academy.uz

The expression "malware" here is being utilized as the conventional name for the class of code that is pernicious, including infections, Trojans, worms, and spyware. Malware writers use generators, fuse libraries, and get code foremothers—there exists a hearty organization for trade, and some malware writers set aside an effort to peruse and comprehend earlier methodologies.

first-historically speaking PC infection (malware) Brain showed up in 1986 [1].

Malware is utilized to send spam messages, to perform web cheats, to take individual data like MasterCard data, and for some, other accursed assignments like Ransomware [2] and counterfeit antivirus programming [3].

Since 1988 [4, 5] the increment in the quantity of PC-based security penetrates affirms that noxious programming has arrived at practically unmanageable levels. Mulling over the degree of potential harm brought about by noxious programming, its discovery alone has caused huge issues for both the agents and the overall public. Recognition frameworks made by examiners are consistently put to broad use in identification works out. This paper is committed to investigating malignant programming location systems.

A. Malware analysis with respect to behaviour and signature-based:

Malware programming identification is divided into two parts: Signature-and Behavior-based advances and every innovation can be utilized with static, dynamic, or hybrid examination [5][16]. The specific technique for an oddity or mark put-together method is based on respect to how the innovation orders the data to distinguish noxious programming [6-9]. How noxious programming discovery is overseen is appeared in Fig. 1.

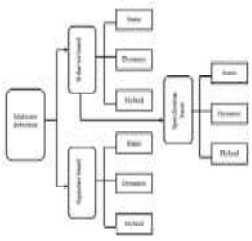

Fig.1 Flow-chart of malware detection

I. Behavior (Anomaly) based:

An anomaly-based put together discovery attracts with respect to its database to decide the presence of ordinary conduct to choose the firmness of malware in examination. Another type of inconsistency-based discovery is called specification-based recognition. This type of discovery examination occurs in 2 circumstances:

In preparing and learning circumstance. While in the preparation circumstance, an indicator attempts to get familiar with a typical conduct. It is very conceivable that an indicator is learning the host's conduct or the PUI's or possibly the two joined. The principal advantage of anomaly-based identification is the capacity to identify 'zero-day" interruptions.

- having location and observing circumstance

The 2 fundamental downsides of such system are:

Immense bogus caution charges: It is apt of unreasonable bogus alert charges, which are characterized as 'typical' yields sorted as (bogus positive) and separated by the all-out figures of 'ordinary' conduct.

- Also getting trouble in guaranteeing what boundaries must realize in preparation circumstance.

Detection by signature-based

Signature based discovery uses technology-based character to observe malicious code and thus affirm a malignant idea of a program in examination. Putting aside, the signaturebased discovery attempts by setting a criterion for utilizing a malicious code and accordingly uses it as the sort of perspective for distinguishing another noxious programming. In gathering every one of these models, the signature-based recognition creates an information base for itself. In an ideal framework, it is basic that the mark ought to perceive any program showing conduct fitting the mark's malignant data set. This data set contains all data required by the mark to recognize vindictive programming. This data set is counseled at whatever point here lies a possible issue with PUI.

The one of most fundamental issues of the signature-based recognition technique is the failure to perceive 'zero day' interruptions. A zero-day interruption means where there is none comparative mark lies in any information base to contrast and additionally, the accomplished individual is likely expected to plan a mark. Besides offering an approach to administrator mistakes it is a dreary cycle if the plan and establishment are not set up to work naturally. The way that certain malware can multiply the capacity to plan and introduce a more exact mark is amazingly basic. Engineers of such marks, which work on a programmed mode, could be found absent a lot of exertion, however altogether more energy should be placed into doing this. Be that as it may, all recognition systems could utilize one of three different philosophies.

B. Analysis Methodologies

The three principal malware investigation procedures are static, dynamic, and hybrid examination. Every examination strategy has its own favorable circumstances and weaknesses which have been talked about in this segment.

Table 1 shows the outline of investigation devices.

Static Analysis

It is an examination of programming executed without truly performing a program [10]. Different procedures are carried out to play out this examination. While a few rely upon the characteristics of the twofold record, by removing “byte code courses” of action from a combined one, and isolating Op code progressions in the wake of destroying the twofold report, to eliminate the “control stream graphical figures” among the party record, and eliminating API calls, from the equal, and the like. Each addresses the rundown of capacities and anyone or many are utilized for malicious area.

Dynamic Analysis:

It is the analysis in which while executing the program, a programming is done [17]. An information segment can achieve by the one of the examinations that is API calls, structure calls, direction follows, dirty assessment, vault changes, memory makes, etc. A part of the malware acknowledgment procedure utilizing dynamic one that has

been investigated formerly is according to the accompanying.

METHODOLOGY

In this section, we are going to apply the methodology for this systematic review that was proposed by 'Kitchenham" [35]. This part presents the technique to achieve the objectives of literature review. The phases of our procedure include:

Data planning

(1) Making of research questions
(2) Searching measures
(3) Addition and rejection measures

A. Data planning

In this phase, to achieve the goals of the current examination we recognized the important ways. In the given underlying process, it was guaranteed that the key and specialized plans were appropriately defined. This is guaranteed that different.

Analysis tools	Purposes	Tools
Static	Use whatever number antivirus recognition motors as could reasonably be expected to help characterization.	"Virus Total" (2008)
	Search the body of the malware for strings.	'Strings" (Microsoft, 2008)
Dynamic	Document respectability check to record gauge setup.	"Winalysis" (2008)
	Record observing. Discover which devices are opening, perusing and composing documents.	"Filemon" (2008)
	Vault observing. Screen vault exercises as they happen.	"Regmon" (Microsoft, 2008)
Hybrid	Dismantling, investigating	"IDA Pro" "OllyDbg" (Yuschuk, 2008)

EURASIAN JOURNAL OF ACADEMIC RESEARCH
Innovative Academy Research Support Center
UIF = 8.1 | SJIF = 5.685 www.in-academy.uz

RQ3. What is the size of datasets?
RQ4. What types of analysis methods are used in this research?
RQ5. Which are the requirements of malware programmin detectors?
RQ6. What are the advantages and disadvantages of malware identification draws near?
RQ7. What are the issues and challenges faced in malware analysis?

C. Searching measures

In the following part, we illustrate the measures to distinguish the given articles for this examination. To remove pertinent SLR concentrates on malware investigation, various electronic information bases were thought of and gotten to. The rundown of information bas looked, and their comparing URL is introduced in Table 2.

TABLE 2 ELECTRONIC DATABASES

Electronic database	Url
Scopus	www.scopus.com
SpingerLink	www.link.springer.com
IEEE Explore	www.ieeexplore.ieee.org
Web of Science	www.webofknowledge.com
ACM Digital Library	www.dl.acm.org
Google scholar	www.scholar.google.com
ScienceDirect	www.sciencedirect.com
Wiley online library	www.onlinelibrary.wiley.com
IET software Digital Library	www.digital-library.theiet.org

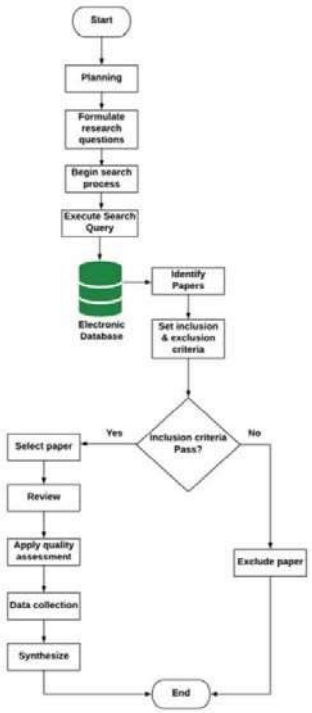

Fig.2. SLR process

TABLE 1 SUMMARY periods of the proposed philosophy were appropriately completed in a coordinated and standard way. This arranging stage shaped the reason for a fruitful usage of the proposed SLR strategy.

A. Forming *research questions:*

In the given section, we present the exploration addresses researched in the flow SLR study. The exploration addresses RQs examined in our investigation are:

RQ1. What number of yearly number of studies on malware examination have there been since 2010? RQ2. What types of datasets are used?

The pursuit string was consequently adjusted to suit the necessities of every information base. We looked through every information base by titles, edited compositions, and

We take out each article from the given electronic databases catchphrases. The figure 2 depicts the rules of deliberate utilizing the customary searching measures from cycle. conferences and journals individually. CONCLUSION AND FURTURE WORK

Malicious arracks lead to the typical danger for PC and correspondence frameworks to harm gadgets or take classified data. The reason for this SLR is to concoct a proficient procedure for malware location that consolidates the upsides of anomaly and signature based detection. This paper identified a definite order of malignant programming detection and evasion programs for specialists to 'nibble into'. Devoted significance was doled out to malignant programming necessities and

acknowledgment given to the basics of malware recognition and avoidance strategies.

Table 5 [Primary studies]

Ref. no.	Author, year	Area of research	Institutions	Publications	Contributions
[PS5]	B.Yu et al. 2018	A survey of malware behavior descriptio n and analysis	National university of defense technolo gy	journal	This paper conducted a survey on malware behaviour descriptio n and analysis considerin g description, analysis and visualizati on methods.
[PS11]	Parmjit et al. 2014	Literature Analysis on Malware Detection	Chandig arh Universi ty	Journal	This paper gives the android architecture, various types of malware and literature analysis for security considerat ions in android smartphon es.
[PS12]	H.M Deyla mi et al. 2016	Taxonom y of malware detection technique s	Universi ti Kebangs aan	Confere nce paper	This gives a comprehe nsive list of sources of strategies to help with battling malware.
[PS13]	Maigid a et al. 2019	SLR and metadata analysis of ransomwa re attacks and detection mechanis ms	Federal Universi ty of Technol ogy	journal	This paper fills in as a benchmark for new analysts in proposing the ransomware discovery technique.
[PS14]	Ya Pan et al. 2020	An SLR of android malware detection using static analysis	Nanjing universit y	journal	A systematic literature review is performed to perform the clarificati on on Android malware detection by using malware detection method

References:

1. Arief, B. & Bernard, D, " Technical and human issues in computer-based systems security", *University of Newcastle upon Tyne*, 2010.
2. H. J. Highland, "A History of Computer Viruses -*The Famous 'Trio'," Computers & Security*, Vol. 60, No. 5, pp. 412-415, 1997.
3. A. Gazet, "Comparative analysis of various ransomware virii", *Journal in Computer Virology*, Vol. 6, No. 1, pp. 7790, 2010.
4. Barbara Guttman, Edward A. Roback, "An Introduction to Computer Security: The NIST Handbook",
5. *Computer Systems Laboratory, National Institute of Standards and Technology*, Gaithersburg, MD 208990001,1995.
6. Howard F. Lipson, "Tracking and Tracing CyberAttacks: Technical Challenges and Global Policy Issues", *PhD CERT ® Coordination Center, Networked Systems Survivability Program*, 2002.
7. Nwokedi Idika, Aditya P. Mathur, "A Survey of Malware Detection Techniques", *Department of Computer Science Purdue University*, West Lafayette, IN 47907, 2007. [7] Hao, S.,

Wang, W., Lu, H. and Ren, P. "AutoMal: automatic clustering and signature generation for malwares based on the network flow", *Security Comm. Networks*, 2014.

8. [8] Muazzam Ahmed Siddiqui, "Data mining methods for malware detection", PhD thesis, *College of Sciences, University of Central Florida*, Orlando, Florida, 2008. [9] Xue, L., Sun, G., "Design and implementation of a malware detection system based on network behavior", *Security Comm. Networks*, 2014.
9. Threat Expert, "Threat Expert," *Threat Expert*, [Online].*Available*: http://www.threatexpert.com/. [Accessed 20 01 2021]. COMODO, "COMODO Automated Analysis
10. System,"COMODO, [Online]. Available:http://camas.comodo.com/.[Accessed 19 01 2021].
11. M. Egele, T. Scholte, E. Kirda, C. Kruegel, "A Survey on Automated Dynamic Malware Analysis Techniques and Tools", *Journal ACM Computing Surveys,* Volume 44, Issue 2, Article No. 6, 2012.
12. M. Christodorescu, S. Jha. "Static analysis of executables to detect malicious patterns", *USENIX Security Symposium*, 2003.
13. F. Leder, B. Steinbeck, P. Martini, "Classification and Detection of Metamorphic Malware using Value Set Analysis", in 4th International Conference on Malicious and Unwanted Software (MALWARE), 2009.
14. M. G. Schultz, E. Eskin, E. Zadok, S. J. Stolfo, "Data mining methods for detection of new malicious executables", *In IEEE Symposium on Security and Privacy*, 2001.
15. P.Deshpande, "Metamorphic Detection Using Function Call Graph Analysis", *Master's Projects*, Paper 336 http://scholarworks.sjsu.edu/etd projects/336,2013
16. G. Shanmugam, R. Low, M. Stamp. "Simple Substitution Distance and Metamorphic Detection," *Journal of Computer Virology and Hacking Techniques*, Volume 9, Issue 3, pp. 159–170, 2013.
17. *Languages and Computing*, Volume 23, Issue 3,pp. 154– 162,2012.
18. Deylami, H.M., Muniyandi, R.C., Ardekani, I.T. and Sarrafzadeh, A., 2016, December. Taxonomy of malware detection techniques: A systematic literature review. In *2016 14th Annual Conference on Privacy, Security and Trust (PST)* (pp. 629636). IEEE.