OPEN ACCESS

50SEA

Check for updates

IJIST
P ISSN : 2618-1630
E ISSN : 2909-6130

Original Article

Computer Malware Classification, Factors, and Detection Techniques: A Systematic Literature Review (SLR)

Asad Hussain[1,2,*], Sunila Fatima Ahmad [2], Mishal Tanveer[2], Ansa Sameen Iqbal[2].
[1]University of Science and Technology, Bannu, Pakistan.
[2]University of Sargodha, Sub-Campus, Bhakkar, Pakistan.
*Correspondence: Asad Hussain, asadhussain@ustb.edu.pk, chasad1303@gmail.com

Citation |Hussain. A, Ahmad. F. S, Tanveer. M, Iqbal. S. A, "Computer Malware Classification, Factors, and Detection Techniques: A systematic literature review (SLR)," Int. J. Innov. Sci. Technol., vol. 4, no. 3, pp. 899-918, 2022.

Received | June 30, 2022; **Revised** | Aug 23, 2022; **Accepted** | Aug 25, 2022; **Published** | Aug 29, 2022.

DOI:https://doi.org/10.33411/IJIST/2022040320

A Systematic Literature Review (SLR) was conducted using tailored searches based on our study topic. We completed all SLR processes, including periodic reviews as SLR. Researchers may find out about the justification, the review procedure, and the research question by using search keywords. This paper describes the trial approach to elaborate the search keywords, resources, restrictions, and validations that were, and explores search strategies made. The reviews are carried out by assessing the publication's quality, devising a data extraction approach, and synthesizing the results. All four research questions were used to analyze the papers concerning the findings. Finally, reports on the categorization of computer malware were analyzed for their detection methods, factors, and how they infiltrate computer systems have been published. SLR identifies the element, characteristics, and detection techniques that are explained in this research paper. Computer malware infects the computer system. This comprehensive literature review's is mainly based on recommendations by earlier studies.

Keywords: Malware Classification, Malicious Software Factors, Malware Detection technique, Malicious Infection

Introduction

Malware is a contraction of two words malicious software, which causes substantial security threats in the computer world[1]. Classification is a collection of related categories used to collect data based on similarity. It is made of codes and descriptions that enable survey results to be organized into relevant groups, resulting in valuable data. Anyone doing statistical surveys will find a categorization beneficial. It is a framework that both simplifies the issue under investigation and makes categorizing all data or replies simple. Viruses, Worms, Trojans, Spam, Phishing, Ransomware, and other similar attack mechanisms fall under the broad term COMPUTER MALWARE. Lacing Malware in a particular malware family is known as malware classification [2]. Malware in the same category has characteristics that may be utilized to develop indicators for detection and categorization. Based on how they are retrieved, described as static or dynamic.

Detection is the practice of finding anything, whereas the technique is a specific way of doing an action, generally one that requires practical abilities. As a result, it refers to the method of finding and practically detecting something. Detection techniques of computer malware [3] described in this paper are behavioral-, anomaly-and signature-based, etc. A factor is one of the variables that impact an event, decision, or circumstance and proactively contributes to creating a result. It's among the cases that impact the outcome of anything[4]. We are providing the human behavior and technological factors.

Infection results in one or more infectious agents establishing themselves on or around the body for a suitable host. Essentially infecting anything or someone is acting or practice of contaminating for something. To accomplish the goal of causing as much harm as possible on many systems as possible, cybercriminals must install malware on the victim's computer, which may form the basis of executable scripts, active content, or software. Malware distribution through hacked websites become one of the most prevalent methods of infecting PCs in recent years. As a result, antivirus software, firewalls, and other security measures are used to protect computers from this danger. Infiltration channels, extraction framework, and infection markers are used to infect the computer system.

We analyzed all four questions by doing a deep analysis of the survey papers, articles, systematic literature reviews, and research papers of all the years from the four selected databases, i.e., IEEE, ACM, Google scholars, and Science direct.

Scheme of paper

In this SLR, section 1 consists of the introduction. Section 2 is about the background of malware. Section 3 describes the review methodology, which is briefly explained in further steps. In the last section, the conclusion is defined.

Background

The first computer virus [5] was discovered over 50 years ago. Virus developers began creating viruses in the early 1980s. Virus authors and hackers began to employ their skills for more professional and illegal purposes in the late 1990s and early 2000s. The term "malware" is a broad phrase that encompasses all threats, including viruses, worms, and Ransomware [6].

ARPANET, or Advanced Research Projects Agency Network, existed before the Internet. ARPANET was founded in 1967 to connect remote computers[6]. The Intel 4004 was designed in 1971 and was the first microprocessor. "The Creeper" was the world's first computer virus. "Elk Cloner" was the first computer virus discovered in the world. It was created by a 15-year-old who used to write programs like this to play a trick on his pals. The virus would stay in memory until it found a blank floppy disc to infect. The first computer virus, called "Brain," was released in 1986. It began in Pakistan and swiftly spread around the world[7]. Because there was no internet then, it was spread by human interaction and floppy disc copying. The world of information security as we know it today was forever transformed. The world's first worm was created more than 30 years ago. The morris worm was named after its creator Robert morris. This worm wasn't harmful in any way. It was made as a proof of concept to find out whether an infection was already present.

The AIDS Trojan was credited to the late Dr. Joseph Popp. About 20,000 contaminated floppy discs were sent by mail (yes, real mail, not email) to AIDS experts worldwide. The Trojan converted file names to encoded strings and concealed data from the user. Dr. Popp's floppy discs were not sent to scholars in the United States, which is an intriguing fact. Michelangelo was a boot sector virus that was designed to attack disk operating system DOS partitions. It was the first time a virus gained widespread coverage in the mainstream media. The mid-90s were groundbreaking for many who grew up in America online AOL chat groups. Numerous cyber attackers were likely to steal account credentials since dial-up internet service was relatively costly and was provided by the minute. Phishing [8] programs began to be distributed on illegal warez chat rooms. 1999 was marked by anxiety over the "Y2K" virus. Digital Subscriber Line (DSL) connections were beginning to gain momentum.

Cybercriminals took advantage of this, bringing in the period of botnets and infections. Botnets conjured up images of Sky net, the fictional corporate evil from the Terminator films. The EarthLink Spam Botnet, which initially appeared in 2000, was the first to be detected. On the other hand, the GT botnet was launched in 1999, constituting the first botnet. It was responsible for 25%, including all email spam, or over 1.25 billion messages. Threat actors still use worms; however, they aren't as prevalent as they once were. A worm [9] may spread far in a short period since it reproduces on its own. They compel an infected machine to grind to a standstill by consuming ever-increasing operating system cycles. Blaster used a Remote Procedure Call (RPC) exploit in Microsoft Windows XP and 2003 to spread globally. This was the first worldwide denial-of-service assault as a consequence of global use of internet access. It also carried out DDoS assaults on pre-determined targets. Mytob was among the first viruses to block or act despite antivirus software.

The virus shut down operations of over 100 companies, including the New York Times. CoolWebSearch, or "CWS," was the first criminal operation to hijack Google search results. In 2007, a similar incident occurred some years later. It was uncovered after a lady in Ohio paid thousands of dollars for a car that was never delivered. The FBI and Semantic patiently waited for the fraudsters to mess up, which ended in their indictments in 2016. Reveton was developed by the National Security Agency to spy on the European Union. Regin was able to adapt to a certain environment as a result of its flexibility. Reveton helped develop the look and feel of up-to-date Ransomware, including the universal lock screen.

Crypto Locker was the very first Ransomware [8] to demand Bitcoin as a payment method. The cost of cracking was two BTC, which was worth between $13 and $1,100 in 2013. This was back when bitcoin was in its infancy and convincing non-technical people to pay was a difficult task. Tech support scams and other browser locks variations debuted in 2015. These annoyance assaults resemble Ransomware in that they cause victims to worry. Other variants included Blue Screen of Death (BSOD) displays and a toll-free number purporting to be Microsoft technical help. They would then seize control of the victim's system and demand fees.

Mirai was the first botnet to attack Internet-of-Things (IoT) devices. It was primarily aimed at network routers but also affected other IoT devices. Mirai was so widespread at one time that cryptologist Bruce Schneier speculated it may have been a nation-state. NSA's Shadow Brokers leak was unprecedented and catastrophic. Hackers expertly recycled the tools and vulnerabilities that were provided. WannaCry, Petya/NotPetya ransomware strains were so destructive that they forced industrial plants all around the world to close down.

Cryptocurrency-related dangers were previously confined to Ransomware and bitcoin wallet thefts, but 2018 saw the introduction of a new technique. XMRig operates by exploiting a machine's unused CPU cycles to assist in the solving of certain mathematical problems that are utilized in bitcoin mining. Various criminal attackers used known attacks in Apache Struts, Oracle Weblogic, and Jenkins servers to take advantage of common vulnerabilities. These assaults were restricted to enterprises that utilized these technologies, as well as the strong CPUs of the machines they operated on. GandCrab aimed to separate itself from real assaults on companies while also increasing income. It mastered the Ransomware as a Service business model (RaaS). The GandCrab writers were able to focus on their code, whereas others handled the real breaches thanks to RaaS. It was eventually used to distribute the infamous WannaCry Ransomware, Petya/NotPetya, with terrible results.

Review Methodology

The goal of this study was to analyze computer malware in a focused manner using a technique called a systematic literature review. Figure 1 shows the review methodology of our SLR The research is based on PRISMA.

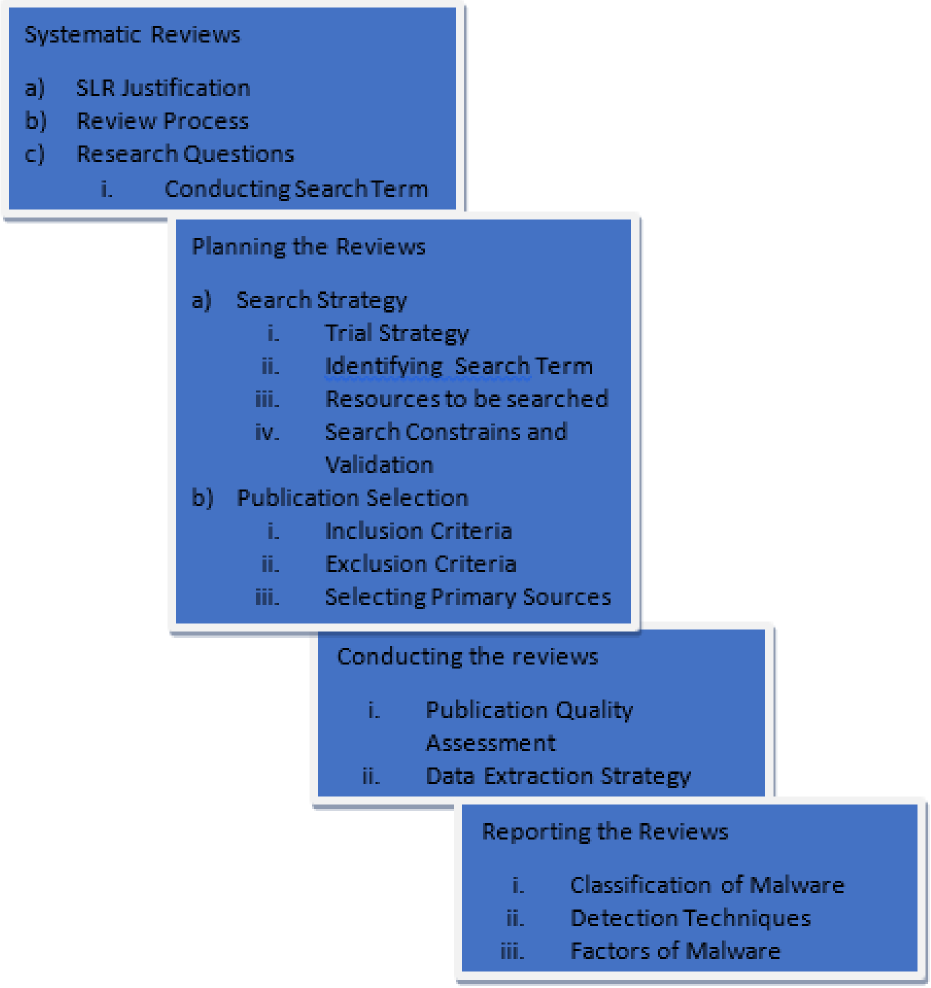

Figure 1. Review Methodology

Systematic Reviews

SLR Justification:

In former eras, malware classifications, factors, and detection approaches were thoroughly discussed in several research articles. We offer a thorough systematic literature review on the categorization, factors, and detection strategies of computer malware such as worms, viruses, Trojan horses, spam, phishing, and Ransomware in this SLR. Four questions address these important issues. The taxonomy of computer malware is the first question. The second question concerns computer virus detection methods. In question three, you'll learn about computer malware factors, and in question four, you'll learn about how computer malware infects a computer system.

Review Process:

This study's review methodology is based on the PRISMA criteria, a well-known review protocol. This comprehensive literature review's structure is mainly based on earlier studies' recommendations. For further information, we looked at surveys, publications, research papers, and studies undertaken.

Research Questions: The work detailed in this technical paper was inspired by four research questions:

a. What is the classification of computer malware attacks?
b. What are the detection techniques of computer malware attacks?
c. What are the factors of malware attacks?
d. How does malware infect the computer system?

Conducting Search Term:

The information here will assist us in creating a search phrase that is appropriate to our research questions.

Population: Computer Malware

Intervention: Classifications factors and detection techniques.

Outcomes of relevance: Way of infecting computer system

Planning The Reviews

Search Strategy

Trial Strategy:

On the Google Scholar (scholar.google.com) digital library, a sample search was done using the following query.

("Classification" OR "Characterization")AND ("Malware" OR "Injurious" OR "Malicious") AND ("Attack" OR "Strom").

("Detection" OR "Noticing")AND ("Techniques" OR "Approach" OR "Method") AND ("Malware" OR "Injurious" OR" Malicious") AND ("Attack" OR "Strom").

("Malware" OR "Injurious" OR "Malicious") AND ("Infect" OR "Harm") AND ("Computer systems").

The publications found using these search keywords will be utilized to help design and validate the primary search phrases.

Identification Search Terms:

The following search approach is employed to create search phrases.

a. Deduce essential words from the research questions by identifying population, intervention, and result;
b. Find alternate spellings and synonyms for significant words using the research questions;
c. Check any relevant paper's keywords;
d. If the database permits it, use the 'OR' operator for concatenation of alternate spellings and synonyms and the 'AND' operator for concatenating essential words.

Resources to be searched: The following digital libraries and databases are searched.

- IEEE Explore
- ACM Digital Library
- Google Scholar (scholar.google.com)
- Science Direct (sciencedirect.com)

Search Constraints and Validation:

On the Google Scholar (scholar.google.com) digital library, a sample search was done using the following search query.

("Classification" OR "Characterization")AND ("Malware" OR "Injurious" OR "Malicious") AND ("Attack" OR "Strom").

("Detection" OR "Noticing")AND ("Techniques" OR "Approach" OR "Method") AND ("Malware" OR "Injurious" OR "Malicious") AND ("Attack" OR "Strom").

("Malware" OR "Injurious" OR "Malicious") AND ("Infect" OR "Harm") AND ("Computer systems").

The publications found using these search keywords will be utilized to help design and validate the primary search phrases.

Publication Selection

The publication selection method will be carried out using publication inclusion and exclusion criteria, and primary source selection. The primary goal of this publication selection approach is to limit our search results to those relevant to our research concerns only augmented reality-related research metrial and vice versa material

Inclusion Criteria:

Only the literature (research papers/reports/books) discovered in the search results is selected using inclusion criteria. Only augmented reality-related research articles will be considered.

The following are the requirements for inclusion:

a. Documentation of the classifications of computer malware assaults.
b. Research that outlines computer malware detection techniques.
c. Research that outlines computer malware factors.
d. Research that outlines how malware infects computer systems.

Exclusion Criteria: Exclusion criteria are used to determine which pieces of literature (research papers, reports, and books) will not be reviewed based on the search term.

The following are the criteria:

- Research work that is not relevant to the research questions.
- Research work that doesn't describe factors, classification, and detection techniques.
- Research work other than computer malware.

Selecting Primary Sources:

Primary sources will be chosen first by looking at the titles, keywords, and abstracts of the material that has been searched. This review will disregard/exclude any literature that does not pertain to the study topics. The main sources picked during this first selection process will be evaluated against the aforementioned inclusion/exclusion criteria by reading the entire text of the research papers.

The case will be submitted to the secondary reviewer if there is any doubt about the inclusion/exclusion decision. The third reviewer will inspect the procedure.

Records of inclusion/exclusion decisions for each primary source shall be kept correctly. This will contain a rationale for including or excluding the primary source from the final evaluation.

Conducting The Reviews

Publication Quality Assessment:

When the final selection of publications is complete, the publication quality evaluation is carried out. This evaluation happens at the same time as the data extraction.

The following questions will be indicated as "Yes" or "No" or "partial" or "NA" on the quality evaluation checklist:

- Has a clear classification/categorization of computer malware assaults been established?
- Is there a clear identification of detection techniques?
- It has been recognized as a factor in computer malware.
- Is it obvious how computers get infected?

Data Extraction Strategy:

The primary goal of this SLR is to collect data from academic articles that are focused on answering our research objectives.

The following information will be retrieved from each study paper.

- Information about the publication (title, authors, journal/conference title, and other required information)
- Information that answers our research questions

The following information will be harvested to answer our study questions:

- RQ1: Background information and computer malware classifications;
- RQ2: Background information and computer malware detection techniques;
- RQ3: Background information, computer malware factors
- RQ4: Infect computer systems, background information

Data Synthesis:

There will be four portions to the data synthesis: Q1, Q2, Q3, and Q4. The first part (for question 1) will categorize computer malware. The detection techniques for question 2 will be maintained in the second portion, while the detection methods will be preserved in the third section, and how malware infects computers will be shown in the fourth section.

Selection Criteria: Figure 2 shows the four phases of selection criteria.

First phase:

We utilized numerous literature studies on computer malware classification, detection methodologies, and aspects to develop the search phrases. We began by going through the words in the next section. After the first phase, 239 items had been gathered.

Key terms: Following key terms and search queries for each question are explored in different repositories.

- Classification: (“Classification” OR “Categorization” OR “Classifying” OR “Grouping” OR “Grading” OR “Ranking” OR “Sorting” OR “Stratification” OR “Systemization” OR “Organization” OR “Codification”)
- Malware: (“Injurious” OR “Mean” OR “Nasty” OR “Virulent” OR “Despiteful” OR “Hateful” OR “Unkind” OR “Awful” OR “Evil” OR “Gross” OR “Bad-natured”)
- Attacks: (“Harm” OR “Hiding” OR “Strike” OR “Violent-act” OR “Wall of fire” OR “Offensive” OR “Hate-crime” OR “Crushing” OR “Rade” OR “Storm” OR “Act of war”)
- Detection: (“Observation” OR “Noticing” OR “Noting” OR “Perception” OR “Spotting” OR “Awareness” OR “Recognition” OR “Distinguishing” OR “Identification” OR “Diagnose” OR “Sensing”)

- Techniques: (“Method” OR “Approach” OR “Procedure” OR “Process” OR “System” OR “Method of working” OR “Tactics” OR “Strategy” OR “Practice” OR “Plan” OR “Manner”)
- Factor: (“Aspect” OR “Cause” OR “Circumstance” OR “Element” OR “Part” OR “Ingredient” OR “Component” OR “Characteristics” OR “Feature” OR “Event” OR “Item”)
- Infect: (“Defile” OR “Effect” OR “Ruin” OR “Crush” OR “Break” OR “Strike down” OR “Do damage to” OR “Mess up” OR “Splash” OR “Make ill” OR “Distort”)
- Computer System: (“Computer devices” OR “Systems”)

Second phase:

The second stage involves tool-based filtering due to the enormous number of publications collected in the previous phase. We were able to find duplicate research from various sources and versions by filtering using the citation tool. We had a total of 72 papers from multiple repositories after this round.

Third phase:

The emphasis during this phase was on "abstracts" rather than the complete metadata. Irrelevant data may be deleted by looking at every paper's abstract; this approach eliminated 8 papers, leaving us with 64 articles in our collection.

Fourth phase:

Finally, all of the remaining articles were thoroughly reviewed. As a result of this approach, articles that did not satisfy the selection criteria or quality, or were beyond the subject of our study, were excluded. Six papers were eliminated at this step, leaving 58 research articles to perform the study. The SLR selection process is given also.

Result And Discussion

The intense, in-depth study we conducted yielded significant results. In this section, we summarize our findings, followed by a discussion. All four research questions were used to analyze the papers concerning the findings. Before the research questions are discussed, selected statistics are presented in Table 1 and Figure 3. Finally, following a detailed analytical discussion, we offer open research challenges to researchers in the field.

Table 1 Data Sources

Sr. No.	Questions	ACM	IEEE	Google Scholar	Science Direct	Total Papers
1	Q1	7	1	2	1	11
2	Q2	6	5	10	2	23
3	Q3	2	1	7	0	10
4	Q4	3	1	9	1	14

Classification of Malware:

Static analysis and dynamic analysis are two types of malware analysis approaches. To examine binary data and develop detection fingerprints, static analysis is utilized. Dynamic analysis involves executing malware in a controlled environment and recording and analyzing the executed instructions. Static anti-malware measures are more effective in detecting malware.

They suggested a novel paradigm for identifying and classifying malware activity that uses a dynamic approach. They do dynamic analysis on malware samples to gather API call sequences, which are then fed into our word embedding module.

Any Artificial Intelligence technology will be used to optimize the categorization process. It will contribute to developing a new categorization approach that incorporates artificial intelligence. This study's results led to identifying important issues that need to be addressed, including the efficiency and scalability of malware categorization to scale up to millions. Few-shot malware classification offered insight into making current classification methods more scalable.

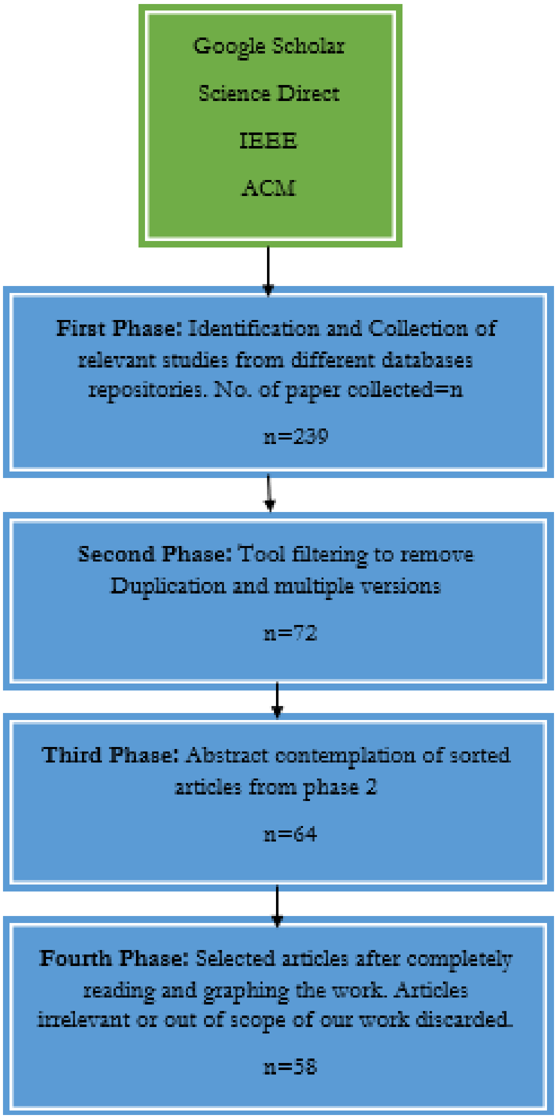
Google Scholar
Science Direct
IEEE
ACM
First Phase: Identification and Collection of relevant studies from different databases repositories. No. of paper collected=n
n=239
Second Phase: Tool filtering to remove Duplication and multiple versions
n=72
Third Phase: Abstract contemplation of sorted articles from phase 2
n=64
Fourth Phase: Selected articles after completely reading and graphing the work. Articles irrelevant or out of scope of our work discarded.
n=58

Figure 2 The SLR selection process and filtering in each phase

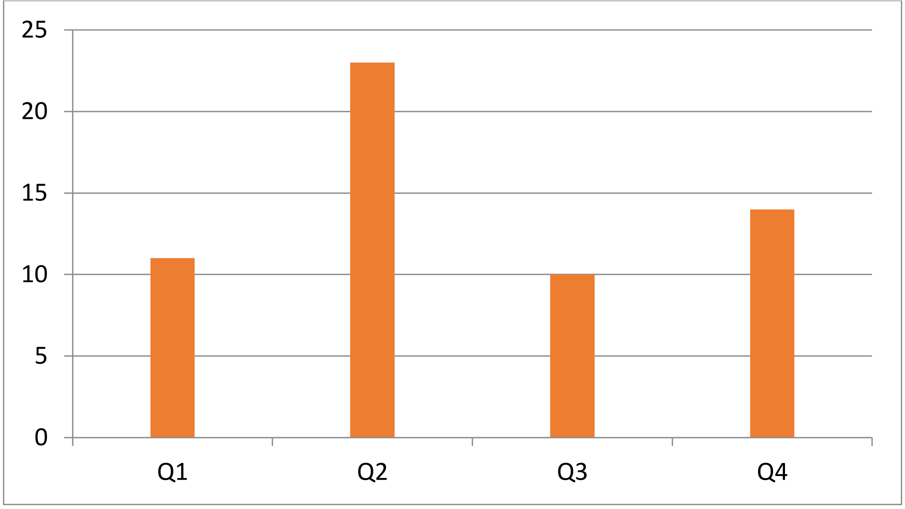

Figure 3 Data source

They're testing the malware classification model's susceptibility by adding an adversarial sample into the datasets to impair the quality of classification presently produced by a trained model. New approaches, such as polymorphic and metamorphic, have substantially boosted the quantity of malware in use today. Similar classification approaches will be used to detect the common behavior of each malware family.

Three primary goals are outlined in the proposed framework. The first goal is to provide a safe environment for malware analysis based on behavior. The second step is to use a thorough approach to malware analysis. The final goal is to classify malware into a new conceivable category. Bazrafshan et al. divide the approach into three categories: 1: signature-based methods. 2: depending on conduct 3: based on heuristics.

Table 2 Classification of Malware

Sr. No.	Papers	Classification
1.	[10], [11], [12], [13]	Static and dynamic classifications
2.	[14], [15]	Dynamic analysis
3.	[11],[14]	Architecture and comprehensive approach, classified malware
4.	[11], [14]	Quantitative measurement
5.	[11], [14]	Artificial intelligent technique.
6.	[16], [15]	Efficiency and scalability
7.	[17]	Adversarial
8.	[14]	Polymorphicand metamorphic techniques
9.	[16],[14]	Signature, behaviorand heuristic-based classification.
10.	[12], [18]	Malware detection

Detection Techniques of Malware:

Because the training set generates many rules and it is impossible to design a classifier that uses all of them, post-processing of correlative classification is essential for enhancing the classifier's accuracy and efficiency. Rule trimming, rule ranking, and rule selection are some strategies used.

Signature-based malware detection is effective for identifying known malware but falls short when detecting new malware. Heuristic-based detection relies on experience and other strategies such as rules and machine learning. Although it can identify zero-day malware with a high degree of accuracy, it cannot detect sophisticated malware. The use of a sandbox allows for automatic analysis. System call monitoring, file change monitoring, registry snapshot comparison, network activity monitoring, and process monitoring.

Signature-based approaches are more efficient and quicker than other methods because they employ patterns retrieved from malware to detect them. These technologies cannot identify new malware variants, and extracting unique signatures requires a significant amount of people, time, and money. Another issue is the inability to combat malware that mutates its programs with each infection, such as polymorphic and metamorphic malware. Behavior-based approaches are resistant to the flaws of signature-based techniques. These approaches aggregate programs that have the same behavior. Because they constantly utilize the same resources and services on your computer. These detection measures may uncover malware that continually produces new mutations.

The ultimate purpose of concealing is to change data and provide a fake response to the higher query function. General programs may scan the system by invoking the API method from the user mode. Once the variations between the detected data and the ones in the raw data library are discovered, it may be determined the data in the top layer has been changed by malware. Malware using concealing technologies may have infiltrated the system, and the higher calling return data should not be believed. SAFE is a malware detection method that uses static analysis to discover harmful executable patterns. This approach builds templates from the assembly code of each known malware program. Templates are instruction sequences specified using variables and symbolic constants. The primary benefit of this technique over SAFE is that it focuses on program semantics and defines program behaviors at a higher level.

Deep convolution neural networks with fine-tuned parameters are used to do the detection. Mal-Detect can be used to detect all types of malware. A dynamic taint analysis technique looks at how tainted data was transported between system calls. This study's findings have ramifications for detecting and visualizing new malware. Malware detection technologies identify and counter harmful programs that may affect a computer system. To identify and ze malware samples into an appropriate family, the input is represented in various ways. There are three primary categories for the suggested malware detection approaches.

Table 3 Detection Techniques of Malware

Sr. No.	Papers	Detection Techniques
1.	[19],[20],[21],[22],[23],[24],[25], [26]	Signature-based
2.	[27],[28], [29], [30]	Machine learning-based
3.	[19], [22]	Anomaly-based techniques
4.	[19],[20], [21], [22], [24], [25]	Heuristic-based methods
5.	[19], [20]	Specification-based
6.	[19],[26]	Hybrid-based
7.	[19], [24]	Cloud-based
8.	[19]	Agent-based Virtual machine Rule-based Network-based
9.	[28], [24], [30]	Deep learning methods
10.	[19],[31]	Host-based
11.	[32]	Using Engine Signature

		Hidden Markov Model (HMM) Detection
12.	[29], [24],[33]	DATA MINING
13.	[24]	MOBILE DEVICES-BASED IOT-BASED
14.	[34]	Graph-based
15.	[35]	Deep Convolutional Neural Network Deep Generative Adversarial Neural Network Mal-Detect
16.	[30]	Convolutional neural network-based malware detection Recurrent Neural network-based malware detection
17.	[36]	SNORT Intrusion Detection System

Factors of Malware:

Investigate users' perceptions and attitudes regarding computer security choices. It largely uses qualitative approaches to learn how and why people interact with computers, including surveys, interviews, and observations. Malware encounters may be predicted using the user's demographic and behavioral characteristics, men and those with technical knowledge are more likely to experience malware.

For each user in our dataset, we evaluate two coarse-grained demographic features their age and the nation from which they access their account. These ages in our dataset are divided into ten-year chunks due to anonymization, beginning with "18 to 24" and ending with "65 or older". Age is a role in vulnerability in previous research, which we believe may contribute to increased rates of attacker targeting. Similarly, wealth correlates with age and country of access which we predict may be a role in targeting owing to economic rewards for attackers.

SCADA (Supervisory Control and Data Acquisition) systems have low computation CPUs, limited memory, and low power consumption. As a result, while preparing for ransomware security, such resources' effective use should be considered. Assailants may easily create numerous sorts of Ransomware against SCADA equipment using Ransomware-as-a-Service (RaaS). (1) Inadequate backup policies and procedures are in place. (2) Inadequate User Awareness (3) Payment that is anonymous and secure.

Factors affecting resources: This form of SCADA affects how attackers could access the systems. Because SCADA equipment is not meant to keep personal data, they are particularly vulnerable to locker ransomware. Because of the resource constraints within SCADA CPU, memory, storage, and battery, IoT device maneuverability is limited. To guard against these assaults, it's critical to combat locker ransomware. On the other side, crypto-ransomware might hold the data produced and gathered by various sensors hostage during collection or transfer. Attackers might take advantage of SCADA systems' vulnerability to take control of services or data. As a result, attackers use several operational-related elements that contribute to the assaults' success, as follows. (1) Data (2) Services (3) inadequate knowledge.

The risk for falling prey to two forms of phishing greater phishing (which uses harmful software) and small phishing (which does not use malicious software). The parallels between (low-tech) phishing and (high-tech) malware assaults are striking. Both types of assaults, for example, follow a nearly identical criminal script. In addition both kinds of criminals are content with whatever they can grab and make no distinction between affluent and poor. There are other distinctions.

To begin with nearly none of the factors influenced phishing victimization when it came to online visibility. Malware victims on the other hand, tell a different narrative. There is a direct link between spending time on the Internet and being a victim. Downloading and online gaming are two sorts of behaviors that increase your chances of being a victim. The second distinction is in terms of internet accessibility. Whenever it came to phishing, none of the factors mattered. However, the malware study reveals that users of specific (popular and extensively used) operating system technology and web browsers are at a higher risk of being hacked.

Value: Respondents' financial qualities had no bearing on phishing victimization. None of the wealth-related characteristics (such as income or savings) can explain why someone is likelier or less likely to become a victim. The findings reveal that phishing attempts aren't only targeted at prospective victims.

Visibility: Certain acts raise the chance of becoming a malware victim. People who shop online are often more likely to get infected with malware. According to a study, online and web browsing (including targeted and untargeted) increases the chance of malware infection.

Accessibility: It indicates that the usage of commonly used software and web browsers has no impact on the probability of becoming a phishing victim. Table 1 illustrates that using the Windows operating system and the Firefox browser increases the probability of becoming a victim. Virus scanners also don't protect consumers against emails that try to convince them to provide personal information (phishing).

Online lifestyle: In this survey, just 15% of police departments (2 out of 13 incidents) could not pay a ransom sought by Ransomware, compared to 85% who paid. Using the most recent version of cyber security and a secure backup system may help reduce the chance of data loss.

This study implies that cybercriminals used spear phishing to attack police agencies' networks with Ransomware as they understood what departments they were going after. Ransomware uses a variety of infection vectors to propagate stealthily and undetected onto victims' computers. Malicious emails, brute-force login credentials, drive-by program downloads, and exploit kits are a few examples. Unaware individuals might get infected with Ransomware via clicking on websites that promise monetary rewards or free software. Ransomware also uses other attack channels, such as exploiting server weaknesses and self-propagation.

Ransomware has been around since the late 1980s. In 2005, the first wave of contemporary Ransomware appeared. The current spike in ransomware assaults has been attributed to several facilitators. Financial income, the availability of cryptographic technology, and undetectable payment methods are among them. Anonymous, peer-to-peer (P2P), and decentralized. Ransomware-as-a-Service (RaaS) has lately appeared in the form of cloud-based services that give malware authors a development environment.

It has identified possible malware victimization risk factors however, its findings are limited, so they are dependent on self-reported rates of infection rather than real malware detections. Microsoft identified technical risk factors associated with malware infestations using telemetry data from the Microsoft Products Removal Tool (MPRT). They discovered, individuals who do not use any antivirus software or who use an out-of-date expired or snoozed antivirus are 5.6 times more likely to be infected than those who use up-to-date security.

Because they lack a reliable technique for correlating user behavior with computer activity, demographic, web, and network traffic studies cannot be utilized to investigate the connection between demographics and online danger.

The susceptibility factors discovered were divided into three categories. Demographics are the characteristics of a population's volume, structure, and distribution (e.g. gender, age, education/training, experience, and so on).

Personality refers to a person's thought process, attitudes, and actions.

Culture is divided into two groups. Culture at the national and organizational levels: Cultural identity refers to the culture unique to a grouping of people in a particular geographical region. In contrast, an organization's culture relates to a culture linked with a certain company or organization.

Demography is an essential aspect in determining the risk of computer infections. Aged, gender, expertise, education, work situation, and economic level are among them.

Understanding the actual value of an individual's personal information (and hence valuing that payment of a ransom is much less valuable even than personal information, e.g., photos or financial information) or trying to make the hush money payment process as seamless as possible are two examples of human factors in this context.

We may look for putative variables connected with an outcome with more than 95% confidence. To properly assess relative risk, the preliminary identification of such variables may be utilized to create stronger epidemiological investigations, including such cohort or randomized control trials.

Malware Infection:

The detected indications should be unaffected by the changing threat environment, including Polymorphic and Metamorphic malware. Malware is known to have a heartbeat and periodic communications that inform the relevant C&C servers of the malware's existence and activities in the target network. The average second-order time differences of periodic messages in a link are used to determine periodicity.

The computer's suspicious files were divided into four categories: harmful, suspect, safe, and unrated. Additional data was gathered when we thought a file was malicious, including browser history, the suspicious file, and any linked files. On 12 different devices, the investigation discovered 20 probable infections.

The availability of a public-key version means that every computer may verify the security of any other machine. it may be appears as useful but there is privacy and security reasons to avoid it so, We conclude that architecture with only one or a few dedicated servers is preferable to one in which any computer may query any other machine.

Table 4 Factors of Malware

Sr. No	Papers	Factor
1.	[37],[4]	Human factor
2.	[4, 37]	Behavioral factors
3.	[37],[38]	Demographics factor
4.	[39]	Potential risk factors
5.	[39]	Technical Risk factors

6.	[4]	Susceptibility factors
7.	[4]	Culture factors
9.	[40]	Operational factors
10.	[40]	Resources factors
11.	[41],[42]	Lifestyle, putative factors
12.	[43]	Risk factor of phishing
13.	[44]	Effective infection vector
14.	[4]	Personality factor

Pretexting assaults include creating fictitious and convincing scenarios to acquire victim's personal information. Bogus software assaults use fake websites to trick consumers into thinking they're dealing with well-known and reputable software or websites. The term "pop-up window" refers to windows on a victim's screen telling them their connection has been lost. Rob call assaults have lately surfaced as large-scale phone calls made by machines to people with known phone numbers.

Email, social software, websites, portable storage devices, and mobile devices are all common places for social engineering malware to spread. Social engineering uses emotions like fear, curiosity, enthusiasm, empathy, greed, and cognitive biases to manipulate people. Multiple infections provide no benefit to the malware's performance, but they may be hazardous to the system's stability since malware often lurks within the operating system. To be effective, an infection marker must be both persistent and predictable. Each kind of infection marker has its characteristics that affect its use as a vaccination. We developed a system that can automatically analyze malware samples and construct a vaccination program. Although this is a proof-of-concept for Microsoft Windows malware, the principle may be applied to other architectures and the operating system's underlying architecture.

Regular ICT malware, including threats that aren't present in traditional business networks, is especially vulnerable to SCADA systems. The Code Red, Nimda, Slammer, and Scalper programs attacked SCADA systems with malware from four well-known cases. Scalper, Nimda, and Slammer successfully infected process network machines. The control flow between the SCADA system and the operator PCs linked through the Intranet was interrupted. The most significant consequence was the possibility of compromised servers executing arbitrary code. Malware may be programmed to work through firewalls and antivirus protection. Malware uses flaws in the Modbus protocol, posing a severe threat to industrial control systems. The more significant problem of malware that mainly targets SCADA systems in process networks is discussed in this section.

Smart malware is highly complicated dynamic virus that uses several strategies to accomplish its goals. The owner of smart malware may change the objectivs at any time and update the infection's components accordingly. Smart malware has three goals: primary, secondary, and fake. Malware is computer software that is meant to infiltrate and corrupt computer systems. Logic bombs, viruses, worms, and botnets are a few examples. Malware may hold harmful payloads that are delivered when susceptible hosts are infected. It can potentially do substantial harm, including network bandwidth usage and host destruction.

Table 5 Malware Infection

Sr. No.	**Papers**	**Malware Infection**
1.	[45]	Early indicators challenges and features
2.	[46]	Identification of threats and suspicious files

3.	[47],[48]	Infiltration channels and tactics
4.	[49], [50],[51]	Extraction framework
5.	[52]	Infection marker mechanism
6.	[53]	ICT and SACADA malware
7.	[52], [54]	Attack classification and description
8.	[54],[50],[55]	Smart malware and malware infection
9.	[55]	Malware propagation

Conclusion

Most of the papers related to SLR are identically explained, but no one provided the complete SLR on Classification, Detection techniques, and factors. This paper presents a complete SLR on the factors affecting the computer system, and the method through which the computer malware infects the computer system also.[56]-[67]

References

[1] D. Uppal, R. Sinha, V. Mehra, and V. Jain, "Malware detection and classification based on extraction of API sequences," in 2014 International conference on advances in computing, communications and informatics (ICACCI), 2014, pp. 2337-2342.

[2] W. Gharibi, "Studying and Classification of the Most Significant Malicious Software," arXiv preprint arXiv:1106.0853, 2011.

[3] S. Divya, "A survey on various security threats and classification of malware attacks, vulnerabilities and detection techniques," International Journal of Computer Science & Applications (TIJCSA), vol. 2, 2013.

[4] A. A. Younis, E. Stronberg, and S. Noor, "User's Susceptibility Factors to Malware Attacks: A Systemic Literature Review," International Journal of Computer and Information Engineering, vol. 15, pp. 543-554, 2021.

[5] T. M. Chen and J.-M. Robert, "The evolution of viruses and worms," Statistical methods in computer security, vol. 1, pp. 1-16, 2004.

[6] F. Syed, "Understanding worms, their behaviour and containing them," Project Report, 2009.

[7] S. Gupta, "Types of Malware and its Analysis," International Journal of Scientific and Engineering Research, vol. 4, pp. 1-13, 2013.

[8] V. Bhavsar, A. Kadlak, and S. Sharma, "Study on phishing attacks," Int. J. Comput. Appl, vol. 182, pp. 27-29, 2018.

[9] B. Rajesh, Y. J. Reddy, and B. D. K. Reddy, "A survey paper on malicious computer worms," International Journal of Advanced Research in Computer Science and Technology, vol. 3, pp. 161-167, 2015.

[10] B. Anderson, C. Storlie, and T. Lane, "Improving malware classification: bridging the static/dynamic gap," in Proceedings of the 5th ACM workshop on Security and artificial intelligence, 2012, pp. 3-14.

[11] B. Kang, T. Kim, H. Kwon, Y. Choi, and E. G. Im, "Malware classification method via binary content comparison," in Proceedings of the 2012 ACM Research in Applied Computation Symposium, 2012, pp. 316-321.

[12] Y. Guo and W. Fan, "Feature collection and selection in malware classification," in Proceedings of the 2019 International Conference on Artificial Intelligence and Advanced Manufacturing, 2019, pp. 1-5.

[13] E. Gandotra, D. Bansal, and S. Sofat, "Integrated framework for classification of malwares," in Proceedings of the 7th International Conference on Security of Information and Networks, 2014, pp. 417-422.

[14] M. F. Zolkipli and A. Jantan, "An approach for malware behavior identification and classification," in 2011 3rd International Conference on Computer Research and Development, 2011, pp. 191-194.

[15] P. Wang, Z. Tang, and J. Wang, "A novel few-shot malware classification approach for unknown family recognition with multi-prototype modeling," Computers & Security, vol. 106, p. 102273, 2021.

[16] A. Abusitta, M. Q. Li, and B. C. Fung, "Malware classification and composition analysis: A survey of recent developments," Journal of Information Security and Applications, vol. 59, p. 102828, 2021.

[17] G. Raju, P. Zavarsky, A. Makanju, and Y. Malik, "Vulnerability assessment of machine learning based malware classification models," in Proceedings of the Genetic and Evolutionary Computation Conference Companion, 2019, pp. 1615-1618.

[18] K. S. Han, B. Kang, and E. G. Im, "Malware classification using instruction frequencies," in Proceedings of the 2011 ACM Symposium on Research in Applied Computation, 2011, pp. 298-300.

[19] I. A. Saeed, A. Selamat, and A. M. Abuagoub, "A survey on malware and malware detection systems," International Journal of Computer Applications, vol. 67, 2013.

[20] R. Tahir, "A study on malware and malware detection techniques," International Journal of Education and Management Engineering, vol. 8, p. 20, 2018.

[21] P. Singh, S. Tapaswi, and S. Gupta, "Malware detection in pdf and office documents: A survey," Information Security Journal: A Global Perspective, vol. 29, pp. 134-153, 2020.

[22] M. Naseer, J. F. Rusdi, N. M. Shanono, S. Salam, Z. B. Muslim, N. A. Abu, et al., "Malware Detection: Issues and Challenges," in Journal of Physics: Conference Series, 2021, p. 012011.

[23] A. Souri and R. Hosseini, "A state-of-the-art survey of malware detection approaches using data mining techniques," Human-centric Computing and Information Sciences, vol. 8, pp. 1-22, 2018.

[24] Ö. A. Aslan and R. Samet, "A comprehensive review on malware detection approaches," IEEE Access, vol. 8, pp. 6249-6271, 2020.

[25] Q.-L. Han, Y.-J. Hao, Y. Zhang, Z.-P. Lu, and R. Zhang, "A new malware detection method based on raw information," in 2008 International Conference on Apperceiving Computing and Intelligence Analysis, 2008, pp. 307-310.

[26] J. Singh and J. Singh, "A survey on machine learning-based malware detection in executable files," Journal of Systems Architecture, vol. 112, p. 101861, 2021.

[27] K. O. Babaagba and S. O. Adesanya, "A study on the effect of feature selection on malware analysis using machine learning," in Proceedings of the 2019 8th international conference on educational and information technology, 2019, pp. 51-55.

[28] D. Gibert, C. Mateu, and J. Planes, "The rise of machine learning for detection and classification of malware: Research developments, trends and challenges," Journal of Network and Computer Applications, vol. 153, p. 102526, 2020.

[29] J. R. S. Alrzini and D. Pennington, "A review of polymorphic malware detection techniques," International Journal of Advanced Research in Engineering and Technology, vol. 11, pp. 1238-1247, 2020.

[30] M. Sahin and S. Bahtiyar, "A Survey on Malware Detection with Deep Learning," in 13th International Conference on Security of Information and Networks, 2020, pp. 1-6.

[31] Z. Bazrafshan, H. Hashemi, S. M. H. Fard, and A. Hamzeh, "A survey on heuristic malware detection techniques," in The 5th Conference on Information and Knowledge Technology, 2013, pp. 113-120.

[32] S. kumar Sasidharan and C. Thomas, "A survey on metamorphic malware detection based on hidden Markov model," in 2018 International conference on advances in computing, communications and informatics (ICACCI), 2018, pp. 357-362.

[33] F. Manavi and A. Hamzeh, "A new approach for malware detection based on evolutionary algorithm," in Proceedings of the Genetic and Evolutionary Computation Conference Companion, 2019, pp. 1619-1624.

[34] Z. Shafiq and A. Liu, "A graph theoretic approach to fast and accurate malware detection," in 2017 IFIP Networking Conference (IFIP Networking) and Workshops, 2017, pp. 1-9.

[35] O. J. Falana, A. S. Sodiya, S. A. Onashoga, and B. S. Badmus, "Mal-Detect: An intelligent visualization approach for malware detection," Journal of King Saud University-Computer and Information Sciences, 2022.

[36] S. Kim, T. Kim, and E. G. Im, "Real-time malware detection framework in intrusion detection systems," in Proceedings of the 2013 Research in Adaptive and Convergent Systems, ed, 2013, pp. 351-352.

[37] F. L. Lévesque, S. Chiasson, A. Somayaji, and J. M. Fernandez, "Technological and human factors of malware attacks: A computer security clinical trial approach," ACM Transactions on Privacy and Security (TOPS), vol. 21, pp. 1-30, 2018.

[38] C. Simoiu, A. Zand, K. Thomas, and E. Bursztein, "Who is targeted by email-based phishing and malware? measuring factors that differentiate risk," in Proceedings of the ACM Internet Measurement Conference, 2020, pp. 567-576.

[39] F. L. Lévesque, J. M. Fernandez, and A. Somayaji, "Risk prediction of malware victimization based on user behavior," in 2014 9th international conference on malicious and unwanted software: The Americas (MALWARE), 2014, pp. 128-134.

[40] M. Gazzan, A. Alqahtani, and F. T. Sheldon, "Key Factors Influencing the Rise of Current Ransomware Attacks on Industrial Control Systems," in 2021 IEEE 11th Annual Computing and Communication Workshop and Conference (CCWC), 2021, pp. 1417-1422.

[41] K.-s. Choi, T. Scott, and D. P. LeClair, "Ransomware against police: diagnosis of risk factors via application of cyber-routine activities theory," International Journal of Forensic Science & Pathology, 2016.

[42] M. Lee, "Who's next? identifying risks factors for subjects of targeted attacks," in Proc. Virus Bull. Conf, 2012, pp. 301-306.

[43] E. R. Leukfeldt, "Comparing victims of phishing and malware attacks: Unraveling risk factors and possibilities for situational crime prevention," arXiv preprint arXiv:1506.00769, 2015.

[44] 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.

[45] S. Karapoola, C. Rebeiro, U. Parekh, and K. Veezhinathan, "Towards Identifying Early Indicators of a Malware Infection," in Proceedings of the 2019 ACM Asia Conference on Computer and Communications Security, 2019, pp. 679-681.

[46] F. Lalonde Levesque, J. Nsiempba, J. M. Fernandez, S. Chiasson, and A. Somayaji, "A clinical study of risk factors related to malware infections," in Proceedings of the 2013 ACM SIGSAC conference on Computer & communications security, 2013, pp. 97-108.

[47] F. Salahdine and N. Kaabouch, "Social engineering attacks: A survey," Future Internet, vol. 11, p. 89, 2019.

[48] S. Mohurle and M. Patil, "A brief study of wannacry threat: Ransomware attack 2017," International Journal of Advanced Research in Computer Science, vol. 8, pp. 1938-1940, 2017.

[49] A. Wichmann and E. Gerhards-Padilla, "Using infection markers as a vaccine against malware attacks," in 2012 IEEE International Conference on Green Computing and Communications, 2012, pp. 737-742.

[50] T. J. Holt, G. W. Burruss, and A. M. Bossler, "Assessing the macro-level correlates of malware infections using a routine activities framework," International journal of offender therapy and comparative criminology, vol. 62, pp. 1720-1741, 2018.

[51] W. Xiong and R. Lagerström, "Threat modeling–A systematic literature review," Computers & security, vol. 84, pp. 53-69, 2019.

[52] M. Jakobsson and A. Juels, "Server-side detection of malware infection," in Proceedings of the 2009 workshop on New security paradigms workshop, 2009, pp. 11-22.

[53] I. N. Fovino, A. Carcano, M. Masera, and A. Trombetta, "An experimental investigation of malware attacks on SCADA systems," International Journal of Critical Infrastructure Protection, vol. 2, pp. 139-145, 2009.

[54] Ş. Bahtiyar, "Anatomy of targeted attacks with smart malware," Security and Communication Networks, vol. 9, pp. 6215-6226, 2016.

[55] M. A. H. Saeed, "Malware in computer systems: Problems and solutions," IJID (International Journal on Informatics for Development), vol. 9, pp. 1-8, 2020.

[56] Y. Ye, Q. Jiang, and W. Zhuang, "Associative classification and post-processing techniques used for malware detection," in 2008 2nd International Conference on Anti-counterfeiting, Security and Identification, 2008, pp. 276-279.

[57] J. R. Nurse, "Cybercrime and you: How criminals attack and the human factors that they seek to exploit," arXiv preprint arXiv:1811.06624, 2018.

[58] D. Kong and G. Yan, "Discriminant malware distance learning on structural information for automated malware classification," in Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining, 2013, pp. 1357-1365.

[59] D. V. Sang, D. M. Cuong, and L. T. B. Cuong, "An Effective Ensemble Deep Learning Framework for Malware Detection," in Proceedings of the Ninth International Symposium on Information and Communication Technology, 2018, pp. 192-199.

[60] Y. Zhang, B. Bhargava, and P. Hurni, "The effects of threading, infection time, and multiple-attacker collaboration on malware propagation," in 2009 28th IEEE International Symposium on Reliable Distributed Systems, 2009, pp. 73-82.

[61] F. Mbol, J.-M. Robert, and A. Sadighian, "An efficient approach to detect torrentlocker ransomware in computer systems," in International Conference on Cryptology and Network Security, 2016, pp. 532-541.

[62] P. P. Kundu, L. Anatharaman, and T. Truong-Huu, "An Empirical Evaluation of Automated Machine Learning Techniques for Malware Detection," in Proceedings of the 2021 ACM Workshop on Security and Privacy Analytics, 2021, pp. 75-81.

[63] S. Saxena and S. Mancoridis, "Malware Detection using Behavioral Whitelisting of Computer Systems," in 2019 IEEE International Symposium on Technologies for Homeland Security (HST), 2019, pp. 1-6.

[64] M. Elingiusti, L. Aniello, L. Querzoni, and R. Baldoni, "Malware detection: A survey and taxonomy of current techniques," Cyber threat intelligence, pp. 169-191, 2018.

[65] J. Kim and B.-R. Moon, "New malware detection system using metric-based method and hybrid genetic algorithm," in Proceedings of the 14th annual conference companion on Genetic and evolutionary computation, 2012, pp. 1527-1528.

[66] A. M. Bossler and T. J. Holt, "Online activities, guardianship, and malware infection: An examination of routine activities theory," International Journal of Cyber Criminology, vol. 3, 2009.

[67] S. Abraham and I. Chengalur-Smith, "An overview of social engineering malware: Trends, tactics, and implications," Technology in Society, vol. 32, pp. 183-196, 2010.