Master of Science in Internetworking

Department of Electrical and Computer Engineering

Project Title:

Analysis of Current Machine Learning and AI Techniques to Perform Automated Hacking

Supervisor:

Leonard Rogers

Provided By:

Vishnu Ramakrishnan V

Fall 2020- Winter 2021

Table of Contents

Table of Figures

List of Tables

Abstract

Automated hacking is a modern approach to conduct computer hacking with a significantly higher performance rate directly or remotely. Automated hacking is carried out by computer programs that are programmed to analyze the vulnerability of a system(s) and carry out hacking attacks based on the analysis. Automated hacking is considered one of the most sophisticated computer programming areas with high integrity in coding and running platform-independent.

In recent years, the advancement of Artificial Intelligence (AI) and machine learning algorithms has influenced many information technology areas. Automated hacking is one such area in which AI and machine learning have shown a drastic improvement in the efficiency and execution power of cyber intrusions with the capability of self-awareness.

This paper focuses on the role of Machine Learning and Artificial Intelligence (AI) techniques in executing various levels of automated hacking. Explains the difference between the normal mode of hacking and automated hacking, how Machine Learning and AI evolved to support automated hacking, how successful these techniques are, and how systems respond to these types of hackings. The research on this topic will also focus on how these various techniques are used for both offensive and defensive purposes in automated hacking as well as how the industries and law enforcement agencies are handling these automated hacking techniques.

Acknowledgement

I would like to thank the following people, without whom I would not have completed this Capstone project.

I would like to express my sincere gratitude to my mentor, *Mr. Leonard Rogers*, for his extensive support and guidance, with which I was able to complete this capstone project successfully. His reviews and suggestions were a real inspiration for me in completing this project.

I am delighted to thank my program coordinator, *Mr. Shahnawaz Mir*, for his guidance, support, and insights into this capstone project.

I would also like to thank our program director *Prof. Mike MacGregor* for the immense support throughout the project.

Last but not least, my biggest thanks to my family and friends for all the support they have shown me through this incredible journey.

Chapter 1: Introduction

1.1 What is Hacking?

Hacking is nothing but finding an alternative way to use any computer hardware or software to fulfill an objective or to solve a problem. These unintended ways are used because of the ineffectiveness of conventional ways to solve such problems. Sometimes these unconventional ways are categorized as legal and sometimes illegal, based on the motive of such actions.

There are several reasons for a person to get motivated to perform hacking. One of the common ones is to gain legal and authorized access to a system to test the security and efficiency of the system and thereby exposing vulnerabilities, if there are any, and fix them. Some other reasons include gaining unauthorized access to a system and thereby stealing, tampering, or even destroying the contained information or sometimes selling such information to a third party such as a company or a government who usually hired them for profit.

The hackers are generally divided into Black hat hackers and White hat hackers. Black hat hackers are people who perform hacking techniques intentionally for stealing information or money. They are also known as malicious hackers or crackers. Whereas on the other hand, White hat hackers perform hacking techniques in a legal way to find any potential vulnerability in the system. They troubleshoot those security vulnerabilities and research ways to avoid or rectify such problems. Their work on such security vulnerabilities is made public so that others can refer and secure their individual systems. There constant work on such vulnerabilities makes the systems more secure and efficient. White hat hackers are commonly referred to as Ethical hackers.

Apart from Black hat hackers and White hat hackers, there is a third category of hackers named Grey hat hackers. These hackers have qualities of both Black and White hat hackers. That is, they are motivated to perform hacking for both profit and because of ethical reasons. They use both legal and illegal ways to exploit systems. Grey hat hackers are hired both by individual companies and government entities to perform tasks that are related to the company's interest, the country's national security reasons [1].

1.2 Evolution of Hacking

The history of computer technology and its development have always made sure that each new technology is better from the previous one in various terms including in terms of security. This

is mainly because of the role of cyber-attacks and various hacking incidents that affected and challenged the existing technology at that time [2][3].

The first known incident of hacking was reported in 1878 at the Bell Telephone Company, where a telephone operator intentionally redirected or disconnected some of the telephone calls [2][3].

World War II was not only a platform for showing military powers but also motivated many of the countries to boost their information intelligence-gathering technology. 1939 to 1945 witnessed the invention of various hacking methodologies that helped to encode and decode information for intelligence and counterintelligence purposes [2][3].

These operations helped in breaking military codes and ciphers used to transmit top-secret information. During this time, Alan Turing, Gordon Welchman, and Harold Keen developed an electromechanical device that had the ability to decipher German Enigma machine encrypted messages. The world's first programmable electronic computer named Colossus was also developed during this time [2][3].

The 1960s was the decade of numerous development and inventions that created the internet and global computer network what we have seen today. American Standard Code for Information Interchange (ASCII) was created in 1963. Followed by the development of the first programming language named Beginner's All-Purpose Symbolic Instruction Code (BASIC) by Thomas Kurtz and John Kemeny in 1964 [3][4].

The research and development on how to hack a computer or how to stop someone from unauthorized access really came forward after the introduction of The Advanced Research Project Agency (ARPA) Network (ARPANET) in 1969. ARPANET influenced not only many as the first network wide-area packet switching network in the world but also influenced many in terms of hacking and network security. ARPANET opened the way in evolving hacking into what it is now [3][4].

Apart from ARPANET, 1969 was also the year for various other technological developments. Intel developed a new RAM with a boosted capacity of 1kb more than its predecessor. Development of the UNIX operating system started in the Bell labs, and more importantly, after almost 100 years of first reported hacking, the technological world witnessed proper computer hacking, which was done by a few MIT students, who hacked their train sets in-order

to modify its operations and later study the new working model. These studies and research based on this kind of hacking in the existing models led to the development and deployment of new computer systems in MIT [2][3][5].

The students who hacked their train sets were part of the TMRC Signals & Power Subcommittee. They used their knowledge and skills to think outside the conventional methods of the current computer systems, helping them to analyze the efficiency of the system. They shared a serious set of core values like dedication, intensity, and focus while solving problems and curiosity about technical system details [6].

In the 1970s, a new set of security breaches started to occur. These were associated with phones and phone networks. These phone hackers (phrackers) used the faults in the existing phone network to break into both regional and international phone networks to make free calls [7][8].

They were using a whistle that generated the same high-pitch tone (2600 hertz) which AT&T was using to access its switching system. The phone hacker named John Draper built a "Blue box," which was later used in conjunction with the whistle to trick the phone system and make free calls [7].

This technique was later published in a magazine which triggered a large increase in phone hacking [7][8]. This incident was another example of how a hacker can effectively make use of an existing vulnerability in the technology to gain unauthorized access to the system and misuse it [7].

The hacking incidents like the AT&T phone hacking and others continued throughout the 70s and 80s. But, one of the major incidents occurred in 1988 in the name of the "Morris Worm" [7]. The Morris Worm is considered as the first worm to hit the Internet, and worms have been a significant presence ever since [9].

This worm was created by a computer science graduate student from Cornell University named Robert Tappan Morris, hence the name Morris Worm. November 2nd,1988, was the start of a new era of hacking and internet attacks on the world [9].

The Morris Worm was the wake-up call to many security experts, system administrators, and engineers to perform research and development technologies that can defend and analyze such attacks in the future. It was a simple program that was designed to be ubiquitous and unnoticed. What made it so dangerous is its ability to cause widespread damage through its secondary

effects like making the system vulnerable to Denial-of-Service attack, and hence raised many questions regarding the safety of user information on the Internet, especially with the graph connectivity (inter-connectivity) among networked systems [9].

As mentioned above, the Morris Worm was just the start of many more cyber-attacks to come. For example, in 1989, the world witnessed the first cyberespionage case in which hackers from West Germany hacked into the U.S government and corporate computers to steal and selling Operating System (OS) source code details to the Soviet Union's KGB [7].

After 1989, there were numerous reports on various cyber-attacks and espionage around the world, both at the government and private levels. In 2001, one of the early Domain Name System (DNS) attacks were reported in Microsoft's servers, which prevented millions of users from reaching the Microsoft webpages [7].

DNS attacks have become one of the most seen common attacks in the 21st century. DNS has become a pivotal technology that provides an efficient way to perform hostname resolution and facilitate Internet communication. DNS has many backdoors or vulnerable sections that can be used for a targeted attack [10].

Cyber terrorism and cyber espionage were essential topics in the first decade of the 21st century. Many cyber-attacks report specifically targeted government institutions of various countries classified as an act of cyber terror. In 2003- 2005 Chinese hackers carried out a series of attacks against the US cyberinfrastructure named operation "Titan Rain," which focused on acquiring documents related to the latest military and non-military technologies [11]. Many such attacks on various governments worldwide focused not only on data breaches but also on creating accidents and human-made disasters for various geopolitical advantages.

As we analyze these histories of hacking that happened across the world, it is noticed that the intensity and the aftereffect of such hackings are based on the technology available at that time. As new technologies emerge in the market, the path and nature of hacking are changing dramatically. One of the reasons for such cyber-attacks was the unawareness of the importance of cybersecurity in earlier days. This made many systems and users vulnerable even to a small hacking scale that could have avoided simple measures.

As we entered this new era of automation, artificial intelligence, and data analytics, hacking mythology is also modernized. Manual or ordinary hackings and attacks have changed into

automated hackings that can initial attacks on a system/network from remote locations without human interference. Hence, it is vital to understand more about automated hacking in detail.

1.3 Artificial Intelligence (AI)

Before defining Artificial Intelligence (AI), it is very important to understand what it means by Intelligence? It is possible to describe intelligence as the capacity to understand, perform, and adapt to various techniques suitable to the situation to solve problems and achieve goals appropriate to the circumstances in an uncertain and changing world. Intelligence can be cultivated or improved by experience, analyzing previous results, or selecting the desired environment [24].

Artificial Intelligence (AI) is the intelligence that focuses on machines. Unlike natural intelligence like human intelligence, AI is a type of intelligence displayed by machines, especially in the current context the computer systems [25].

In 1956 AI was first introduced by John McCarthy at an academic conference. Before that itself, the concept of machine intelligence and its ability to progress logic was an understudy. In 1950 Alan Turing developed an empirical test of AI, which later came to know as the "Turing Test." The test was to examine an artificial entity is intelligent or not by comparing the answering capability of human being involved in the test [26].

The emergence of AI started with research involving AI modelling the neurons in the human brain. This research was based on the concept of binary variable representation of artificial neuron signals that switched to either on or off (according to programming concept, these signals are represented as 0's and 1's). This further helped develop Hebbian Learning for neural networks in 1949 by Donald Hebb [25].

The first neural network computer was built in 1951 by Marvin Minsky and Dean Edmonds, named Stochastic Neural Analog Reinforcement Calculator (SNARC). In the upcoming years, AI has emerged as a new discipline in the computer and network technology industry whose goal was to create computer systems that could learn, react, and make decisions in a complex, changing environment [25][26].

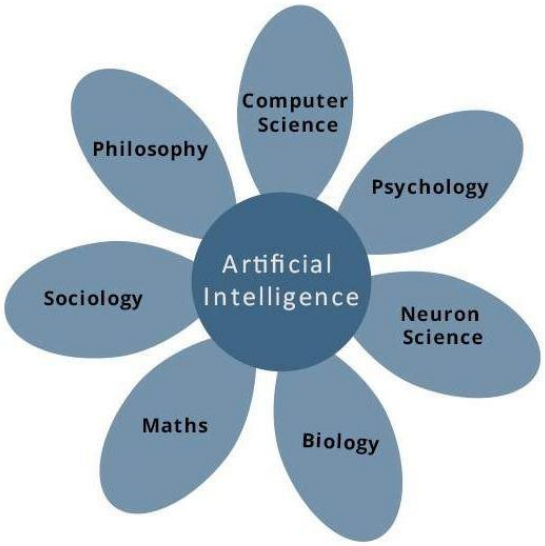

Figure 1.1: Disciplines Contributes to AI [28]

The factors contributing to AI are always debatable and continuously change based on the emergence of new technologies and how these new technologies are inherited into AI techniques. Figure 1.1 presents the science and technology disciplines such as Mathematics, Computer Science, Biology, Engineering, Linguistics, and Psychology that are the most commonly known contributed disciplines for AI. These areas help develop intelligent systems that can power the AI for performing human intelligence-related machine features, such as reasoning, understanding, and problem solving [28].

1.3.1 History of Artificial Intelligence [28]

Year	Milestone / Innovation
1923	Karel Capek's play Rossum's Universal Robots' (RUR) opens in London. Which marked the first English use of the term "robot".
1943	Foundations for neural networks laid.
1945	Isaac Asimov, a Columbia University alumnus, coined the term *Robotics*.
1950	In analysis to define intelligence, Alan Turing implemented the Turing Test and published Computing Machinery and Intelligence. Detailed Study of Chess Playing as a Search was published by Claude Shannon.
1956	The word Artificial Intelligence was invented by John McCarthy. Demonstration of Carnegie Mellon University's first AI program running.
1958	John McCarthy invents LISP programming language for Artificial Intelligence.

1964	The dissertation by Danny Bobrow at MIT showed that machines can comprehend natural language well enough to correctly address algebra word problems.
1965	ELIZA, an interactive problem that conveys on a conversation in English, was developed by Joseph Weizenbaum at MIT.
1969	Stanford Research Institute scientists have developed Shakey, a robot fitted with locomotion, vision, and problem solving.
1973	Freddy, the Popular Scottish Robot, was designed by the Assembly Robotics group at Edinburgh University, capable of using vision to find and build models.
1979	The first autonomous computer-controlled vehicle, the Stanford Cart, was constructed.
1985	Harold Cohen created and demonstrated the drawing program, *Aaron*.
1990	Major advances in all areas of AI − • Important Machine Learning simulations. • Case-based reasoning • Multi-agent planning • Scheduling • Data mining, Web Crawler • Comprehension of natural languages and translation • Vision, Virtual Reality • Games
1997	The Deep Blue Chess Program would defeat Garry Kasparov, the then world chess champion.
2000	Interactive robot pets are being available commercially. MIT features Kismet, a robot with a face that conveys thoughts. The Nomad robot visits Antarctica's isolated areas and locates meteorites.

Table 1.1: History of AI during 20th Century [28]

1.3.2 Symbolic Artificial Intelligence

Symbolic artificial intelligence is the earliest and most common representation of AI. The intelligent system can be explicitly described in Symbolic AI, knowledge is expressed symbolically, and intellectual operations can be identified as formal operations over symbolic expression and structures [27].

The symbolic AI is sub-grouped into two generic models of knowledge representation and intelligent operations, where the AI approach like cognitive stimulation and logic-based reasoning is defined. The second sub-group concentrates on specific applications and is based on representations of domain knowledge. The rule-based representation, structural knowledge representation, and the mathematical linguistics approach of AI are defined in this part of symbolic AI [27].

1.3.3 Computational Intelligence

Computational Intelligence is another group of methods in AI. The common features include numeric information is fundamental in a knowledge representation, and numeric computation is used to process the information. Unlike symbolic AI, the information is not represented straightforwardly. Not all the computational intelligence models include these features; the model like Bayes network model does not fulfill all the characteristics. The models are presented in a generic way to avoid this misunderstanding and grouped under three groups: connectionist models, mathematics-based models, and biology-based models [27].

1.3.4 Weak Artificial Intelligence

Weak Artificial Intelligence, also known as Narrow Artificial Intelligence, is the basic and most common type of AI we use in our computer systems, smartphones, or the internet. It is non-sentient machine intelligence with not-so-complex programming that is instructed to perform narrow tasks with less complexity. These limited tasks include facial recognition, internet searches with given keywords [29].

1.3.5 Artificial General Intelligence (AGI)

Before briefing about AGI, it is important to understand the concept of general intelligence. General intelligence is the ability to accomplish a variety of objectives and initiate a variety of tasks in various environments and platforms. These systems should be able to deal with issues and circumstances that are very different from those predicted by their developers [29][30].

The term "Artificial General Intelligence" has emerged as a synonym for "narrow AI" to relate to systems with the same kind of strong generalization ability. The AGI method views 'general intelligence' as a primarily specific property from the task or problem-specific capability and mainly focuses on knowing this particular property and developing display systems. AGI can also close the gap between the narrow AI and the advanced AI programs, including the complex AI programming and functionality seen in robots [29][30].

1.3.6 Superintelligence

Artificial Superintelligence or just Superintelligence is considered the most advanced type of AI ever developed/developing. Most researchers view it as a futuristic approach to build advanced multi-functional, complex AI. Superintelligence is defined as any form of intelligence that is much greater than current general intelligence and exceeds humans' cognitive performance across all platforms of operations. This is pretty ambiguous now. Under this concept, various systems with very disparate performance attributes may classify as superintelligences [31].

It is beneficial to disaggregate this basic notion of superintelligence by separating various bundles of intellectual super-capabilities to advance the study. There are several ways that such decomposition can be accomplished. Thus, we can classify superintelligence furthermore as speed superintelligence, collective superintelligence, and quality superintelligence [31].

Speed superintelligence is a type of superintelligence that can process information much faster than a human. In comparison, collective superintelligence is another form of superintelligence that achieves superior performance by aggregating large numbers of smaller intelligence. And the third form of superintelligence is the quality superintelligence, which can be defined as a system that can process information as fast as the human mind and vastly qualitatively more intelligent [31].

1.4 Machine Learning

The influence of machine learning technologies has dominated the 20th-century Information Technology infrastructure. It has a promising result that can efficiently integrate big data algorithms with the knowledge-based artificial intelligence techniques in various computer systems platforms. Some many attributes and techniques are combined to represent machine learning technology. But it is important first to understand the proper definition of machine learning and how machine learning has a massive impact on the current computer and network infrastructure.

So, what is machine learning? There are many definitions and explanations that were introduced over the years on machine learning. Machine learning is the area in computer science based on understanding pattern analysis and the principle of machine learning in artificial intelligence. Machine learning deals with developing algorithms and studying algorithms that can understand a particular human and non-human behavior and predict the data generated. These algorithms use the available data to derive predictions and decisions. This is different from standard programming, as standard programming comes to conclusions based on static programming instructions, whereas machine learning algorithms are based on data-driven methodology [32].

In 1959, Arthur Samuel defined machine learning as a "Field of study that gives computers the ability to learn without being explicitly programmed." Tom M. Mitchell provided a widely quoted, more formal definition: "A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P, if its performance at tasks in T, as measured by P, improves with experience E" [32].

Machine learning is generally defined as actions taken in relation to improvements in artificial intelligence-related systems. This task includes information gathering, analysis, recognition, diagnosis, planning, robot control, prediction, decision making, etc. Figure 1.2 represents a simple artificial intelligence system known as a typical AI agent [33].

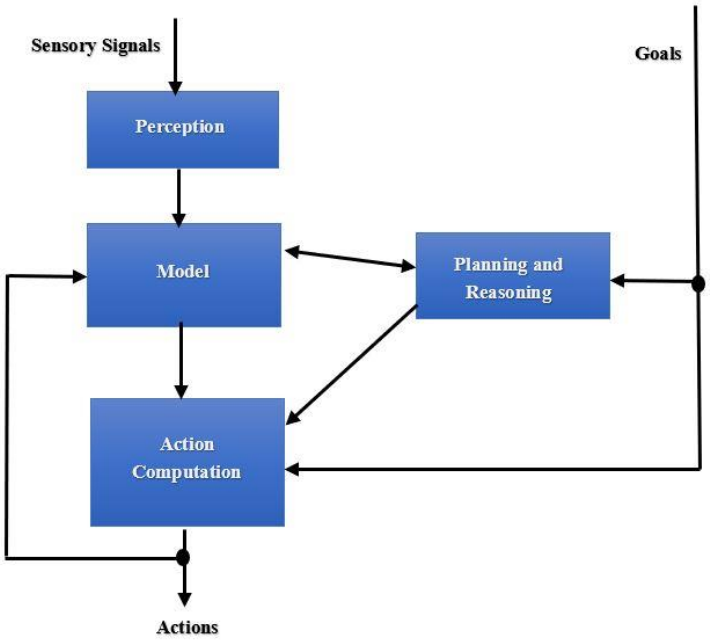

Figure 1.2: AI System [33]

This AI system will build its environment and computes necessary actions based on the results accumulated from each of these components shown in figure 1.2. The decisions are made in anticipation of their effects on the results. The changes made based on the previous results can be counted as learning, and the AI systems use various machine learning mechanisms depending on which subsystem is being changed [33].

1.4.1 Important of machine learning

There are several reasons to understand why machine learning is important. Some of them include:

- The importance of establishing a relationship between input and output data for a greater set of sample inputs, machine learning algorithms can reorganize the internal structure to generate accurate outputs. Thus, the functionality is extended to form a connection among inputs and outputs.
- Machine learning can perform data mining, which will help extract meaningful relationships and correlations hidden in a large pile of data.
- Machine learning methods help to adapt to the changing working environment quickly. This will avoid the inefficiency that can cause a system in a production environment before and after that system is designed and commissioned.
- The volume of information about these tasks can be too high for explicit encoding by humans. Machines that ultimately obtain this information would capture more of it than people would like to markdown.
- Machine learning methods can help the AI systems adapt to the redesign based on new knowledge gained and the development of new technologies associated with the systems [33].

1.4.2 Machine Learning Algorithms

Machine learning is applied in a wide range of fields like computer games, surveillance and security, data mining, virtual personal assistants like Amazon “Alexa,” Google home & personal assistants’ programs like Nest mini, and other areas like share market, search engines, medical diagnosis [35].

Various machine learning algorithms are deployed to different areas of interest-based on the approach it needs to take, the type of input and output data, and the type and complexity of the problem the algorithm is placed to solve. The required machine learning algorithm which need to be chosen from the available techniques of "supervised learning", "unsupervised learning",

"semi-supervised learning," and "reinforcement learning" depends upon the availability of styles and categories of training data. Given below are the most commonly used machine learning algorithms across various fields [35].

- **Decision Tree Learning**

  Decision tree learning participates in a decision tree as a predictive model, which plots an item's expectations to predictions about its target value. A decision tree is a mathematical model that is used for classification. This machine-learning algorithm uses a tree structure technique to divide data into groups and represent the flowchart outcomes. This model classifies knowledge in a dataset by flowing through a request system from the root before reaching the leaf, which implies one class. The root is the grouping feature that plays a primary role, and the leaf determines the class [36][38].

- **Rule-Based Learning**

  Rule-based learning is another algorithm used in machine learning which is almost similar to a decision tree, and hence they can be converted to each other without much complexity. Identifying and utilizing a set of relational rules that collectively describe the information captured by the system is the defining characteristic of a rule-based machine learning environment. Unlike other machine learning algorithms, this is generally defined by a singular model that can be uniformly applied to any example to make a prediction [37][38].

- **Artificial Neural Networks**

  An artificial neural network algorithm is a machine learning algorithm inspired by natural neural networks' configuration and functional characteristics, widely referred to as the 'neural network.' Simulations are structured using a connectionist approach to estimating, processing information of an interactive group of artificial neurons. Modern neural networks are computational data processing tools that are non-linear. Typically, they are used to model dynamic relationships between inputs and outputs, to locate data correlations, or to capture the statistical structure of an undefined joint probability distribution function between variables observed [36][38].

- **Inductive Logic Programming**

  Inductive logic programming is a rule-learning method that uses logic programming as a consistent representation for input cases, context knowledge, and hypotheses. The known history encoding an inductive logic programming scheme can derive a hypothesized logic program that includes both positive and no negative examples, with information and a collection of examples described as a rational database of evidence. Inductive programming is a similar discipline that considers the expression of theories in some form of programming language, such as functional programs [38].

- **Support Vector Machines**

  Support Vector Machines (SVMs) can be related to supervised learning methods used for classification and regression. The algorithm produces a model that predicts whether a new norm falls under one or another group, presenting a collection of training instances, each marked as belonging to one of two classes, an SVM training course [38][39]. It is possible to classify SVM models into four distinct groups, namely:

  a. Classification SVM Type or C-SVM classification
  b. Classification SVM Type 2 or nu-SVM classification
  c. Regression SVM Type 1 or epsilon-SVM regression
  d. Regression SVM Type 2 or nu-SVM regression

- **Clustering**

  The clustering algorithm of machine learning is the analysis of a cluster. It is the allocation of information into subsets so that information within the same cluster is identical according to some pre-designated parameters or criteria. In contrast, observations from different clusters are dissimilar. Other clustering methods make various assumptions about the data structure, often described by a metric of similarity and evaluated, for example, by internal compactness and separation. It's between multiple clusters. Other approaches are based on the approximate density and connectivity of graphs. Clustering is a tool for unsupervised learning and a popular method for evaluating statistical results [38][39].

- **Bayesian Networks**

  A Bayesian network is a stochastic graphic model relating a series of random variables and their conditional independence through a directed acyclic graph, also known as a belief network or directed acyclic graphical model. For instance, the probabilistic interactions between diseases and symptoms may be described by a Bayesian network. The network can measure the probabilities of different illnesses' occurrence and provided the signs [38].

- **Reinforcement Learning**

  To optimize any notion of long-term reward, reinforcement learning is concerned with how an agent belief to take action in an environment. Reinforcement learning algorithms focus on finding a policy that guides the agent's actions to the states of the universe in those states. Reinforcement learning varies from supervised learning in that sub-optimal action neither presents nor specifically corrects good input/output pairs. [38].

- **Representation Learning**

  Several learning algorithms, mostly unsupervised learning algorithms, seek to find better representations of the inputs provided during training. Typical examples include analysis of critical components and cluster analysis. Representation learning algorithms also attempt to conserve the data in their input but transform it in a way that makes it usable, often as a pre-processing step before classification or predictions are carried out. This will enable the information from the uncertain distribution of data generation to be reconstructed, while not always being faithful to configurations that are unplausible under that distribution [38].

- **Genetic Algorithms**

  Genetic algorithms are more stable algorithms and can be used for different problems of optimization. These algorithms do not deviate readily in the face of noise, unlike other machine learning and AI algorithms. Genetic algorithms may be used in large-area or multimodal space searches. They are algorithms that rely on the evolutionary principle of natural and hereditary preference. Genetic algorithms are adaptive heuristic

search algorithms, i.e., the algorithms adopt an iterative pattern that differs over time. It is a form of reinforcement learning where it is important to provide feedback without specifying the correct direction. The input may be either positive or negative [40].

1.4.3 Applications

Machine learning is one of the most discussed and applied technologies in this era. A combination of AI and machine learning algorithms are always made breakthroughs and increased the efficiency of many known computer technologies. Thus, machine learning has numerous applications and fields of interest in a wide variety of platforms. Some of them are listed below:

- **Webpage Ranking**

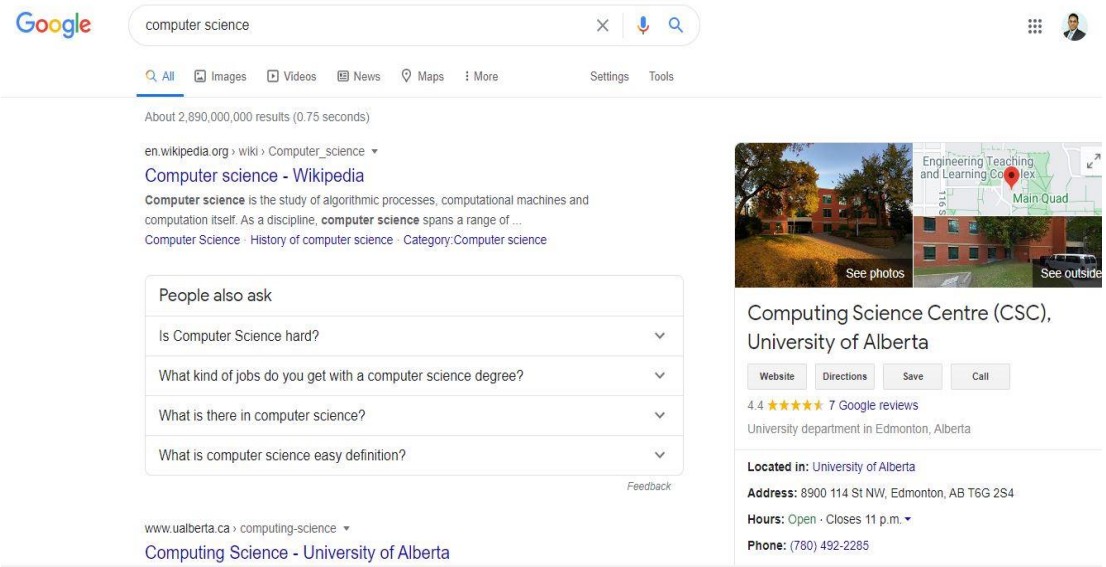

Figure 1.3: Webpage ranking on a search engine.

One of the features most commonly seen on the internet is search engines. A wide range of search engines is available across the internet platform based on the user's requirements. The webpages in a search engine are loaded based on the query submitted on the search engine, which then compares the webpages across various domains in the search engine database with the keyword embedded in the query [34].

The figure 1.3 is an example of such search engine results (in this case, the search engine is Google) for the keyword "computer science." So how the search engines produce these sorted lists of websites and related resources? The search engines use

machine learning algorithms to search through various website structures and analyze these with various sub-factors like the user's geo-location, the frequency of keyword used, the number of times the websites are viewed, etc. After running through different automated machine learning algorithms and tools, the final result is displayed [34].

In figure 1.3, the search engine has also given the University of Alberta's academic building one of the results. This is mainly because of two factors: the user's geo location and the frequency of visits made to the University of Alberta's website and academic buildings (based on the google maps statistical data) [34].

- **Collaborative Filtering**

Another important yet similar to that of webpage ranking is the collaborative filtering of data. This data can be of anyone, of any type, or any level of complexity. Collaborative filtering is the process of filtering a result based on analyzing the users' or computer systems' previous behavior. This can be explained with a comparative example, like the suggestion list the users usually get on their Netflix page, or the suggestion of similar genre books in the Amazon e-commerce website, or the case of Facebook, the friends' suggestions based on mutual friends or with another common factor. These are all because of the collaborative filtering process powered by various machine learning and AL techniques with various complexity levels [34].

- **Facial and Named Entity Recognition**

One of the main applications in which machine learning has proven highly effective is the technique of facial recognition. This machine learning application has a huge role in various security levels in various governmental and non-governmental institutions and individual capacity. Various security applications like access control use this technology [34].

This technology categorizes the faces of an individual and identifies his or her security level of authorization. This is achieved by converting the facial blueprint into complex mathematical values based on the distance between the two eyes, the shape of the face, etc., and process these values in machine learning algorithms to authenticate the facial image. There are various complexity levels in this process, and a relative number of times AI techniques are adapted with machine learning algorithms [34].

Named entity recognition is another application which sometimes used along with facial recognition. Named entity recognition deals with the problem of identifying materials or a particular entity. These identifications are very important for data mining and automatic extraction, and understanding of documents. For example, some latest applications like Apple Mail uses this technique to identify and auto-fill the address of shipping and fill them automatically to the address book. Although systems that use hand-crafted rules can lead to satisfactory results, using examples of marked-up documents to automatically learn those dependencies is much more effective, particularly if we want to deploy the framework in several languages [34].

Figure 1.4 is an example of named entity recognition in which a tool called LingPipe is used to tag a person's name. The relevant information like location, organization, etc., is tagged along for further information extraction [34].

```
HAVANA (Reuters) - The European Union's top development aid official
left Cuba on Sunday convinced that EU diplomatic sanctions against
the communist island should be dropped after Fidel Castro's
retirement, his main aide said.

<TYPE="ORGANIZATION">HAVANA</> (<TYPE="ORGANIZATION">Reuters</>) - The
<TYPE="ORGANIZATION">European Union</>'s top development aid official left
<TYPE="ORGANIZATION">Cuba</> on Sunday convinced that EU diplomatic sanctions
against the communist <TYPE="LOCATION">island</> should be dropped after
<TYPE="PERSON">Fidel Castro</>'s retirement, his main aide said.
```

Figure 1.4: Named entity tagging of a news article [34].

- **Automated Hacking**

Automated hacking is another important machine learning application with a wide range of positive and negative impacts in this world. Adaptation of machine learning and AI technology into computer hacking has created a significant difference in approaching such automated hacking offensively and defensively. Machine learning is used in intrusion detection, random & continuous injections, Distributed Denial of Service (DDoS) attacks. This paper will further discuss these two aspects of offensive and defensive ways in automated hacking and more about machine learning in the coming chapters.

Apart from these applications, given below is the list of more technologies that adopt machine learning as the critical technique of execution and processing.

- Affective computing
- Bioinformatics
- Brain-machine interfaces
- Cheminformatics
- Classifying DNA sequences
- Computational advertising
- Computational finance
- Detecting credit card fraud
- Game playing
- Internet fraud detection
- Machine perception
- Medical diagnosis
- Natural language processing
- Optimization and metaheuristic
- Recommender systems
- Robot locomotion
- Sequence mining
- Software engineering
- Speech and handwriting recognition
- Stock market analysis
- Structural health monitoring
- Syntactic pattern recognition

Chapter 2: Automated Hacking

2.1 Automated Hacking - Then and Now

Automated hacking is a new form of computer hacking that can be carried out directly or remotely with a much higher success rate. Automated hacking is carried out by computer programs that are programmed to analyze the vulnerability of a system(s) and carry out hacking attacks based on the analysis. This method replaces the manual techniques of hacking in which an attacker manually performs the steps in an attack sequence. Manual Methods take time and need much preparation to carry out a successful attack. This includes extensive research on the target, implementing new algorithms, writing down new source codes, or developing new tools to attack a specific target [12].

Automating the different phases of an attack provides a structured overview of vulnerabilities in a system. Automating an attack includes running scripts that can collect information on the target system (reconnaissance), executing programs that can hack into a system without manual intervention [12].

Automated hacking is considered one of the most sophisticated computer programming areas with high integrity in coding and running platform-independent. Programmers have created many such automated attacking tools that concentrate on achieving specific tasks like network exploitation, system analysis, phishing, penetration testing, and surveillance and are flooded in the market as paid and freeware tools [12].

Automated hacking tools provide freedom to its users of not studying the attacking methods and their techniques. Instead, these tools help them to attack or defend a system/network in a short period without any literature review. This enhanced the number of cybersecurity incidences and its response to depend on such automated hacking tools for future references and to develop countermeasures.

Automated hacking tools and scripts are used in different stages of a hacking process. In 2019, Managed Threat Detection and Response (MTR) team, who conducted various cyber threat operations, noticed that the attackers were automating the earlier stages of their attacks to access the target system and then complete the objective of their attack [13].

Once the automated tools led them into a target system, the following main objective is to remain stealth. The remaining stealth will keep their presence unnoticed, and thereby the attackers can analyze the environment and collect valuable intelligence from that. Tools like

Nmap and netstat can be used to move the operations into higher priority targets without being detected [13].

It is easy to notice that the concept of automated hacking has been changed or upgraded over the years. Every year, a new concept in technology and security has been introduced into the world. So, the users who use the automated hacking methodology to defend and attack a network or a system are forced to conduct research on these new concepts and develop new automated programs to achieve their hacking objective. For example, a hacker uses a vulnerability in an organization's network to create and deploy an automated hacking tool to the network. Over time, the engineers will correct this vulnerability with the help of new or existing technology. This will force the hacker to upgrade his hacking concepts to gain access to the network again.

Earlier the automated scripts were introduced into the IT world as part of automated programs and testing. This was followed by internet bots and automated hacking tools powered by strong AI technology and machine learning concepts. A more detailed explanation about these methodologies and technics is included in this chapter.

2.2 Automated Scripting

Automation scripting is a process of automating an existing script or a set of programs that can be executed and managed automatically without having to do custom script development and maintenance going forward. This will drastically reduce the human efforts in writing and managing codes and facilitating the development costs and timelines [14].

An automation script contains various sections, including launch points, variables with a corresponding value in it (values that will be used while the automated scripts are running), and the source code [15]. One of the main things to consider in automated scripting is that most automated scripts are platform-dependent. Most of the automated scripts are developed to run on a particular framework or a framework with similarities.

Scripting always had its challenges in every stage of processing. When it comes to automating an entire enterprise system using scripts, it requires more advanced scripting commands that can target the particular area in the system and use other systems execution commands, including the use of the Command Line interface, to achieve automation. Also, automated scripts require a unique environment to run and execute their tasks. This creates problems in getting adapted to dynamic changes [14]. Some of these challenges were overcome by

deploying AL and machine learning techniques into these automated scripts that triggered dynamic responses based on the targeting path form changes and made the scripting faster and efficient.

Software testing is one of the areas in which the automated scripting had shown significant use. Because of its' integrity and efficiency, automated software testing became so popular in the software development and testing industries. Automation testing reduced the overall effort of testing software and thereby reduced software deployment time. The test scripts used in automation testing consist of a series of commands and events executed based on the logical decisions scripted in the test cases [16].

There are different scripting techniques used in automation testing; some are the Linear scripting technique, Structured and Shared, Data-Driven scripting technique, Keyword-Driven scripting technique, etc. All these scripting techniques are used based on the complexity of the testing and the software's behavior [16]. Table 2.1 provides a detailed comparison of all these scripting techniques

Property	Linear	Structured	Shared	Data-Driven	Keyword-Driven
Ability to use reusable functions	No	No	Yes	Yes	Yes
Data separation from test script	No	No	No	Yes	Yes
Logic steps separation from test script	No	No	No	No	Yes
Access to code required	No	Yes	Yes	Yes	Yes
Use structured programming instructions	No	Yes	Yes	Yes	Yes
Ability to compare test results with expected	No	Yes	Yes	Yes	Yes
Ability to using script in regression testing	No	Yes	Yes	Yes	Yes
Special framework required	No	No	No	No	Yes
Programming skills level	1 (Low)	2	3	4	5 (High)
Effort needed to create test script	1 (Low)	2	3	4	5 (High)
Maintenance costs needed to update test script	5 (High)	4	3	2	1 (Low)
Reusability of test script	1 (Low)	2	3	4	5 (High)

Table 2.1: Comparison Between Different Scripting Techniques [16]

The integrity, efficiency, and security of software and the system in which the software runs depend on how rigorously the software has been tested with various test scripts. A security vulnerability can be discovered in such software testing, or different behavior of the software model can be analyzed and can be corrected based on the test results. Suppose the software testing is not conducted correctly. In such cases, there is a possibility that the security

vulnerability (if any) to be used against the software as a gateway to access the entire system or the network in case.

2.3 Bots and Botnets

Bots are defined as a script or a group of correlated programs developed to repeatedly and automatically operate and achieve specific functions. Most of the time, the bots are activated remotely or otherwise activated automatically based on the function scripted into it while developing. Bots are always used both for legal purposes as well as for illegal (malicious) purposes. Those bots used for legitimate purposes like supporting search engines are called Benevolent bots, and those bots used for unlawful purposes are called Malicious bots [17].

The categorization of bots as benevolent and malicious depends on various factors like the core objective of deploying the bots, who is deploying it, and the impact the bots can cause on the platform being deployed. For example, a law enforcement agency can use a benevolent bot to run its systems and make some functions efficient, i.e., for a defensive purpose. Simultaneously, they can also deploy malicious bots to a target system or network to again information about a person or to again access the system or for surveillance purposes [17].

Botnets always pose a security threat to the digital world in various ways. With the growing sophistication of botnets representing different system intrusion techniques, it’s still challenging for individuals and organizations to defend against it [17].

Over the years, the botnets' capability has increased substantially, and the result was the increase in reports of cyber-attacks caused or initiated by botnets. It has been estimated that the number of systems used in botnets is over in millions [18].

Over the years, the objective of deploying a botnet was changed, and thus the overall architecture of a botnet changes based on the purpose. This always makes it challenging for the security software (e.g., antivirus) and personals to defend and detect botnets' operations in a system/network [18].

In earlier days of the botnet, the objective of deploying botnets was to express “script kiddie” vandalism and demonstration of programming skills over the black hat and white hat community. This slowly changed to profit-based attacks on organizations and extortion that may or may not be backed by organized crimes. This profit-based use of botnets has motivated the development of more sophisticated botnets, as mentioned before [18].

Apart from being categorized as a security problem, bots are also used in various situations that help engineers and analysts find and fix many issues, especially when it comes to network security and software development. The development of software bots powered by explainable Artificial Intelligence or “explainable AI” has made a practical approach in targeted analysis and bug fixing, which further reduced the development and testing time of software and minimized vulnerabilities in a system software [18].

2.4 Information Security Management Automation

As defined in ISO 27001, Information Security Management (ISM) deals with establishing, implementing, operating, monitoring, reviewing, maintaining, and improving an Information Security Management System (ISMS) [19]. ISM provides the process model which can be used to apply automation technology in security systems and networks [20].

Automating a security system starts with planning the risk. Risk management of automating an information security system can be demonstrated with the help of security ontology. Figure 2.1 represents the security ontology concepts and their relationship to each section in an information security system.

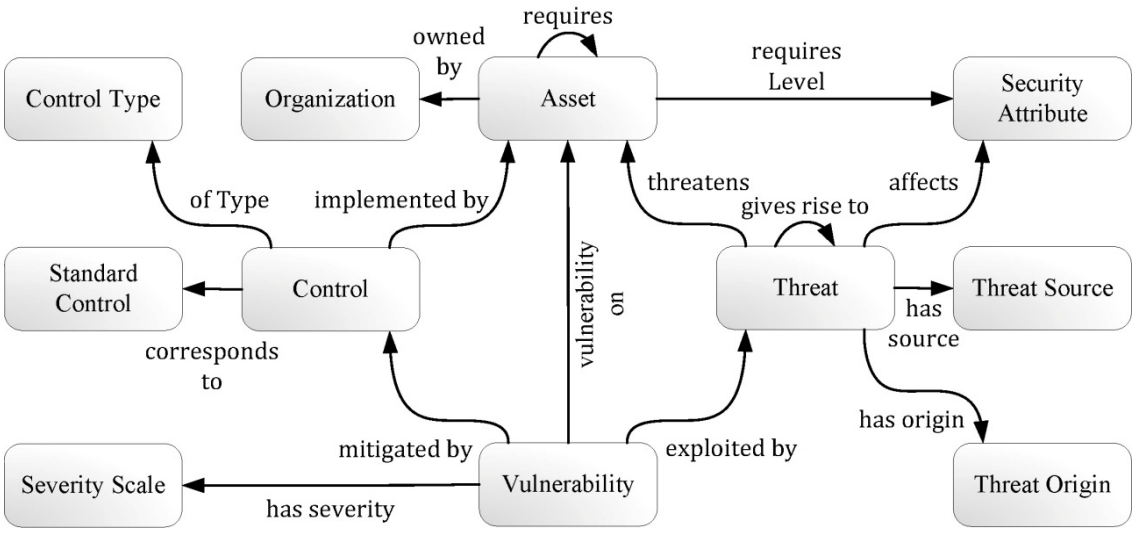

Figure 2.1: Security ontology - main concepts and relationships [20]

The Security ontology [4] is based on the security relationship model presented in the special publication 800-12 of the National Institute of Standards and Technology. The above figure (figure 2.1) clearly projects a security ontology in which threats, vulnerabilities, controls, and implementations are pivotal elements [21].

A vulnerability can be led to the creation of another vulnerability or a threat, which further compromises a system's security in many ways. The security ontology also differential the type of vulnerabilities as physical, technical, or administrative. The origin of these threats (human or natural source) also plays a crucial role in finding a solution.

Once a vulnerability is captured, it is important to take immediate corrective actions so as to protect the respective assets by preventive, curative, deterrent, recovery, or detective measures (control types). Each control type is derived based on the international security standards, and thus it further helps in mobility and reusability. The coded ontology follows the OWL-DL (W3CWeb Ontology Language) standard and ensures that the knowledge is characterized in a standardized and formal form to enable it's employed by automated systems [21].

By having a standard definition of the domain and the device security platform, the security ontology enables interoperability. A formalization of common knowledge that facilitates computer processability and allows reusing knowledge already gathered by intelligence collection offers additional security ontology (enables reusability due to adaptation of international security standards). These all together support the risk management automation of an information security system [20].

Domain	**Information Security Controls**			
	Controls that can be automated	*Total controls*	*Percent*	*Examples of controls*
Security policy	0	2	0.0%	-
Organization of information security	0	11	0.0%	-
Asset management	1	5	20.0%	Inventory of assets
Human resources security	1	9	11.1%	Removal of access rights
Physical and environmental security	2	13	15.4%	Physical entry controls
Communications and operations management	15	32	46.9%	Controls against malicious code
				Information back-up
				Audit logging
Access control	13	25	52.0%	Unattended user equipment
				Network connection control
Information systems acquisition, development and maintenance	4	16	25.0%	Key management
				Control of technical vulnerabilities
Information security incident management	0	5	0.0%	-
Business continuity management	0	5	0.0%	-
Compliance	1	10	10.0%	Technical compliance checking

Table 2.2: ISO 27001 Controls That Can Be Automated [20]

When it comes to automating the security controls, the ISO 27001 has clearly mentioned the automation possibility and how it can affect security system management. Table 2.2 shows the various controls that can be automated based on the ISO 27001 standards and provides examples for these controls. If the control functions can be performed without human involvement, a security check can be automated. Some controls can be completely automated, as machine-readable and processable resources are needed for operations and control monitoring. Simultaneously, some controls can only be automated partially (or to a certain level) because the control system still requires human resources to process [20].

2.5 Security Content Automation Protocol

Security Content Automation Protocol (SCAP) is developed by the National Institute of Standards and Technology to overcome the deficiencies like standardization and interoperability across security tools and reduce the security administration costs on covering the vulnerabilities caused due to these. SCAP is a suite specification that standardizes the format and nomenclature by which the security software tools transfer data related to the software identification, flaws, and security configurations. SCAP supports vulnerability checking using automated technology, control compliance activities, and security measurements related to that [22].

SCAP framework allows automation of managing security checklists, examine various vulnerabilities in different security platforms, and report it instantly [22][23]. This framework replaces the traditional manual way of witting the security checklists, and the actions need to be taken based on the audited results of those checklists [22].

2.5.1 SCAP Technical Specifications

Category	Specification	Definition
Language	Extensible Configuration Checklist Description Format (XCC DF)	An XML specification for structured collections of security configuration rules used by an operating system (OS) and application platforms.
	Open Vulnerability and Assessment Language (OVAL)	An XML specification for exchanging technical details on how to check systems for security-related software flaws, configuration issues, and patches.
Enumeration	Common Platform Enumeration (CPE)	A naming convention for hardware, OS, and application products.

	Common Configuration Enumeration (CCE)	A dictionary of names for software security configuration issues, such as access-control and password-policy settings.
Vulnerability	Common Vulnerabilities and Exposures (CVE)	A dictionary of names for publicly known security-related software flaws.
	Common Vulnerability Scoring System (CVSS)	A method for classifying characteristics of software flaws and assigning severity scores based on these characteristics.

Table 2.3: Current Security Content Automation Protocol (SCAP) specifications [22]

SCAP technical specifications are categorized into three sections, as seen in Table 2.3. They are

a) Languages for specifying the security checklists, defining the testing procedures for these checklists.
b) Enumerations for security and product-related information.
c) Vulnerability analysis and scoring system for checking the vulnerability's behavior and generate scores based on those analyses [22].

2.5.2 Limitations of SCAP

It is clear that SCAP deals with different categories of threats, which can mitigate other threats even if only one solution is employed. The main problem with the SCAP solution is that they are always deployed directly into the system. The SCAP tools should be run locally or through privileged access, which is not suitable for cloud systems. This will compromise the SCAP solutions as the user cannot rely on the tools or the results developed

Chapter 3: Automated Hacking for Defensive Purposes

3.1 AI and Machine learning for defence:

Automated hacking using AI and Machine Learning techniques has a huge role in defending a system from various cyber-attacks. Over time, multiple studies showed that recovering from a cyber-attack is very expensive than establishing a countermeasure to avoid such attacks. These cyber-attacks not only affect an organization financially, but they will affect their reputation too. Also, sometimes these organizations have to pay the government's penalty for failing to counter such cyber-attacks. These factors pushed various industries and organizations to invest in building tools and systems to defend against automated hacking.

Various organizations develop automated hacking tools to test it on their systems and analyze how vulnerable the system is and the drawbacks of their active security measures. These tools provide a digital footprint to the user on what he needs to protect, the scale of the digital attack the system can handle, and the vulnerabilities present. These assessments with automated hacking tools will not provide 100% protection from cyber-attacks but rather offers proactive detection and mitigation of the system's risks.

Over the years, deep research and development boosted machine learning and AI in taking part in computer engineering and networking technology. One such advancement is the role of machine learning and AI techniques in making automated hacking efficient. Machine learning and AI techniques can analyze a system by faster calculations and running millions of permutations. AI uses complex mathematical progressions with too many If Else statements integrated with some exceptions and handlings. The concept is that the Artificial Neural Network does the handling and the genetic algorithms do the learning part.

Machine Learning and AI techniques include various simple and complex algorithms widely used to perform various tasks, including stimulating attacks, performing various reconnaissance steps, etc. These tasks help in attacking a system effectively and potentially helping analysts check, defend, and secure vulnerability with no physical brainstorming required.

Machine learning and AI techniques help from simple malware analysis to the most complex vulnerability analysis. They are being used to improve protection, perform repetitive analysis scans, and generate log reports. On the other hand, these techniques can be programmed to perform fast cyber-attacks, replicate, make them undetectable, perform deep data manipulations and destruction, and launch attacks on multiple sources.

3.2 Automated Cyber Threat Intelligence

As discussed in previous sections, the number of cyber threats is increasing at an alarming rate. New models of attacks and vulnerability exploitations are causing huge problems and financial burdens across the globe. It is crucial to have always an upper hand in dealing with these cyber threats to avoid and defend these attacks from happening and reduce the collateral damage in case of a successful attack. This is where an automated cyber threat intelligence gathering has made significant importance in the white hat community [41].

Cyber threat intelligence refers to “the set of data collected, assessed and applied regarding security threats, threat actors, exploits, malware, vulnerabilities, and compromises indicators.” Cyber threat intelligence has helped security professionals understand cyber threat signs, extract information about the attack methods, and react correctly and on time to the attack [41].

So, the question is; “How can this related information be collected, and how can we process such a large amount of information to extract all the vital intelligence?” This is where the AI and machine learning techniques are applied with the efficiency of automating the overall process of cyber threat intelligence.

To understand, analyze, learn, and operate intelligently against sophisticated cyber threats, cyber threat intelligence's evolving field considers the implementation of artificial intelligence and machine learning techniques. Researchers have taken various artificial intelligence approaches into account in recent years to provide cybersecurity personnel to identify signs of cyber-attacks. In particular, related to their demonstrated productivity in malware analysis (in both static and dynamic) and network anomaly detection, thus creating a growing trend in the use of machine learning (ML) and data mining techniques [41].

The automated cyber threat intelligence is achieved through five necessary and correlated steps. Figure 3.1 represents five steps of automated cyber threat intelligence and shows how they are correlated with automation.

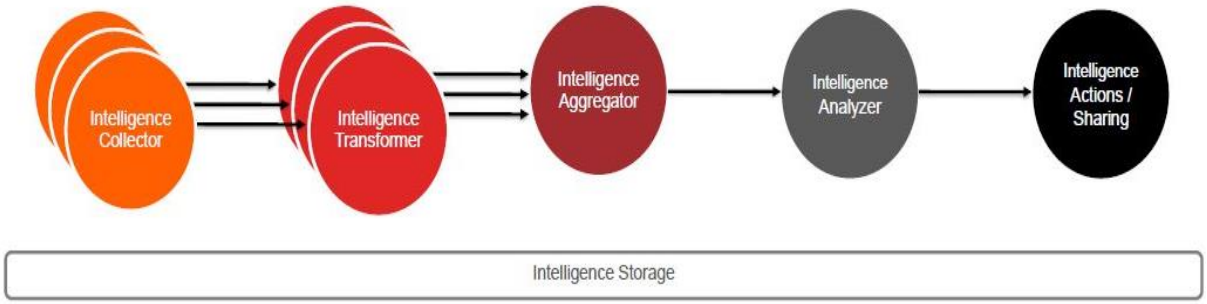

Figure 3.1: Five steps of automated cyber threat intelligence [42]

3.2.1 Automated Intelligence Collection

Intelligence collection is the initial and vital stage of cyber threat intelligence. The intelligence data can be collected from any source with different levels of complexity. Basically, these sources are categorized into two:

- External Intelligence Sources

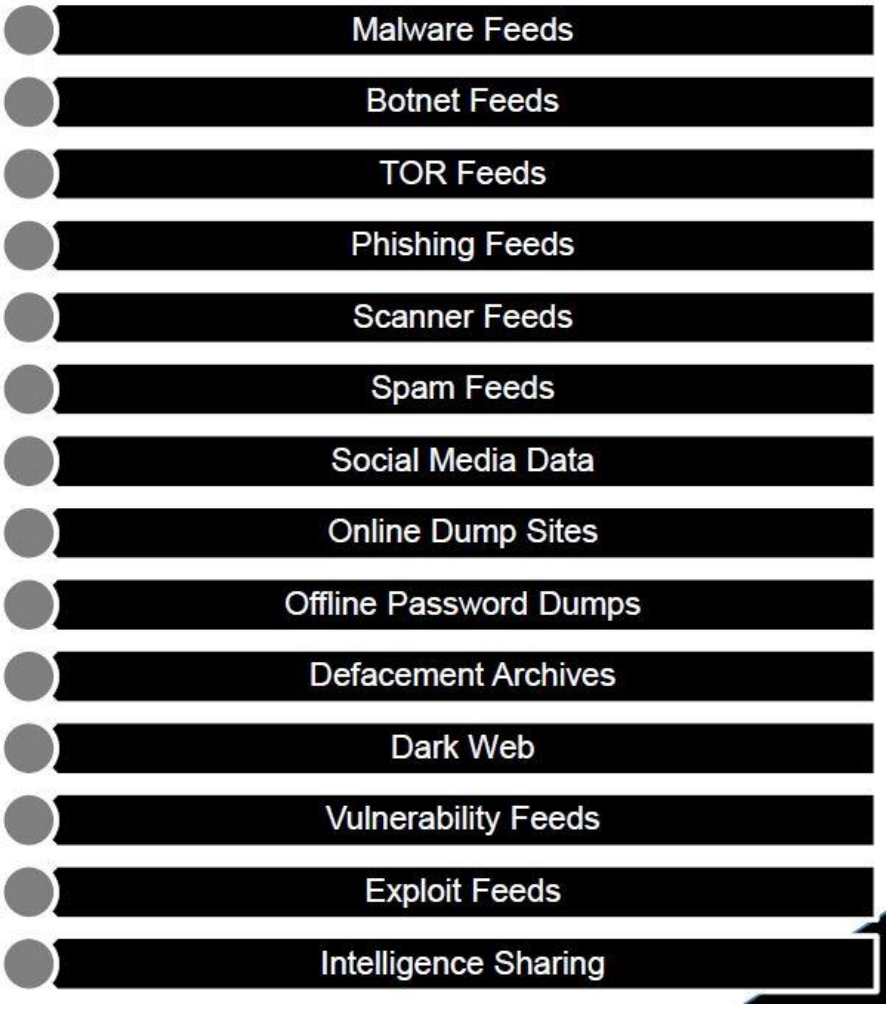

Figure 3.2: External intelligence sources [42]

External intelligence sources are always vulnerable to tampering, and the successful extraction of vital intelligence from these sources is time taking procedure. Figure 3.2 represents the list of external sources that can be relied on to prevent attacks, detect and neutralize security breaches, identifying the vulnerable systems and employees, etc. [42]. The data extracted from these sources can be processed using simple machine learning algorithms like decision tree algorithms (refer to section 1.4.2) and deploying advanced state-of-the-art AI-powered data collection tools like PRISM. Even though, with the potential to data mining this intelligence, the integrity of these data is always questionable because of the sources the intelligence is extracted.

- Internal Intelligence Sources

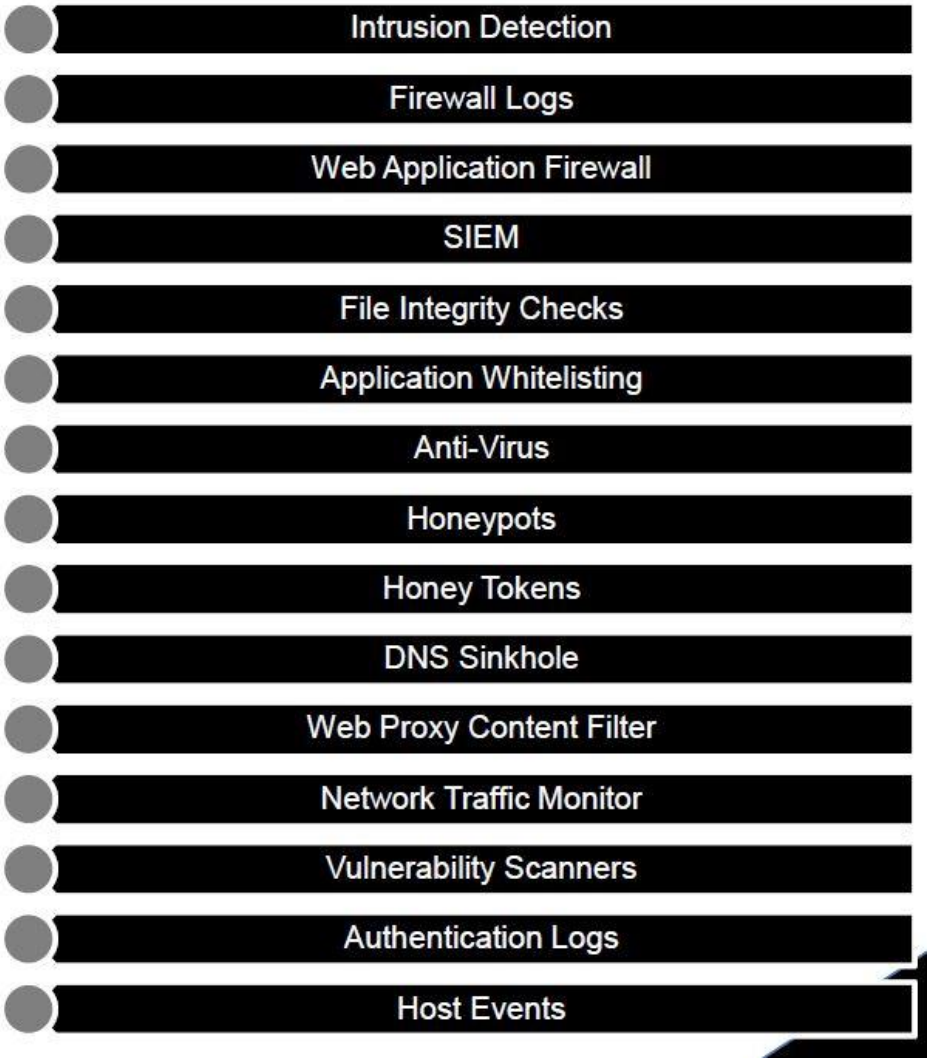

Figure 3.3: Internal intelligence sources [42]

Internal intelligence sources consist of areas located inside the designated organization or a company, or even a system. These data can be used to identify security

vulnerabilities, provide context to threat activity, detect a wide variety of security incidents like malicious network traffic, anomalous credential usage, abnormal traffic flow from the system, etc. Like the external intelligence sources, with the help of internal intelligence sources, we can detect the vulnerable devices and employees and generate internal intelligence feeds and industry-based threat data [42]. Figure 3.3 provides the list of internal intelligence sources that can be used to extract vital information.

3.2.2 Automated Intelligence Transformation

Intelligence transformation can be defined as the process of extracting valuable information from a set of data collected from various intelligence collection sources. This process is further expanded by deploying automated scripts and automation tools to increase efficiency. Automated intelligence transformation is powered by data mining, machine learning algorithms, and AI techniques [43.].

Machine learning facilitates automatic and continuous learning by tracking, evaluating, and eventually presenting potential measures through data analysis to alter or boost performance. Its potential to detect patterns or deviations in big data provides much-anticipated processing power while enabled with AI techniques. Three main elements are comprised of intelligence transformation are [43].

- **Big Data**

  Big data is the engine behind the intelligence transformation because trends, interactions, and observations are buried inside the increasing bounty of organized and unstructured data: the more information, the better. And the more detailed, the more granular, the more comprehensive the data (with robust metadata), and the less latency, the better.

- **Advanced Analytics**

  Advanced Analytics incorporates advanced computational methods (statistics, data analysis, predictive analytics, machine learning, big data, reinforcement learning) to pull out or discover business, product, and organizational knowledge from this increasing volume of data.

- **Intelligent Applications**

  Applications are the medium for offering actionable insights (think prescriptive and preventive recommendations) and documenting as they communicate with the program the subsequent experiences and behavior of humans and machines.

3.2.3 Automated Intelligence Aggregation

A centralized database guarantees the capacity to respond and align quickly promptly against the current cyber threats. In a common point of fact, aggregating threat analysis keeps the most precise and up-to-date record of observations and enrichments from multiple intelligence outlets, including cooperation between defense teams and enrichments from internal networks [44].

Automated intelligence aggregation consists of aggregating data from external and internal sources. Aggregating of internal intelligence deals with collecting selective packet data, incident-response reports, internally derived intelligence reports. Whereas external intelligence aggregation deals with ingesting from multiple sources of data like feeds of indicators, structured data with context, etc. [44].

The automated intelligence aggregation platform should also query other predictor servers and credibility folders, such as blacklists, VirusTotal, etc. It should obtain information on the enrichment of markers, such as IPGEO (geographical data).

An automated intelligence aggregation platform requires external pivoting to be analyst-driven and automatic. By comparing it to other vulnerability operations, pivoting is a way to bring knowledge into perspective. External pivoting aims at external information such as pDNS, repositories of malware, and domain intelligence [44].

3.2.4 Automated Intelligence Analysis

Once the collected intelligence is aggregated, before an appropriate action plan can be created, it needs to be processed and put into perspective. This is when it comes to automated intelligence analysis. The data has no meaning without analysis. So, it is important to process this available intelligence in the right way so as to make it useful in the future [44].

The aggregated data may be manually or automatically analyzed. For successful analysis, human analysis using a validated technique such as the Diamond Methodology is important.

The pitfall with manual research is that it takes so many hours for people to overlook important connections. Therefore, wherever possible, analysis needs to be automated. Automated processing achieves findings more efficiently and thus in more significant amounts. The approach is versatile and gives a higher degree of scientific detail [44].

3.2.5 Automated Intelligence Actions & Sharing

The final step of cyber threat intelligence is the sharing or transfer of processed intelligence into relevant areas. As well as taking security decisions based on the intelligence analyzed. These security actions should focus on preventing threats to the system or containing already occurred breaches.

As mentioned before, the cumulative collection of intelligence from various sources helps in taking real actions and thus facilitates detailed metrics on information management of high thrust threat-related incidents. The distribution, either for internal use or external sharing of merged threat information, threat evaluation, or other reports, enables an entity to engage in more extensive security culture, thus improving its protections against adversaries [44].

3.3 Orchestrated Security Infrastructure

Cybersecurity threats have significant corporate and socio-economic implications, such as loss of sales, credibility, information technology disruption, theft of confidential data, and customer-sensitive information. To prevent known and unknown attacks and mitigate the repercussions typically associated with technical flaws and risks, companies use multiple security solutions [45].

Providers of security solutions use multiple technology and paradigms to design, deliver and run their security solutions that are not readily compatible and interoperable to serve the Security Operation Centre (SOC) effectively and efficiently [45].

Orchestrated security infrastructure aims to implement technological and socio-technical strategies to incorporate multi-vendor security resources as a single whole to support SOC security personnel. Organizations are rapidly implementing responsive, autonomous, and collaborative security orchestration platforms to allow security professionals to reliably and efficiently execute their responsibilities [45].

A security orchestration initiative encourages individuals, practices, and technology to collaborate to strengthen enterprises' security knowledge for improved control and security

operations. Security orchestration is a requirement for security automation, which uses information technology, automated machine learning algorithms, and artificial intelligence to track, deter, and recover automatically from cyber threats without human intervention [45].

3.3.1 What is Orchestration [46]?

According is Red Hat Inc, “Orchestration is the automated configuration, management, and coordination of computer systems, applications, and services. Orchestration helps IT to more easily manage complex tasks and workflows.” [46].

Orchestration can be defined as a system administrative methodology that can perform automated configuration, coordination, and management of a computer and software system. A centrally controlled collection of workflow logic makes interoperability between two or more separate programs more straightforward in these systems. A typical orchestration implementation is a hub-and-spoke model that allows a central orchestration engine to communicate with several external participants [47].

One of the guiding criteria behind these technologies' production was to handle the convergence of broad business processes. Through orchestration, it is possible to link various processes without redeveloping the solutions that originally automated the processes. By adding fresh workflow logic, orchestration fills this void. Also, the use of orchestration will greatly minimize the difficulty of environments with solutions. The workflow's logic is abstracted and retained more effectively than when incorporated within elements of individual solutions [47].

3.3.2 Security orchestration vs. Security automation

Security automation can be defined as the ability to perform single/multiple security operations independently without intervention from a human being. In contrast, security orchestration is the ability of a system to perform various automation operations across various platforms. This suggests that automation activities are part of the overall orchestration process, which encompasses broader, more complicated situations and assignments [48].

With this being said, we may assume that orchestration involves the automated integration and control of processes, middleware, and resources. Security orchestration integrates various automated and semi-automated procedures to execute a complex process or workflow automatically, and these may consist of various automated tasks or programs [48].

By distinguishing between a single task and an entire process, automation and orchestration can best be understood. Only a single task is handled by automation, while orchestration uses a more complex set of functions and processes. It accelerates things when a task is automated, particularly when it comes to repeating basic tasks. But with simple automation, optimizing an approach is impossible, as it only handles a single task. A process is not restricted to a single function, so only with orchestration is optimization possible. If done right, orchestration achieves the primary objective of speeding up the entire process from beginning to end [48].

Optimizing a process is the main aim of orchestration. Although security automation is limited to a single approach being automated, orchestration goes well beyond this. On the other hand, with automation supplying the requisite speed processes, orchestration offers a simplified solution and process optimization [48].

3.3.3 DNS Sinkhole

Domain Name Service (DNS) is the directory that provides access to various websites. DNS is considered as one of the foundations of the internet. DNS service is provided worldwide using DNS servers worldwide with different configurations and security settings complexity levels. Because of its heavy traffic flow, DNS is always prone to various attacks.

With high-volume messages from DNS resolver servers, DNS reflection attacks will swamp victims. Attackers use the victim's spoofed IP address to request huge DNS files from all the open DNS solvers they can identify to do so. The victim experiences a stream of unrequested DNS data as the resolvers respond, which overwhelms their computers.

DNS spoofing is achieved by initiating a DNS sinkhole attack or otherwise known as a DNS blackhole attack. This prevents specific targeted URLs from getting resolved in the network. This can be done by configuring the DNS forwarder to return a particular URL with a fake IP address. DNS sinkhole should be used to avoid enterprise-level links to malicious URLs. By inserting a false entry to the DNS, malicious URLs may be blocked, and there would be a second level of security. Firewalls and proxies are normally used in the enterprise to block malicious traffic [49].

DNS sinkhole is considered one of the powerful methods to avoid Botnet (refer to section 2.3) attacks across the internet. To prevent any such attacks, we reply to extracting malicious or potentially vulnerable URLs from the internet and applying them to the DNS sinkhole system. With the new domain blacklist used by the DNS sinkhole scheme, we will be able to identify

and prevent the most malicious activities on the Internet early on [50]. Figure 3.4 shows the DNS sinkhole system and how it works.

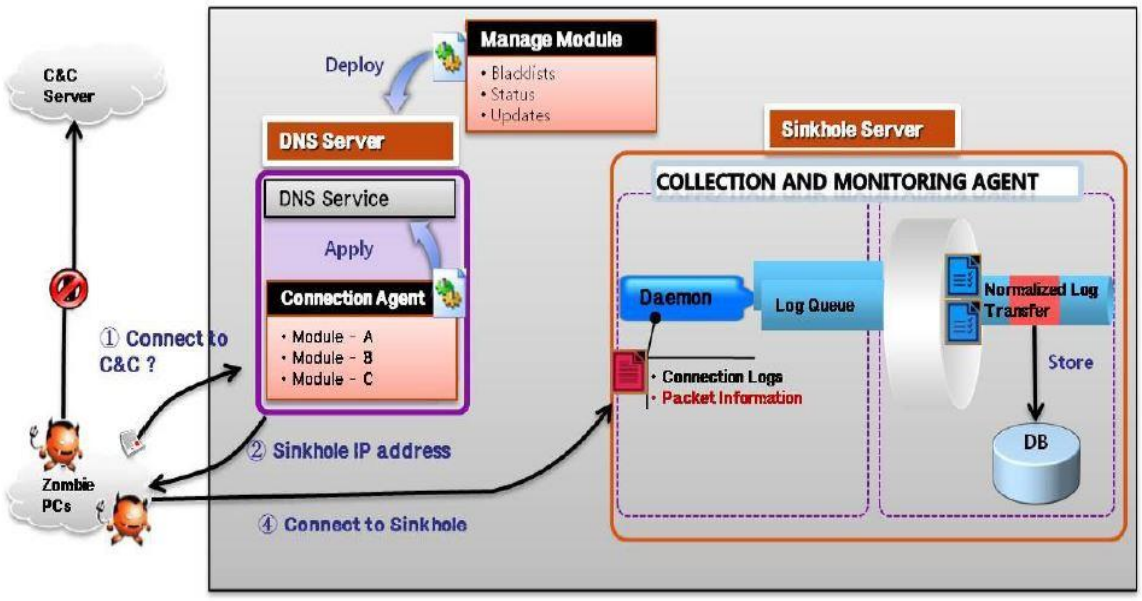

Figure 3.4: DNS Sinkhole System [50]

The main objective of the DNS sinkhole system is to block the chain of communication between the command and control (C&C) servers and so-called “Zombie PCs” (computers that are connected to the internet and are compromised by botnets). If zombie PCs send a DNS query to a DNS server without a DNS sinkhole server, they can obtain the C&C server's IP address. This means that they will connect to the C&C server and will be supplied with malicious commands. But, when the DNS sinkhole system server runs, the DNS server will return the sinkhole server's IP address, not the C&C servers. Thus, since zombie PCs link to the sinkhole server, we will prevent the transmission of malicious commands to zombie PCs [50].

3.3.4 Syslog Collector

Syslog collector is a system tool that can record system log data. It can be operated as a central log collection tool for the entire security infrastructure, thus provides a feather of evidence collection and thrust protection. Syslog collectors are used as a storage option for a long period. It helps in automated intelligence-integration for system log analysis and reporting of automated breach detection. Thus, act as a booster for machine learning algorithms to detect and trigger security incident responses from the syslog collector's logs.

Most of the Syslog collectors consist of mainly three core components [51]:

- **Syslog Listener**: Takes data sent to UDP port 514 from Syslog messages and stores it in a consolidated area.
- **Syslog Internal Database**: Syslog data is maintained by the strongest Syslog servers inside an internal database.
- **Syslog Filtering System**: The amount of time taken to sift through all this information is reduced by filtering log messages.

Some of the Syslog collectors available are:

- SolarWinds Kiwi Syslog
- Paessler PRTG Syslog
- SnmpSoft Syslog Watcher
- Splunk Light
- Fastvue Syslog

3.4 Automated Incident Response (AIR)

Incident response is a term used to describe the method by which an entity responds with a data breach or cyber breach and how the organization tries to handle the results of the attack or violation (the 'incident'). Ultimately, the aim is to handle the event efficiently so that the damage is minimized. All recovery time and expenditures are kept at a minimum, and collateral damage such as brand image [52].

Organizations should have a specific emergency management plan in place, at the very least. This strategy should identify an organization's occurrence and include a simple, guided structure to be implemented when an incident happens. Besides, it is advisable to identify the departments, personnel, or representatives responsible for overseeing both the overall effort for incident response and those responsible for taking each action stated in the incident response plan [52].

3.4.1 Importance of automation

Automating the response to security incidents helps the organization’s security operations staff triage alerts more quickly, adapt quicker to critical incidents, and incorporate your current defense technologies seamlessly into a more effective and robust incident management

program. To recognize and respond to security threats and events, automation speeds up standard responses and routine activities; thus, no human interaction is required. The goal of automation in incident management is to help organizations develop around-the-clock protection systems [53].

The effect of automatic incident response can mostly be felt in the real-time identification and response of attacks. Ninety-one percent of cyber-attacks, for example, begin with a phishing email and an automatic incident response in place. However, without any human interference, these warnings and risks can be addressed efficiently. Automation removes researchers' need to regularly dig through hundreds of alerts, from obtaining ransomware intel to following set processes and remedying risks [53].

According to the SANS Institute [54], the most automatic processes are to remotely execute custom content or signatures from security suppliers and block command and control on malicious IP addresses, followed by deleting rogue files. Techniques that are least likely to be automated include isolating contaminated computers during remediation from a network, locking down networks, and taking them offline [53][54]. The effective incident response, according to the SANS Institute, is six. These six steps are discussed below as sub-topics of AIR.

3.4.2 AIR Preparation

Preparing for an unavoidable security breach is the most critical step of incident response. Preparation allows organizations to evaluate how best their Computer Incident Response Team (CIRT) will respond to an incident, including policy, response plan/strategy, communication, documentation, CIRT team, determination, access control, tools, and training [54].

The policy includes a written collection of guidelines, laws, or practices; it is one of the main components that guide when an event has occurred in an organization. The policy is followed by a response plan and strategy that takes care of circumstances and prioritizes them based on their impact on the organization [54].

Since it will be essential to reach particular people during an incident, having a communication plan is crucial. When it is necessary to contact them, and why, the whole CIRT should know who to contact. The most important justification for reporting an incident is that it may be used as evidence to put the suspect(s) to trial if the incident is viewed as a criminal offense. The other explanation that is just as important for recording is for lessons learned [54].

3.4.3 AIR Identification

This stage deals with the diagnosis and assessment of whether an incident is a deviation from regular activities within an organization, and its scope implies that the deviation is actually an incident. This basic phase allows one to obtain incidents from multiple sources, such as log files, error messages, and other tools, such as firewalls and intrusion detection systems, that can provide information to decide whether an incident is an incident. Suppose an occurrence is confirmed to be a specific case. In that case, it can be reported as quickly as possible to allow ample time for the CIRT to gather information and prepare for the previous measures [54].

CIRT representatives should be alerted at this point of an event, and contact between members and authorized command center personnel should be organized. To handle an incident, it is recommended that at least two incident handlers are present so that one can be the primary handler who can identify and examine the incident and the other to aid in collecting information. It is important to connect and coordinate between CIRT participants, especially if the event's complexity can substantially affect business operations [54].

As mentioned above, these records should be able to address the who, when, when, when, and how questions whether the evidence was to be used to prosecute the perpetrator(s) in court. This is also the stage in which emergency victims should record everything they are doing. The CIRT team will continue with the next step after assessing the incident's extent and reporting the facts [54].

3.4.4 AIR Containment

Once an incident is detected or established, it is a top priority to contain it. The primary focus in an AIR containment is to control and mitigate the damage and prevent the possibility of further damages in the future. This is because the faster the containment is controlled; the fewer damages can cause. It is important to remember that all of the suggested steps of SANS should be taken during the containment process, especially to "prevent the destruction of any evidence that may be required for prosecution later." These steps involve short-term containment, back-up of the device, and long-term containment [54].

3.4.5 AIR Eradication

The actual removal and reconstruction of affected systems are discussed in this process. To assess the expense of human hours and other services to evaluate the organization's total effect,

continued reporting of all steps taken would be critical, as with any of the previous phases of incident management. It is also essential to ensure that sufficient measures have been taken to remove the infected devices from harmful and other illegal material to ensure that they are thoroughly safe [54].

In general, this will involve a complete re-imaging of the hard drive(s) of a system to ensure that all malicious material has been deleted and that reinfection is avoided. This process is also performed when, after learning what caused the incident, protections can be strengthened to ensure that the system will not be breached again [54].

3.4.6 AIR Recovery

The key activities involved with this phase in incident response are inspecting, tracking, and validating systems before bringing them back into production to ensure that they are not re-infected or damaged. To resume processes, validate and check corrupted systems, track for suspicious behaviors, and use resources for checking, tracking, and validating system behavior, this process also involves decision-making in terms of time and date [54].

3.4.7 AIR Lessons Learned

After all the others, the essential step is Lessons Learned. The goal of this stage is to complete any paperwork that has not been completed since the incident and any new documentation that might help future accidents. To have a play-by-play analysis of the whole event, the text should also be published in the form of a paper; the report should be capable of addressing the questions that can emerge during the lessons learned meeting: Who, When, What, Why, and How [54].

The main aim is to learn about the events that happened within an organization in the case of a similar occurrence and enhance the staff's efficiency and provide reference resources. The papers may also be used as instructional manuals for new team members or as a benchmark for future reference emergencies [54].

The lessons learned should be conducted as quickly as possible; within two weeks of the incident, a reasonable rule of thumb is. In an executive summary style, the conference should run over the incident response report with finalization. It should be kept short not to lose the crowd's interest and stay professional [54].

3.5 Automated Penetration Testing

When it comes to protecting a system or a network with automated hacking fundamentals, automated penetration testing and its advanced features play a critical role. So, what is penetration testing? Is it that much important to have a penetration testing capability in cybersecurity?

The U.S. Department of the interior defines penetration testing or pen testing as “a controlled attack simulation that helps identify susceptibility to an application, network, and operating system breaches.” This will help in implementing defensive strategies to defend critical systems and intelligence [55].

Penetration testing can be explained as the practice of testing applications by qualified security professionals for their security vulnerabilities (e.g., penetration tests or ethical hackers). Such a test aims to improve the security bugs that might be present in the app so that the hacking community cannot take advantage of them quickly. During Web App pen testing, the program being tested is a web application stored on a remote server that clients can access over the Internet. Pen testing is mostly performed on web applications as they are more prone to automated attacks [56].

Companies that conduct penetration testing can be divided into three distinct penetration testers: grey hat, black hat, and white hat. In the white hat, the tester is an ethical hacker who follows the organization's guidelines and uses the penetration tests for research and development purposes. While the black hat is primarily used to figure out how the employees of a particular organization deal with the undesired attack [57].

Here only the organization's administrative body are the only people who know the evaluation is ongoing with this strategy. Also, we can make a Gray hat into a custom test plan that is a hybrid solution to the previous two types [57].

Penetration testing is an important process that needs to be performed in a security evaluation of a system or an application. Hence, most organizations perform this testing during its production release or when a significant product upgrade has been carried away. However, it is essential to conduct penetration testing while installing new software to the system, or after modifying policy details, new security patches, or after adding new infrastructure to the overall system network topology [57].

Penetration testing has increasingly been used to define the flaws that occur in the device to learn how to prevent them. Typically, the test simulates different kinds of attacks on the target machine. The administrator would have a coordinated and regulated way to recognize the security vulnerabilities through this exercise [57].

The money and time taken for rigorous testing would make the expense of penetration testing intensive. As a result, the penetration system model for specific protocol-based attacks has often been automated. The application includes several attacks that support hypertext transfer protocol (HTTP), SIP, and TCP/IP for automatic penetration testing. This work aims to provide a rapid, accurate, and automatic testing platform that is easier to use than the current tools [57][58].

3.5.1 Internal Penetration Testing

Internal penetration testing is a type of penetration testing conducted within the internal network of an organization. The applications hosted within the network are tested and help the security team find and analyze any security vulnerability. This helps to estimate the damage that an attack can create within the organization's infrastructure. Internal penetration testing focuses on the attack that can cause by internal sources (refer to section 3.2.1). It involves attacks by angry staff or contractors who have withdrawn but are aware of internal security protocols and codes, attacks on social engineering, simulations of phishing attacks, and attacks exploiting user rights or misuse of an unlocked terminal [56].

3.5.2 External Penetration Testing

These attacks are carried out from outside the organization remotely and require web application monitoring on the Internet. Testers look like hackers who don't know the internal structure very well. Testers are supplied with the target system's IP to simulate these attacks, and no further detail is provided [56]. This helps to understand how deep a malicious attacker might burrow into the network and the business effect of a successful attack; they would also measure the scope of any vulnerabilities found.

3.5.3 Penetration Testing Standards

Penetration testing standards provide an overall roadmap of penetration testing. These standards give a detailed step-by-step explanation of how to conduct successful penetration testing on various platforms and conclude these tests' results. Currently, there are many penetration testing standards available, each with its pros and cons. These standards are

selected for penetration testing are directly dependent on the objective of the penetration testing operation.

Some of the penetration standards which will be discussed here include the Information Systems Security Assessment Framework (ISSAF), the Open-Source Security Testing Methodology Manual (OSSTMM), the NIST SP 800-115, the Penetration Testing Execution Standard (PTES), and the Open Information Systems Security Group (OISSG) [58].

The ISSAF standard methodology in penetration testing provides detailed guidance for initiating and conducting a successful pen-testing. This standard consists of advancing through different phases of pen-testing, which are [59]:

- Information gathering
- Network mapping
- Vulnerability identification
- Penetration
- Gaining access and privilege escalation
- Enumeration
- Compromising system via remote access
- Maintaining access
- Covering the tracks

The OSSTMM is a peer-reviewed security evaluation and research manual that checks information. These details include actionable evidence that will strengthen organizational protection measurably. OSTMM allows one to consider and evaluate how well defense performs. Organizations no longer need to focus on generic best practices when using the OSSTMM because you've already checked information relevant to the organizations' requirements to base their protection choices [58].

The OSSTMM standard is created for both developers and testers who are engaged in the field of cybersecurity. Even though the OSSTMM manual will not help understand network protocols or applications, the users can find it helpful in developing better networks, firewalls, applications, and network testing tools [58].

The standard NIST (SP800-115) offers instructions for preparing and performing checks and reviews of information security. Besides, the effects should be evaluated, and mitigating strategies developed. It is not meant to provide a comprehensive testing or evaluation scheme

but to outline the core elements of both safety testing and evaluation to assure particular methods that explain their advantages and disadvantages. In addition to that, it also includes guidelines for their use and studies. As per this standard, the overall penetration testing process can be grouped into mainly four processes, as shown in figure 3.5 [57][58].

Figure 3.5: NIST Standard Phase [57]

The PTES is a relatively recent standard that started implementation in 2010. One of the most significant aspects of this standard is that industry professionals implement it in particular areas of the process in which they specialized [58]. It is a very comprehensive framework for penetration testing that includes the technological and other essential aspects of a penetration test, such as scope creep, reporting, and security.

According to the PTES, there are seven stages of penetration testing that are shown in figure 3.6. These stages help in customizing the PTES standard according to the testing process's objective and the testing platform [58].

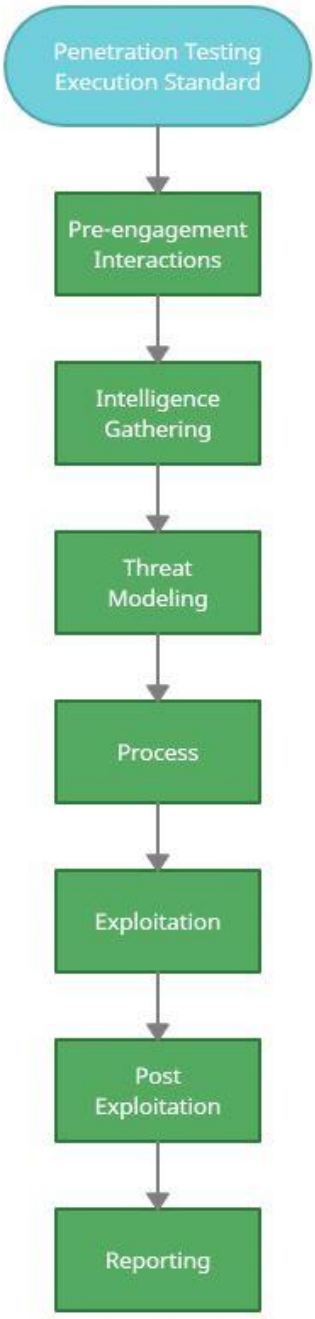

Figure 3.6: PTES Operational Stages [58]

3.5.4 Manual and Automated Penetration Testing- Comparison

For an extended period, penetration testing was carried out manually, including a complicated step-by-step process based on the penetration testing standard. The developers have to manually script the exploitation code based on their testing goal using the tools available in the market, which was time-consuming and complex. Manual penetration testing required a team of professional who was experts in a diverse area of cybersecurity. This skill set requirement

was always a challenge for the organizations in terms of money and effort need to put into this. To overcome the dilemma caused by manual penetration testing, a team of developers created automation tools that users can use with low skill sets. And thus, solved the problem of teams with high knowledge set requirement.

Table 3.1 provides a detailed comparison of manual and automated penetration testing.

	Automated Penetration Testing	Manual Penetration Testing
Testing Process	Replicable tests with fast and standard processing;	Manual, non-standard process, customization cost is high;
Attack Database Management	Maintaining the attack archive and writing modified attack codes for several platforms;	Manual maintenance, code changes across various platforms are manually done.
Exploit Development and Management	The product vendor develops and maintains all exploits. For optimum efficiency, exploits are revised continuously. Exploits are designed carefully, carefully tested, and easy to run. For several platforms and attack vectors, exploits are written and optimized.	Developing and managing an exploit database is time-consuming and requires considerable skills. Public exploits are suspect and may be dangerous to run. Re-writing and porting code is important for cross-platform features.
Reporting	Automated and user optimized.	Manual collection of intelligence.
Clean-up	Tools contain clean-up features in inbuild.	Testers perform clean-up manually.
Network Modification	No changes to the network config	Can cause network config change.
Logging	Automatic logging of system activity.	Slow logging process with chances of mistakes.
Training	Easier and user-friendly UI	Requires high skill set and understandings.

Table 3.1: Comparison Between automated and manual penetration testing [57]

3.6 AI-powered security tools

Artificial Intelligence has proved to solve cybersecurity threads to an extent in a variety of environments. This is achieved by using a wide variety of security applications based on AI techniques with machine learning algorithms' data processing power. In the current world of advanced cybersecurity industries and a community that has built a culture of defending their data from active and passive attacks, it is essential to discuss some of the tools that keep the systems protected from malicious activities.

3.6.1 Intercept X tool

Intercept X tool was developed by a British security software company named Sophos. Intercept X is powered by AI neural network that empowers deep learning techniques, like what we see in the human brain. The US Defence Advanced Research Projects Agency (DARPA) developed its first Cyber Genome Program in 2010 to discover the 'DNA' of ransomware and other cyber threats, contributing to the Intercept X algorithm's development [61].

Intercept X will extract millions of features from a file until a file executes, perform an in-depth inspection, and decide whether a file is benevolent or harmful in 20 milliseconds. The model is trained by access to millions of samples provided by data scientists on real-world feedback and bi-directional threat intelligence exchange. This results in a high accuracy rate and a lower false-positive rate for both current and zero-day malware. To limit new ransomware and boot-record threats, Intercept X utilizes behavioral analysis [61].

3.6.2 Symantec’s Targeted Attack Analytics (TAA) Tool

The Targeted Attack Analytics (TAA) Tool was developed by Symantec, which is powered by AI and machine learning techniques that will be applied to Symantec’s security experts and researchers’ capabilities and processes. The TAA tool has been effective in countering attacks like Dragonfly 2.0, which targets various energy companies' operational networks. TAA tool analyses the vulnerabilities exploited in the Symantec systems and uncovers malicious traffic in the individual endpoints. This malicious traffic is compared with the incidents found before to check any hidden security breach is present or not [61].

3.6.3 Darktrace Antigena

Darktrace Antigena is the active self-defense product of Darktrace. Antigena extends the critical functionality of Darktrace to detect and propagate the feature of automated antibodies that recognize threats and viruses and neutralize them. Antigena uses the Enterprise Immune Systems of Darktrace to detect malicious behaviors and, based on the severity of the threat, acts to them in real-time. Darktrace Antigena detects and defends against significant risks as they evolve with the aid of underlying machine learning technologies. This is achieved without any human intervention or intelligence based on previous attacks. Organizations will respond to threats easily with such automatic response capabilities without disturbing the usual business operation routine [61].

3.6.4 IBM QRadar Advisor

The IBM Watson technology is used by IBM's QRadar Advisor to fight against cyber threats. It uses AI to auto-investigate any compromise or exploit signs. To offer critical insights, QRadar Advisor uses cognitive logic, which further accelerates the response cycle. Security analysts will analyze hazard events with IBM's QRadar Advisor's aid and reduce the chance of ignoring them. IBM QRadar Advisor consists of the following features [61]:

- Provides Intelligent reasoning
- Identifies high priority risks
- Key insights on users and critical assets

Chapter 4: Automated Hacking for Offensive Purposes

Just like a coin have two sides, automated hacking can not only be used for defensive purposes, as seen in chapter 3 but can also put into work in the field of offensive attacks. The offensive method in which automated hacking using AI and machine learning tools are used is always debatable. The ethics and objective of using such offensive methods are questionable and concerning. But the understanding is that if the offensive techniques of automated hacking are used by some government entity or a law enforcement agency, or even an organization under the governing law of that country, it is considered a defensive method.

On the other hand, if the offensive techniques are used by a Blackhat hacker(s) or an organization listed as illegal, to attack any entity like a computer system or an organization to damage, they are considered illegal and remain as an offensive method.

Both offensive and defensive automated hacking using AI and machine learning techniques can be used by attackers like Blackhat hackers and the security professionals (government and law enforcement agencies) who try to stop such attacks. This means that the way any automated hacking is categorized as offensive and defensive is based on various factors like the attack's objective, who is in charge of such attacks, and the mode of operation. Thus, the automated hacking techniques that are used for defending an attack can be turned out to be an offensive attack technique if it is used in an illegal way and vice versa.

Automated hacking for defensive purposes (Chapter 3) dealt with various AI and machine learning techniques used to defend against different automated hacking attacks. These techniques were explained in chapter 3 from an organization or a law enforcement agency's perspective, which depends on such strategies in defending public and private networks, computer systems, and vital intelligence.

This chapter, “Automated Hacking for Offensive Purposes,” deals with offensive techniques used to carry out automated hacking attacks in various environments successfully. Unlike chapter 3, this chapter describes most of the offensive techniques used from a Blackhat hackers’ perspective.

4.1 AI and Machine learning in offensive attacks

In cybersecurity operations, machine learning has proved to be highly effective, particularly for quick pattern detection. There is also an application of machine learning in cyber-offensive operations to boost human operators' current capability set or enable large-scale offensive operations, otherwise complex to do, without a substantial workforce. Machine learning can be

a solution to today's world system's complexities, can support completely autonomous operations in the job mode, and can be the ideal unseen adversary to support asymmetric operations. Machine Learning enabled malnets to have the ability to target local service infrastructure, telecommunications, communications, social services, card systems, hive-networks/botnets [62].

Machine learning is different from the rule-based approach of automation. The critical difference between conventional rule-based and machine learning systems is the way they achieve an objective goal. Machine learning systems can learn and change their actions depending upon the circumstances. By contrast, machine learning is not more efficient or more advanced than automation based on rules. It is simply new and in ways that have thrilling, curious, and disturbing possibilities. A single measurable target is required for machine learning systems to review and maximize their performance [63].

With regard to the fact that open-source AI resources are already being developed, it is reasonable to expect AI technology to be leveraged to build new kinds of advanced and sophisticated risks. Various ways in which AI can be used to perform malicious activities are highlighted in figure 4.1 [64].

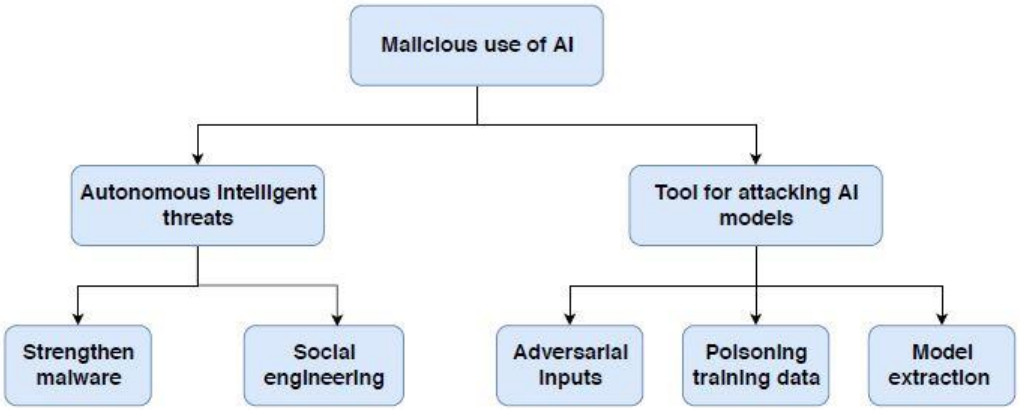

Figure 4.1: AI malicious activities [64]

With the potential to automate human intelligence and procedures and overcome existing human capacities, the risks of AI-enable threats emerge. Threat actors may develop their weapons with AI techniques to make them more autonomous, complicated, and challenging to detect. Many such malicious actors are powered by AI autonomous threats that can be deployed for offensive purposes [64]. Some of them are discussed in the subtopics below.

4.1.1 AI-Powered Malware [64]

AI technology can be armed to improve the malware's effectiveness, making it more autonomous, more advanced, quicker, and more challenging to detect. The latest generation of malware is smarter and able to function autonomously with the help of AI. Intelligent malicious programs may spread themselves based on a set of autonomous choices in a network or computer device, intelligently tailored to the host system parameters, and autonomous malware capable of choosing lateral movement strategies, thus raising the probability of completely breaching the targeted networks [64].

In addition, malicious hackers may implement the ability to adapt to a new environment or use the intelligence learned from past AI incidents to build intelligent viruses and malware or model adaptable attacks. As a result, malware becomes autonomous, blends into its environment, takes countermeasures against security tools, and can use previously acquired data to target the device. Hiding its existence and destructive intent to prevent being discovered by anti-malware solutions is one of the malware's ultimate objectives. Cybercriminals would undoubtedly find ways to incorporate the latest complex technologies into evasive tactics [64].

4.1.2 Social Engineering Attacks

To collect publicly identifiable information, which can be used to hack user accounts, AI can be leveraged to mine vast volumes of large databases containing social network data. Malicious actors may implement AI-based user information to create personalized malicious links or build automatically customized phishing emails [64].

Over the recent years, there has been researching on powering AI in complex social engineering attacks, like the development of Long Short-Term Memory (LSTM) neural networks that can be deployed to manipulate users in social media posts to click illusory URLs [64].

4.1.3 Attacking AI Models

As we know, various AI and machine learning techniques are used to power security solutions, intrusion detection, etc. Some cyber-attacks use automated AI techniques to attack and manipulate these defensive mechanisms. Basically, the logic is that AI techniques are used to attack AI-enabled security systems. These offenses on AI systems are categorized into three areas:

- **Adversarial inputs**: This is a strategy where malicious actors design the inputs to allow models to forecast erroneously to prevent detection. Recent research has shown how adversarial malware samples can be produced to avoid detection [64].

- **Poisoning training data**: The cybercriminals could damage the data sets from which the algorithm was studying about the same kind of attack to detect and recognize functionality is exhausted. Different domains are vulnerable to attacks by poisoning, such as network interference, spam filtering, or analyzing malware [64].

- **Model extraction attacks**: These methods are used via black-box analysis to recreate the detection models or recover training data. On this occasion, through reversing strategies, the attacker discovers how machine learning algorithms work. From this information, the malicious actors can analyze and summarize what the intrusion tools are looking for (attack patterns) and how they initiate actions to stop them [64].

4.1.4 Unauthorized Access

Machine learning can be used to obtain unauthorized access to systems, such as captcha-related systems. Machine learning has significantly influenced machine vision, whereby a machine is programmed to recognize objects. With computers capable of recognizing objects in images, they can be programmed to circumvent a captcha-based framework that depends on a user before being allowed to recognize objects in an image [65].

Machine learning algorithms, such as neural networks that aim and mimic the human brain, can also be programmed to speed up and automate social engineering strategies, such as guessing user passwords by analyzing the model with a massive database that contains data from previously compromised user information, including their usernames and passwords, and any data that can be used to develop the model [65].

4.1.5 Spear Phishing

Another application of machine learning is the ability to carry out advanced spear-phishing attacks by gathering legitimate email data from targeted individuals in an unauthorized manner and feeding the data to a machine-learning algorithm that can then learn, extract meaning from the data and produce emails that look close and genuine to those from which they knew. This can then be integrated into an automated procedure, improving the effectiveness and speed at which targeted cybercriminals can conduct phishing attacks [65].

Several Phishing attacks exploit social engineering to collect data about their intended users unlawfully. Social engineering attack (refer to section 4.1.2) is a standard method of attack technique that uses deception to exploit people to get their personal data. Related approaches can also be used to carry out phishing attacks based on addresses [65].

4.2 Automation Attack Framework

The automation attack framework is used to build a tool to manipulate a network system's vulnerabilities automatically. This approach can explain the feasibility of cyber-attack security techniques in real networks by launching real attacks. Figure 4.2 represents the different phases of the automation attack framework [66].

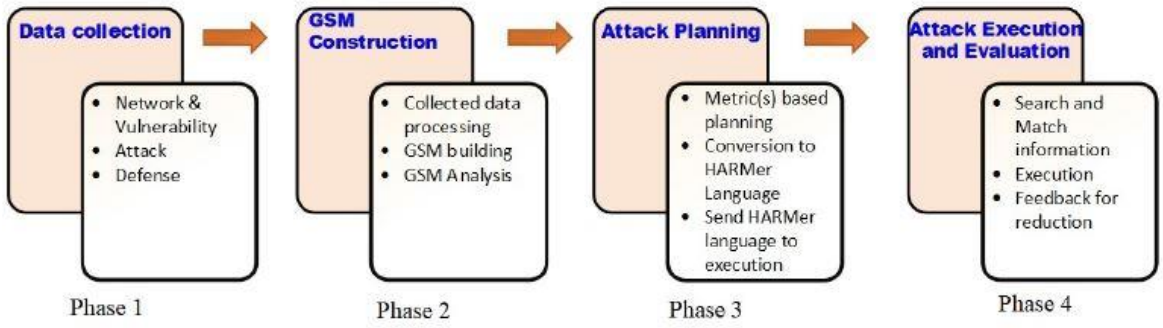

Figure 4.2: Automation Attack Framework Phases [66]

4.2.1 Data Collection

It is possible to gather a lot of data from a network environment. This framework, however, uses only the necessary details needed to create the model of the attack. To gather information automatically, the automation attack platform provides a vulnerability scanning application and a network and open port discovery tool. However, the system is not restricted to knowledge obtained by scanning and exploration tools, as network managers are often empowered to include other information. Security applications such as OpenVAS, Nessus, and Nmap can also be used to gather vulnerability data from hosts, operating systems, port providers [66].

4.2.2 GSM Construction

A two-layered hierarchical attack representation model of the network is created in the second phase using the information gathered in phase 1. In the hierarchical attack representation model, all potential attack paths are captured and enumerated, from which the possible attack scenarios are well captured. The security decision-maker will choose the security metrics to use in the security model for security analysis. Here, to decide the attack strategy, the computed

security metrics via the model will be used. Therefore, depending on the chosen protection metrics, this step evaluates each attack path based on its risk, threat, or the possibility of a successful attack [66].

4.2.3 Attack Planning

For the adversary agent, this stage is responsible for planning and producing actions. Here, a strategy may involve a response from the next host/target, port scanning, IP range, or targeted acts, such as leveraging a host's software flaw and sending a spear-phishing email. With the hierarchical attack representation model, different methods can be used to produce potential attack plans strategically [66].

Here, the model of metrics-based attack strategy and hierarchical attack representation model create attack plans in the form of attack scenarios for attackers. Based on the metrics used in the hierarchical attack representation model, we can produce deterministic attack plans. Attack proposals should be conceived in the language of the attack [66].

The primary reason for using an attack language is to make it possible for the generated attack plans to be universal and useful for numerous forms of attacks and defensive instruments. It is also possible to translate attack plans written in a universal attack language to an appropriate format for specific attacks and protection tools [66].

4.2.4 Attack Execution & Evaluation

In this phase, the attacking tool's output, which was executed in the attack planning phase, is transferred to the execution and evaluation phase as values for assessment. These attacks and evaluations are based on the intelligence collected during phase 1, i.e., information collection [66].

4.3 Cyber Kill Chain [63]

The cyber kill chain is a well-recognized structural concept of intellectualizing cybersecurity operations by providing a sequence of procedures that attackers work through on their journey to their target. Lockheed Martin adapted the cyber kill chain in 2010 from the concept of traditional military kill chain methodology. The kill chain methodology consists of a targeted step-by-step attack on the targets. Kill chain comes with some limitations, like portraying cyber operations as overly linear [63].

Attackers are going to execute any or all the steps of the kill sequence. Using widely used exploit tools and techniques, attackers can integrate multiple steps depending on their overall goal. Six necessaries, such steps will be discussed as the sub-topics below. All these steps are continued to evolve over the years with automation [63].

4.3.1 Reconnaissance

Attackers must first select their target because the reconnaissance and mission selection process depends on the targets. Some attackers will be focused on attacking broad user categories and will largely be forgoing this process. In their identification of victims, some are more selective, involving a more thorough reconnaissance campaign. At this point, attackers first identify people and computers worth exploiting and then collect data on those targets' technological vulnerabilities [63].

Using search engines like google and other sources like social media research and scrapping technological web forums, attackers will collect valuable information about an entity and its employees to inform their hunt for human goals. These passive strategies have the bonus that they are mostly undetectable [63].

Traditional automation strategies provide a means of storing, processing, and reviewing information obtained in this manner, significantly shortening the time taken in this process and helping attackers prepare their next step. To classify the victims most vulnerable to a range of social engineering tactics, those techniques may be enhanced by very straightforward machine learning-enabled methods [63].

Attackers may use more effective tactics such as automated scanning that evaluate targeted networks for specifics of their linked servers, network protections, and related device settings to inform their search for computer targets. This type of aggressive reconnaissance is incredibly popular, and a wide range of malicious actors are actively searching most computers on the internet, all of which are looking for vulnerabilities to attack [63].

4.3.2 Weaponization

With their targets established, to obtain unauthorized entry, attackers have to identify and manipulate technological vulnerabilities in the software of their target. To build a package that is later sent to their target, attackers must then combine their malware with a vulnerability. This mechanism is called weaponization. The correct exploit code takes advantage of the

vulnerabilities recently found and allows the attackers the ability to operate on the network of the target, even by staying undetected [63].

To hack them, automated weaponization software can easily find bugs and compile code. These tools also feature vulnerability libraries that attackers can browse through to find ones that fit their goal's apparent vulnerabilities. To uncover new vulnerabilities, attackers frequently start by analyzing the code that runs on the device of their target, again using information gathered during the reconnaissance process. In this method, a tool named Fuzzers can help. Fuzzers check out glitches and flaws by attacking and tracking the results of a chosen piece of software with multiple inputs. These inputs may be totally random or customized to the tested program [63].

4.3.3 Delivery

The attackers must now complete the somewhat trivial and often challenging process of deploying the code into their target device after executing reconnaissance and weaponizing a software package. This can be achieved in many ways, like an intrusion into the targeted device using a computer or human vulnerabilities [63].

Malicious codes can be delivered using attacks like watering hole attack, in which attackers hack a legitimate website and infect all of its users with an exploit targeted at their browser. Other operations spread through malicious code-infected USB drives. Whereas some other operations are also conducted by third parties with whom the target communicates. Such operating strategies depend on the attacker going "upstream" to a trustworthy entity for which the victim has little influence, such as a business that provides the target with IT services or other applications. These include compromising the Wi-Fi router or any other access gateways [63].

Attackers also use social engineering to trigger actions that undermine an institution's security while exploiting human vulnerabilities. The assault tactics used against persons are as complex as our perceptions. These include faked calls in which an emergency is created to extract vital information from the employees or spear-phishing or even spoofing email addresses [63].

4.3.4 Command & Control

The next move is to create a safe line of contact with the code they have put after attackers successfully penetrate their targeted device. Attackers will pilot their malicious script and execute the command through this network; this is known as command-and-control (C2).

Attackers build and design their C2 infrastructure based on the victim's cybersecurity setup, the malicious code's goal, and the frequency they need to exchange commands. There are trade-offs between acceleration, stealth, and durability in C2 structure variants. Attackers favor speed and the potential to exfiltrate vast quantities of data in some instances. To prevent detection by adversaries, other cyber practices emphasize secrecy and use delay-tolerant C2, transferring information across circuitous networks. When attackers' aims evolve with each hacking effort, their methods and strategies often do so [63].

4.3.5 Pivoting

The process of infecting multiple devices using a device that the attackers already compromise is called pivoting. Pivoting is an important step of the cyber kill chain to achieve an operational objective and dominions. Most of the time, pivoting operation's primary purpose is to extend to as many machines as possible. However, indiscriminate pivoting also raises the possibility of discovery. Many tasks systematically concentrate massive effort in pivoting to find the most successful follow-on systems in steady progress towards the ultimate target [63].

Two distinct components are used in each pivoting method: privilege escalation, which involves obtaining additional access to a compromised device and privileges, and lateral movement, which uses privileges or program bugs to gain access to other devices [63].

Pivoting can use tools that take advantage of the same attackers with technological or human vulnerabilities to obtain initial access to a network. For example, attackers who have infiltrated trustworthy email accounts within a network, such as IT staff manager accounts or senior staff accounts, can use them to participate in more spear phishing, specifically targeting accounts with even higher administrative rights to gain additional passwords and access. Otherwise, a software vulnerability that allows unauthorized access to one computer on a network will operate almost as well against other devices as well [63].

4.3.6 Actions on target

Actions on target are the final step on the cyber kill chain. The steps from reconnaissance to pivoting were performed to reach this final step. If the attackers succeed at all stages in the kill chain, they will eventually act against their target to achieve their mission. When the intruder has checked that they have entered their target machine, various objective activities will begin, including frequently validating the computer's name, files on the computer, or its location within the network [63].

Other offensive attacks, including operations that encrypt sensitive data before a ransom (Ransomware attack, section 4.5) is paid or breach financial networks and allow illegal payments, may be motivated by benefit, a strategy that North Korea has used to lavish itself with tens of millions of dollars. The purpose of asset hijacking, in which a network is manipulated to exploit its computing resources for some other aim, like turning the resources into cryptocurrency mining machines [63].

4.4 Automated Intrusion Attack

A network intrusion can be explained as unauthorized access into a computer system or a network device to steal/modify or destroy data for various purposes. These unauthorized activities can be part of one or more cyber kill chain (section 4.3) steps for accomplishing multiple objectives.

The attacker initially generates malware payloads and then hosts them on the vulnerable web server to gain complete access to the target users. One of the most important techniques that can be used to build malware is the Social Engineering Toolkit (SET) with the Metasploit System built-in. Also, various methods such as "Java Applet Attack," "Browser Exploit," and "Spear-Phishing Attack" are used by attackers to mask their attacks [67].

The "Java Applet Attack" strategy consists of building a malicious applet that, when deployed, totally exploits the victim. In the “Exploit Browser” attack technique, the attacker integrates the malware code with a common website page code, where the code will be hidden entirely from the victim. On the other hand, Spear-Phishing attacks, as discussed before, approach by phishing/masking emails with malware [67].

Another important technique in intrusion attack is the use of a Meterpreter. The Meterpreter is a versatile shell of commands used to communicate with victims. A significant benefit is that Meterpreter runs in the victim computer's memory exclusively and does not use the hard drive. This technique makes it easy to circumvent several antivirus systems and guard against digital forensics. The next step of the intrusion is to convince the user to access the malicious web server. Phishing email messages are one of the most commonly used ways to fool their victims. This approach can be used on both a local or a WAN infrastructure to attack victims [67].

Another powerful social engineering strategy is DNS spoofing. Hackers use this technique to compel users of a LAN to access a malicious server. If efficient, a script that will be executed will be inserted, and an interface window will open immediately from which the hacker will

take control of the victim's computer. At this stage, the intruder would be able to monitor the victim's machine and then transform it into a permanent zombie. This can be accomplished either by running scripts from Meterpreter or by inserting resources that act as a backdoor. Netcat is a very powerful tool that can be deployed to achieve this objective [67].

4.5 Ransomware Attack

4.5.1 Introduction to Ransomware

Ransomware is a type of malware that stops users from accessing different computing resources and the data stored in that computing resource. This is achieved by various techniques, which will be discussed in sub-section 4.5.3.

Ransomware is a type of malware that prohibits people from accessing different methods to access their operating system resources and personal data. Ransomware does not plan to harm the machine file system but simply leaves it accessible to reveal the ransom note on the victim's monitor and provide the victim with a way to pay the ransom. The data on the victim's electronic system remains inaccessible before the computer owner pays a ransom to lift the constraint [68].

Ransomware-infected operating machines will typically display a display notice that threatens the recipient that they must pay the ransom within a short amount of time; otherwise, irreversible data loss will occur. Ransom money is then obtained by the attackers behind the malware using anonymous financial transactions to narrow their monetary track [68].

Initial ransomware variants used traditional techniques to restrict user access to resources/personal data for computer devices, including refusing access to machine tools or the desktop. Modern variants of ransomware use cryptography to firmly lock the user's private files, rendering it irreversible without the cryptographic signature correlated with it [68]. The key difference between various types of ransomware is the technique they use to restrict users from accessing their data and computer resources.

Figure 4.3: Ransomware message displayed on an infected device [68]

Ransomware extended its activity to attack suppliers of services and continued to use offensive methods to extort money from its victims. For example, hackers of ransomware compromised websites of pornography and pirated apps and thereby targeted their users by claiming that they had broken the law by accessing child pornography content or downloading pirated technology they had made a copyright violation. Therefore, the local law enforcement department has locked their confidential files, and they can pay the police a fine to restore access to the locked computer [68].

Figure 4.3 is an example of how the ransomware attack warning will be displayed in the victim's system.

4.5.2 Types of Ransomware

There are two main types of ransomware in use. The main objective of both the type is to block the user from gaining access to his/her computer resources as well as personal data will remain

so until they try to pay a ransom. But the approaches adopted by each type of ransomware are very different, despite sharing common aims [69]. The two main types of ransomware are:

- **Locker Ransomware**

  Locker ransomware is intended to restrict access to the resources of computing. This usually takes the form of locking the computer or device's user interface and then requiring the user to pay money to access the system back.

  Even though the computer is locked, the hackers will leave behind some computers' capabilities, like allowing the user to communicate with the ransomware and pay the ransom. This means that access to the mouse could be disabled, and the functionality of the keyboard could be restricted to numeric keys, enabling the victim to indicate the payment code only by typing numbers [69].

  Usually, Locker ransomware is only intended to block access to the computer interface, keeping unaffected the underlying device and data to a large degree. This ensures that returning a device to anything similar to its original form when the malware is removed. This makes locker ransomware less powerful than its more damaging relative crypto-ransomware in collecting ransom payments [69].

  Since locker ransomware can typically be efficiently eliminated, it appears to be the kind of ransomware that goes out of its way to combine social engineering strategies to force victims into paying. As law enforcement agencies, this sort of ransomware also classifies and claims to issue fines to users for suspected digital indiscretions or illegal activities [68][69].

  The effectiveness of the locker ransomware is very high on those devices with minimal options of user interaction. In this contemporary world full of different IoT devices, locker ransomware is a huge problem.

- **Crypto Ransomware**

  Crypto ransomware is another type of ransomware designed to find and encrypt data stored in the infected system. This makes the entire data un-processable, thus making it useless. The only way to retrieve the original information is by decryption suing a

decryption key obtained from the attacker by paying ransom amount through different financial channels [69].

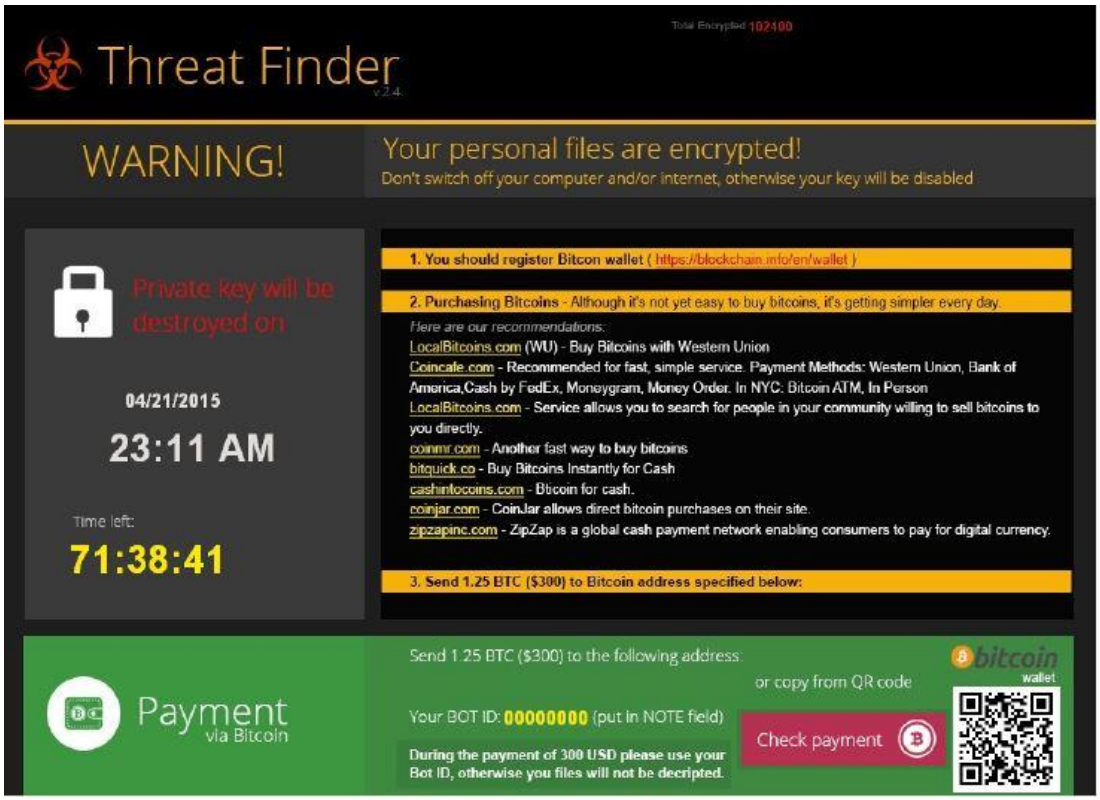

Figure 4.4: Crypto ransomware demand screen [69]

A common crypto-ransomware threat secretly looks for and encrypts data after download. It aims to remain off the radar until any of the files used by the customer can be found and encrypted. The data will be encrypted by the time when the user is confronted with the ransomware warning that tells them that their data is encrypted. The impacted computer continues to run like most crypto-ransomware attacks, as the malware does not target sensitive system files or refuse access to the system's functionality. This ensures that, aside from data access that has been encrypted, people will also use the device to perform several tasks [69].

Figure 4.4 shows how the ransom warning message will be displayed in the victim's system if crypto-ransomware attacked it. It clearly shows how locker ransomware is different from crypto-ransomware in terms of denying the system's capabilities.

4.5.3 Ransomware Techniques

As the different types of ransomware are present with one primary objective, the techniques used in ransomware attacks may vary. Still, all the methods are executed only to achieve one primary task: to lock the system in such a way that the victim will be forced to pay the ransom amount. Given below are the various techniques used to execute ransomware attacks.

- **File Encryption**

  Typically, conventional crypto-ransomware employs both symmetric and asymmetric methods of encryption. A single key is used in symmetric encryption to encrypt the data, and then the same key is used to decrypt the encrypted data. Identifying the secret helps the user to decrypt data that has been encrypted.

  Ransomware typically produces a key on the compromised device using symmetric encryption and transfers it to the attacker or demands a key from the attacker until the user's files are encrypted. Since encrypting their data, the intruder has to guarantee that the key is not accessible to the victim, or the victim would be able to decrypt the data themselves without paying [69].

  The advantages of using symmetric data encryption are that they are usually simpler than asymmetric algorithms and use small keys to perform the encryption. Standard crypto-ransomware has a vast amount of data to scan and encrypt rapidly. Hence, efficiency and speed are important for encrypting files before the victims notice the abnormal activates in the system [69].

  In asymmetric encryption, the attacker uses two keys, a public key, and a private key. The public key is used to encrypt the data, whereas the private key is used to decrypt the data. There is a significant advantage for the attacker in employing asymmetric encryption. The symmetric encryption technique is easy to decrypt the data if the public key got exposed. In contrast, in asymmetric encryption, the victim cannot decrypt the data even with the correct pubic key [69].

  A mixture of symmetric and asymmetric encryption methods usually utilizes more sophisticated crypto-ransomware. Unique public-private key pairs for each victim computer can also be created by versions that use asymmetric encryption. This helps the attacker decrypt files on one victim computer without exposing the private key that

could theoretically be used to decrypt files with much the same pubic key on any other infected computer [69].

Some of the crypto-ransomware encryption approaches are:

- Downloaded public key
- Embedded public key
- Embedded symmetric key

- **Screen Locking**

Figure 4.5: Browser locking ransomware source code [69]

Locker ransomware seeks to restrict the compromised users' connections to the operating system and applications running on their server or computer. The most widely used approach is to show in a repetitive cycle a ransom note to the victim. This provides the idea that, while there might be minor moments where the user may close the message's current window, the message is continuously displayed.

The ransomware that uses the screen locking technique will employ the API or other features associated with the operating system deployed in the victim's computer. The screen-locking technique is further classifies into Windows locker ransomware, Browser locking ransomware or Android locker ransomware based on which feature is used.

Figure 4.5 represents the source code from the browser locking ransomware with functions to display ransom messages multiple times [69].

Chapter 5: Industries & Law Enforcement Agencies

By now, it is clear how automated hacking is treated from a defensive point of view and offensive point of view. And how AI and machine learning have become the catalyst in increasing the efficiency and penetration capabilities of automated hacking. Automated hacking has evolved over the years with new emerging technologies, increased connectivity, and digitalization over many operational areas.

One of the most significant areas in which automated hacking has affected or was a huge topic of concern is the government agencies, including law enforcement and various industries, who were dependent on the computer and network infrastructures for their daily operations. It is always challenging for these organizations to maintain their system infrastructure security and integrity, and confidence. The cost of doing this is still high and requires dedicated professional help to achieve that.

Nowadays, most government agencies, departments, and industries have a dedicated cybersecurity team who monitor, develop and defend (some cases offense as well) the network and system integrity of their organization from targeted automated hackings using various tools powered by AI and machine learning techniques.

5.1 Government Agencies and Cyber Security Challenges

To achieve productivity and improve public participation, the government depends heavily on information technology. A rise in cyber threats and data leaks that impact government activities has created a perfect combination of challenges and difficulties.

Cyberattacks appear to have an outsized effect on departments and agencies' activities, which may undermine public interest and decrease the capacity to execute vital mission functions. The U.S. government is transforming how it handles cyber challenges and broader market risks in response to disruptive cyber-attacks, data breaches, budget constraints, and consumer perceptions [70].

The U.S. government came up with a new set of strategies that can be applied in different branches of the government and private sectors to safeguard and defend sensitive data and networks from cyber-attacks. These strategies are [70]:

- Active threat hunting.
- Collaborate with different departments to share intelligence data.
- Non-stop network scanning and monitoring of active malicious actions.
- Automation and Orchestration of security procedures.

5.1.1 Cyber Threat Hunting and Intelligence Sharing [70]

As a logical way of detecting active risks, the federal government turns to cyber threat hunting because conventional preventive and intervention strategies are frequently unsuccessful against dedicated opponents. The ability to successfully scan for endpoints and recognize sophisticated threats is an evolving task involving specialized software, technologies, and individuals to uncover both the external sources of applications and data breaches and internal exploits. To allow cyber threat hunting operations in complex environments, it is necessary to obtain and retain maximum visibility of malicious attackers targeting a particular area [70].

Intelligence obtained from sharing sensitive information across various departments is now proactively integrated into Indicators of Compromise (IOCs) to check for suspicious behavior signals, such as malicious users that may collect data and escalate power. Such activity is likely to originate from risks that have not been adequately classified or contain previously undisclosed threats. For IOCs connected to nation-state threat actors, this allows researchers the opportunity to analyze multiple machine objects [70].

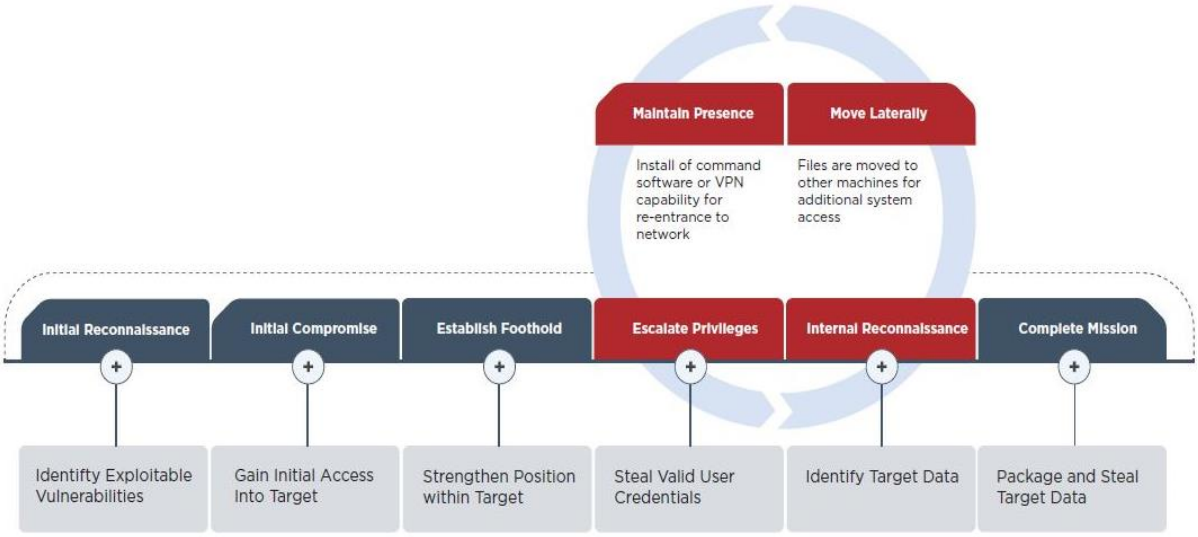

Figure 5.1: Attack Lifecycle [70]

The application of advanced identification systems to scan for unique IOCs and conduct sweeps directly identified with sophisticated malicious attackers targeting federal agencies is used in modern hunting strategies. This technology allows researchers to analyze multiple device objects associated with the nation-state terrorist and other specialized threat actors for IOCs.

In addition to the automatic IOC sweeps, researchers use event analysis frequency to gather and analyze data to discover occurrences that may have gone unnoticed with previous steps. This approach helps analysts to identify anomalies not observed by IOCs in the environment [70].

The IOCs used to check for other malicious behavior indicators, such as data mining and privilege escalation by illegal access, are conveniently formalized with intelligence gathered from these hunting strategies. Figure 5.1 shows the overall attack lifecycle and how it is integrated into IOC's information gathering. Such approaches can facilitate proactive scanning for other malicious behavior signs, such as non-target and commodity-based malware, which can also have catastrophic consequences [70].

5.1.2 Monitoring & Management of malicious actions

The Office of Management and Budget (OMB) developed a broad cross-agency goal in 2013 to allow all federal computer systems to be continuously operated and controlled. This change characterized the primary identification of the shortcomings of a static, decade-old continual appraisal and authorization approach within a diverse and integrated information infrastructure. It also accepted the limits of the government's ability to protect against some of the more severe risks clustered against such structures [70].

A new federal Continuous Diagnostics and Mitigation Software (CDM) was authorized by the Department of Homeland Security (DHS) to promote this systematic improvement in the government's plans to protect sensitive systems and intelligence. [70].

The CDM software helps government departments and agencies to broaden their continuous surveillance and diagnostic capabilities by increasing their sensor power, automating data collection, and prioritizing risks. The program was designed to combine government network infrastructure with commercial technologies [70].

The first two main phases, which occurred from 2013 to 2017, concentrated on the essential asset and vulnerabilities management technologies and identity, credentialing, and access management capabilities. Phase three is committed to border security, which started in 2017 and will proceed for many years to come [70].

Figure 5.2 represents the CDM program with its three phases introduced over the years. The three phases are highlighted in red, followed by the areas of capabilities that must be applied to all assets throughout the CDM program processes.

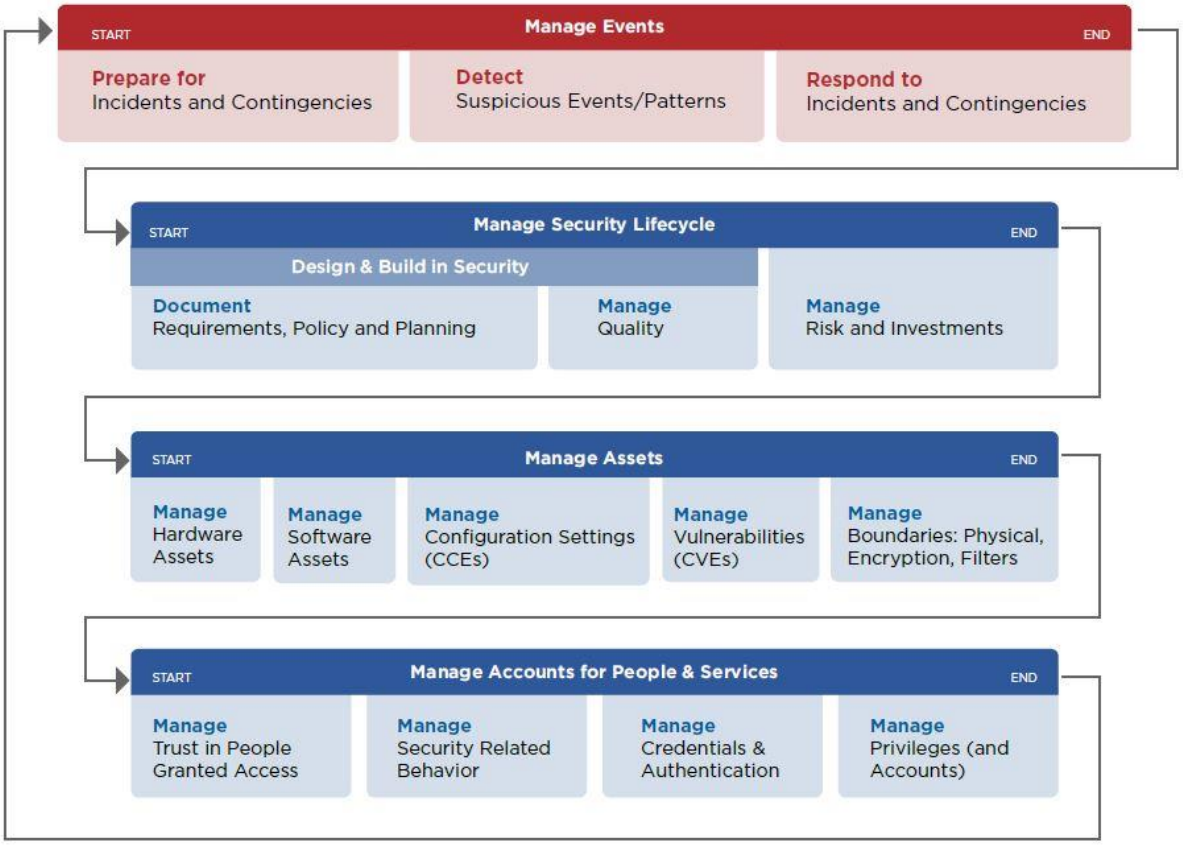

Figure 5.2: CDM program with different phases [70]

5.1.3 Automation and Orchestration

There many agencies that are involved in protecting the cyberinfrastructure of the government. These agencies have different objectives and modes of operation to secure the infrastructure. Due to this, the agencies face multiple limitations on a daily basis. These limitations include skilled human resources required to perform these operations and data analysis, slow incident resolving time, error-prone and unpredictable manual remediation procedures, and limited time to respond [70].

To overcome all these limitations, the concept known as security orchestration is introduced. The security orchestration will enable the process in which the security resources and the integrated disparate security systems can be linked to each other to reduce the human analysis and interaction and allow automation. It will enable the company to have a mature safety environment and correctly identify actionable accidents [70].

5.2 Red Team as a Service

In recent years, the number of automated hacking has increased tremendously, which made various companies worldwide enforce strict security procedures and data handling culture throughout their organization. These security procedures can be implemented by introducing the concept of "Red Team as a Service."

Figure 5.3 picturise how Red Team as a Service work with its four elements in a cooperative environment.

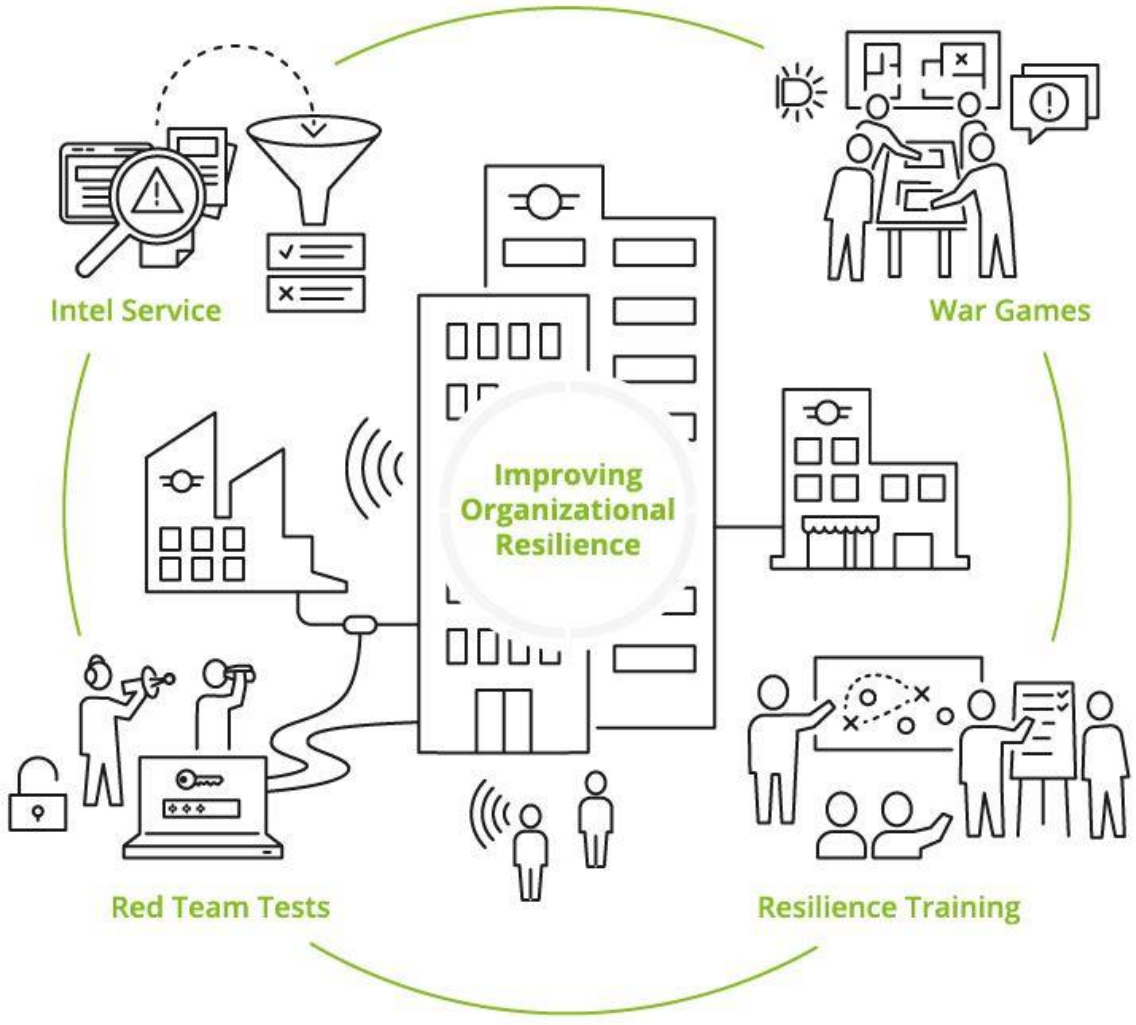

Figure 5.3: Red Team elements [72]

5.2.1 What is Red Teaming?

Red Teaming is a multi-layered, full-action simulation intended to test how well the employees and networks, applications, and physical security controls of an organization can withstand an

enemy's attack in real life. Red teaming is always known as groups with ethical hacking ideology. They use various defensive and offensive hacking techniques to collect intelligence and thereby securing their company's data and network from a cyber-attack [71].

The red team focuses on finding the vulnerability in different areas of a company. These vulnerabilities can be seen from a company's technical aspects, including the way networks are configured, the firewall configurations, applications, etc. Or it can be found from people aspects (employees) of the company or from the physical point of view like building infrastructure. Using social engineering and physical, software, and network penetration experiments to identify ways to improve businesses' defenses, red teaming enables an organization to remain competitive while maintaining its corporate interests [71].

5.2.2 Intelligence service

A good understanding of the potential risk and the number of threats that can face is always vital in successfully executing read team operations in a company. Intelligence service in red teaming focuses on gathering, processing, and dissemination of data for specific purposes. These specific purposes can identify potential vulnerabilities from inside and outside the company, analyze the latest malware on the internet, etc. Accurate awareness of dynamic and rapidly moving trends in this field forms the basis for all resilience-focused behavior [72].

5.2.3 Red Team Tests

Red Team tests focus on simulating a realistic adversarial attack on a company. The objective of these realistic tests is to identify and exploit the vulnerabilities in the company's network infrastructure. This is also known as the process named penetration testing (Section 3.5). By doing this, the red team can demonstrate and analyze the potential effect of such attacks on the business-critical assets [72].

The Red Team performs these attacks by referring specific attack scenarios and behaviour derived from the previous attacks and the intelligence collected. Declarations about probable hacker groups with their corresponding purpose, skills, and capabilities are an important component of the scenarios. An attack scenario defines a clearly specified target, including a description of the negative impact that a successful attack will have on the organization concerned [72].

The Red Team Tests aim to explore and exploit undisclosed weaknesses and threat actors that can be used to target business assets by, for example, simulating process disruption or personal

data theft. The Red Team assesses the physical protection of buildings, the security of networks and software, and the possibilities for targeted manipulation of people while designing these scenarios and carrying out the experiments. The attack scenarios involve elements of physical infiltration strategies, social engineering, and intrusion techniques integrated to enable the Red Team to achieve the goals set [72].

5.2.4 War Games

War Games are scenario-based exercises that challenge and measure the responsiveness of the company. In these stages of Red Team threat evaluation, the core participants are the representatives of crisis teams. So why this War Game is so important?

Corporate incidents are complicated and can suddenly and unexpectedly strike any company. Various previous crises have shown that a problem's magnitude can be dramatically minimized by swift, targeted intervention. An essential pillar of a successful company is a well-rehearsed, fast, and self-assured crisis management team. Daily training of those involved and proper documentation of the required processes and procedures is indispensable for developing and sustaining an adequate capacity to respond [72].

Crisis teams are equipped and educated for incidents in a regulated way within the War Games context. The simulations' scope and scenario are clearly decided in advance and tailored to the company's interests and capabilities. The scenarios selected for the War Games are based on the intelligence gathered. This intelligence analysis decided the complexity, intensity, and duration of the War games [72].

5.2.5 Resilience Training

Resilience training includes the full-scale training of the employees on specific topics of interest. The training program structure will be based on analyzing the outputs derived from the rest of the three Red Team elements. This separate training program will meet the requisite participants and provide maximum added value in a tailored manner. The variety and training style are personalized and continuously explicitly adapted to the company's needs [72].

Chapter 6: Future of Automated Hacking

6.1 Covid-19 and Cyber Security Impacts

The Covid-19 has created a drastic change in how cybersecurity should operate in the future, especially with automated hacking. There is no doubt that the pandemic has made an immense challenge in all the fully or partially dependent organizations on information technology for their day-to-day operations. These organizations who were long dependent directly on various IT facilities like data centers, servers, cloud systems, digital services are now (in the future as well) have to be accessed remotely. This new operation mode comes with enormous security challenges, especially to the organization with a small or limited capacity to deal with various cybersecurity incidents [73].

Simultaneously, the growth in connectivity and the widespread transition to run businesses online have increased the chances of security breaches by an enormous margin. The perimeter protection of companies is at risk of being violated. At both digital and physical entry points, they always need monitoring and real-time risk assessment for threats. Experts in cyber risk management now have to secure their businesses on a large scale and rapidly. They must ensure that their companies' web applications and digital channels are resistant to cyberattacks [73].

These remote working have led to installing new applications to the systems without any security checks. This software increases the risk of hacking sensitive data from less secure workplaces. To execute operations remotely, business executives, administrators, and their employees need access to internal resources and applications. Because many businesses haven't made these apps and data accessible previously over the internet or virtual private networks (VPN), security professionals are hesitant to allow access without strict access protocols [73].

Implementing strict security measures to overcome this challenge will take time and can only be implemented by continuous threat analysis and behavior challenges of the latest malware created after the pandemic has started, designed to target remote workstations. The attackers are taking advantage of this time-inconsistency to attack vital systems and create multiple backdoors to access those systems in the future.

Automated hacking in the future will be much more sophisticated that can take advantage of this current cybersecurity vacuum created in many individual and organizational systems. From the attacker's point of view, the AI and machine learning algorithms can run specific programs to identify such vulnerable systems to carry out target specific automated hackings. Whereas in the case of security professionals, this situation has created a new platform for studying and developing AI-based tools to defend such automated hacking and analyze the intelligence

gathered from the recent attacks based on those newly developed automated hacking algorithms.

6.2 Quantum Computing and Cybersecurity

Quantum Computing is a new computing method in which the computer uses the quantum mechanical phenomena of superposition and entanglement to construct states that scale exponentially with the number of qubits or quantum bits [74].

Traditional computers use the binary system of 0's and 1's to carry out all the data processing. But, in Quantum Computing, the qubits are used to process data. Because of their unique characteristics and behavior, these qubits can exist in more than one state at a time. For example, a qubit can be represented as 1, 0, or 1 and 0 simultaneously for a given time [74].

This helps Quantum Computers operate in parallel on multiple computations and thus pick up the time it takes to process a task tremendously. As a result, Quantum Computers can resolve almost impossible problems to calculate with currently available computing resources, including supercomputers [74**]**.

The computational power of Quantum Computing is beneficial in many areas like health sciences, space explorations, disaster management, etc. But, one of the fields where Quantum Computing poses a significant threat is the area of cybersecurity. One can argue that Quantum Computing will be so beneficial in cybersecurity and provide a great deal of protection against automated attacks. Even though it is valid, the full picture shows the other side of Quantum Computing.

Quantum Computing can be an unsolved threat to cryptography. To limit who can process the information, cryptography defines the practice of translating simple data into ciphers. Two primary forms of cryptography are available: symmetric and asymmetric cryptography. The difference between symmetric and asymmetric cryptography is that symmetric cryptography uses the same key for encrypting and decrypting data. In contrast, asymmetric cryptography uses a publicly shared key (public key) for encryption and a privately shared key (private key) for decryption [74].

Cryptography has great use in transferring data across an unknown network (e.g., Internet) securely. But the issue is that most cryptographic techniques rely on computational calculations, and the security that this calculation is cannot be solved. However, as soon as sufficiently powerful Quantum Computers exist, these cryptographic calculations can be

solved and thus most of today's traditional cryptographic systems will no longer be safe and object to attack, hence eavesdropping and theft of digital identity [74].

There is no proper Quantum Computer ever developed globally with sufficient power to break the cryptographical calculations. But, in the future of automated hacking and Quantum Computing will have a compelling and significant role. There will be a time when these Quantum Computers can automate and launch a targeted attack on a network or a system that can break the asymmetric cryptography, thus completely wiping out the encryption security we have developed now to access sensitive information without any resistance.

Chapter 7: Conclusion

Machine Learning and AI technology have made noticeable changes in implementing and processing various computer engineering & networking devices. And thus, there is no reason not to believe how machine learning and AI have improved the success rate of automated hacking for offensive and defensive purposes.

Hacking has evolved very much from the day it was first reported to this day. The advancement of new technologies and the digitalization of everything worldwide has become the catalyst for increased cyber-attacks. This also provides a unique platform for creating new types of attacks, from the basic DDoS attack to the advanced WannaCry ransomware attack.

One of the advancements that the cyber-attacks have illustrated is Artificial Intelligence and machine learning techniques in powering automated hacking attacks. This upgrade in automated hacking has created a new era of advanced remotely operated and autonomous hacking tools. The complex algorithms used to collect the intelligence on a system or a network and thereby assess the system's vulnerabilities will help any attackers with the upper hand in a successful infiltration. The AI and machine learning algorithms are so advanced that they can carry out surveillance, intelligence analysis and attack a system without any human intervention.

Now, automated hacking is so advanced and has reached a point in which the tools are aware of the steps that need to be taken based on the scenarios; this has created a new much of complexity for the security professionals who are dedicated to defending their respective systems and networks from any form of cyber-attack.

Chapter 3 and 4 help to understand the difference between how automated hacking can be used in an objective defensive way and for offensive purposes. The attackers with malicious objectives or otherwise called Blackhat hackers are generally categorized as people who employ automated hacking for offensive purposes. In contrast, the security professional, or otherwise called White Hat hackers, tend to be the people who use automated hacking for defensive purposes.

Chapter 4 discussed how an organization or an individual uses automated hacking for both offensives and defensive purposes. Any automated hacking is categorized as offensive and defensive is based on various factors like the objective of the attack, who is in charge of such attacks, and the mode of operation. This is clear when the government departments, law enforcement agencies, and significant companies approach cybersecurity from a Red Team point of view.

The government and law enforcement agencies depend on automated hacking to attack any potential enemy and defend the public infrastructures and sensitive data from domestic and international cyber-attacks. Overall, the National Security Agency (NSA) and the US government have used mass surveillance tools like XKeyscore and PRISM to conduct targeted automated hackings on domestic and international entities for national security.

As discussed, the techniques, concepts, and the number of hacking incidents have drastically changed over the last few decades. Likewise, it is sure that automated hacking capabilities and potential threats will increase over the years. The fast-growing Internet across the world and the introduction of satellites that can bring high-speed Internet even to the remote parts of the world, and the consistent presence of human errors and quantum computing will become the catalyst for automated hacking to increase and evolve in the future.

References

[1]. Hall, Gary, and Erin Watson, “Hacking: Computer Hacking, Security Testing, Penetration Testing and Basic Security”, CreateSpace Independent Publishing Platform, Dec 20th 2016

[2]. J. Fell, "Hacking through history," in Engineering & Technology, vol. 12, no. 3, pp. 30-31, April 2017, doi: 10.1049/et.2017.0320.

[3]. Janwar, I. The Complete History of Hacking. www.Academia.Edu. https://www.academia.edu/4903685/The_Complete_History_of_Hacking

[4]. S. Lukasik, "Why the Arpanet Was Built," in IEEE Annals of the History of Computing, vol. 33, no. 3, pp. 4-21, March 2011, doi: 10.1109/MAHC.2010.11.

[5]. Security, H. N. (2019, March 11). The History of Hacking. Help Net Security. https://www.helpnetsecurity.com/2002/04/08/the-history-of-hacking/

[6]. “The Hacker Mystique” [ONLINE]: http://www.cs.mun.ca/~harold/Courses/Old//CS1401.F17/Diary/lect8.pdf

[7]. PCWorld.com staff. (n.d.). CNN.com - Timeline: A 40-year history of hacking - November 19, 2001. Copyright (c) 2001 Cable News Network, Inc. All Rights Reserved. https://edition.cnn.com/2001/TECH/internet/11/19/hack.history.idg/?related

[8]. The History of Phone Phreaking. Historyofphonephreaking.Org. http://www.historyofphonephreaking.org/faq.php/

[9]. H. Orman, "The Morris Worm: A Fifteen-Year Perspective" in IEEE Security & Privacy, vol. 1, no. 05, pp. 35-43, 2003. doi: 10.1109/MSECP.2003.1236233

[10]. Rasmussen, R., & Vixie, P. (2013). Surveying the DNS Threat Landscape. 1–6.

[11]. Lakomy, Miron. (2013). Cyber Threats at the beginning of the 21st Century. Przegląd zachodni. 191-208.

[12]. Krenn, R. (2019, June 13). The Rise of Automated Hacking. Infosecurity Magazine. https://www.infosecurity-magazine.com/infosec/the-rise-of-automated-hacking-1-1-1/

[13]. Sophos Labs. (2019). Sophos 2020 Threat Report Sophos 2020 Threat Report Contents.

[14]. How To Build Automation Scripts without Code | HelpSystems. (n.d.). Helpsystems. https://www.helpsystems.com/blog/how-build-automation-scripts-without-code

[15]. IBM Knowledge Center. (n.d.). Www.Ibm.Com. https://www.ibm.com/support/knowledgecenter/SSWT9A_7.6.1.2/com.ibm.mbs.doc/autoscript/c_automation_scripts.html

[16]. Hanna, M., El-Haggar, N., & Sami, M. (2014). A Review of Scripting Techniques Used in Automated Software Testing. International Journal of Advanced Computer Science and Applications, 5(1), 194–202. https://doi.org/10.14569/ijacsa.2014.050128

[17]. Banday, M. T., Qadri, J. A., & Shah, N. A. (2009). Study of Botnets and their threats to Internet Security. Working Papers on Information Systems, January 2009.

[18]. M. Monperrus, "Explainable Software Bot Contributions: Case Study of Automated Bug Fixes," 2019 IEEE/ACM 1st International Workshop on Bots in Software Engineering (BotSE), Montreal, QC, Canada, 2019, pp. 12-15, doi: 10.1109/BotSE.2019.00010.

[19]. "ISO/IEC 27001: Information technology - Security techniques -Information security management systems – Requirements," 2005.

[20]. R. Montesino and S. Fenz, "Automation Possibilities in Information Security Management," 2011 European Intelligence and Security Informatics Conference, Athens, 2011, pp. 259-262, doi: 10.1109/EISIC.2011.39.

[21]. S. Fenz and A. Ekelhart, "Formalizing information security knowledge," Proceedings of the 4th international Symposium on information, Computer, and Communications Security, 2009, p. 183–194.

[22]. S. Radack and R. Kuhn, "Managing Security: The Security Content Automation Protocol," in IT Professional, vol. 13, no. 1, pp. 9-11, Jan.-Feb. 2011, doi: 10.1109/MITP.2011.11.

[23]. Aslam, M., Gehrmann, C., & Björkman, M. (2013, November). Continuous security evaluation and auditing of remote platforms by combining trusted computing and

security automation techniques. In Proceedings of the 6th International Conference on Security of Information and Networks (pp. 136-143).

[24]. Sternberg, R. J. (2012). Intelligence. Dialogues in Clinical Neuroscience, 14(1), 19–27. https://doi.org/10.7551/mitpress/9780262062749.003.0018

[25]. Duch, W. (2007). What is computational intelligence and where is it going? Studies in Computational Intelligence, 63, 1–13. https://doi.org/10.1007/978-3-540-71984-7_1

[26]. Neapolitan, R. E., & Jiang, X. Artificial intelligence : with an introduction to machine learning. Second edition.

[27]. Flasiński, M. (2016). Introduction to Artificial Intelligence. Springer International Publishing. https://doi.org/10.1007/978-3-319-40022-8_2

[28]. Tutorialspoint. (n.d.). Artificial Intelligence - Overview - Tutorialspoint. Www.Tutorialspoint.Com. Retrieved January 20, 2021, from https://www.tutorialspoint.com/artificial_intelligence/artificial_intelligence_overview.htm

[29]. Saleh, Ziyad. (2019). Artificial Intelligence Definition, Ethics and Standards.

[30]. Goertzel, B. (2014). Artificial General Intelligence: Concept, State of the Art, and Future Prospects. Journal of Artificial General Intelligence, 5(1), 1–48. https://doi.org/10.2478/jagi-2014-0001

[31]. Bostrom, N. (2014). Superintelligence: Paths, Dangers, Strategies (Illustrated ed.). Oxford University Press.

[32]. Dönmez, P. (2013). Introduction to Machine Learning, 2nd ed., by Ethem Alpaydın. Cambridge, MA: The MIT Press2010. ISBN: 978-0-262-01243-0. $54/£ 39.95 + 584 pages. Natural Language Engineering, 19(2), 285–288. https://doi.org/10.1017/s1351324912000290

[33]. Nilsson, N. J. (2005). INTRODUCTION TO MACHINE LEARNING AN EARLY DRAFT OF A PROPOSED TEXTBOOK Department of Computer Science. Machine Learning, 56(2), 387–399. http://www.ncbi.nlm.nih.gov/pubmed/21172442

[34]. Smola, A., & Vishwanathan, S. V. N. (2008). Introduction to machine learning. Cambridge University, UK, 32(34), 2008.

[35]. S. Ray, "A Quick Review of Machine Learning Algorithms," 2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon), Faridabad, India, 2019, pp. 35-39, doi: 10.1109/COMITCon.2019.8862451.

[36]. Mohammed, M., Khan, M. B., & Bashier, E. B. M. (2016). Machine Learning. Amsterdam University Press.

[37]. Weiss, S. M., & Indurkhya, N. (1995). Rule-based machine learning methods for functional prediction. Journal of Artificial Intelligence Research, 3, 383-403.

[38]. Omer, O. (n.d.). Introduction to Machine Learning The Wikipedia Guide. Https://Www.Academia.Edu.https://www.academia.edu/41157657/Introduction_to_Machine_Learning_The_Wikipedia_Guide

[39]. Bonaccorso, G. (2018). Mastering Machine Learning Algorithms: Expert techniques to implement popular machine learning algorithms and fine-tune your models. Packt Publishing.

[40]. Introduction To Genetic Algorithms In Machine Learning. (2021, January 18). Software Testing Help. https://www.softwaretestinghelp.com/genetic-algorithms-in-ml/

[41]. Dehghantanha, A., Conti, M., & Dargahi, T. (2018). Cyber Threat Intelligence. Springer Publishing.

[42]. Intelligent Security Automation Introduction. (Miller, 2018), https://www.blackhat.com/docs/webcast/2018-08-23-intelligent-security-automation-by-ty-miller.pdf

[43]. It’s Not Digital Transformation; It’s “Intelligence Transformation” We Seek. (2018). Data Science Central. https://www.datasciencecentral.com/profiles/blogs/it-s-not-digital-transformation-it-s-intelligence-transformatio-1

[44]. ThreatConnect. (2015). Threat Intelligence Platforms. Everything You’ve Ever Wanted to Know But Didn’t Know to Ask. 1–52. http://cdn2.hubspot.net/hubfs/454298/ebook/Threat-Intel-Platform-ebook-ThreatConnect.pdf

[45]. Islam, C., Babar, M. A., & Nepal, S. (2019). A Multi-Vocal Review of Security Orchestration. ACM Computing Surveys, 52(2), 1–45. https://doi.org/10.1145/3305268

[46]. What is orchestration? (n.d.). Www.Redhat.Com. https://www.redhat.com/en/topics/automation/what-is-orchestration.

[47]. Erl, T. (2005). Service-Oriented Architecture (SOA): Concepts, Technology, and Design. Prentice Hall.

[48]. DFLabs. (2021, January 18). Security Automation vs Security Orchestration – What's the Difference? https://www.dflabs.com/resources/blog/security-automation-vs-security-orchestration-whats-the-difference/

[49]. Understanding DNS Sinkholes - A weapon against malware. (2020, October 30). Infosec Resources. https://resources.infosecinstitute.com/topic/dns-sinkhole/

[50]. Kim H., Choi SS., Song J. (2013) A Methodology for Multipurpose DNS Sinkhole Analyzing Double Bounce Emails. In: Lee M., Hirose A., Hou ZG., Kil R.M. (eds) Neural Information Processing. ICONIP 2013. Lecture Notes in Computer Science, vol 8226. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-42054-2_76

[51]. Keary, T. (2020, September 15). 17 Best Syslog Servers (System Logging) tools for Windows and Linux. ITPRC. https://www.itprc.com/best-syslog-servers/

[52]. What is Incident Response? (2018, September 7). Digital Guardian. https://digitalguardian.com/blog/what-incident-response

[53]. S. (2019, July 10). Automated Incident Response - How Enterprises Benefit from it? Siemplify. https://www.siemplify.co/blog/how-enterprises-benefit-from-automated-incident-response/

[54]. Kral, P. (2011). Information Security Reading Room Incident Handler' s Handbook. The SANS Institute.

[55]. Penetration Testing. (2018, July 31). U.S. Department of the Interior. https://www.doi.gov/ocio/customers/penetration-testing

[56]. Funk, M. (2019, March 19). Web Application Penetration Testing Checklist (* New* Updated 2019). Cybers Guards. https://cybersguards.com/web-application-penetration-testing-checklist-updated-2019/

[57]. Abu-Dabaseh, Farah & Alshammari, Esraa. (2018). Automated Penetration Testing: An Overview. 121-129. 10.5121/csit.2018.80610.

[58]. K. P. Haubris and J. J. Pauli, "Improving the Efficiency and Effectiveness of Penetration Test Automation," 2013 10th International Conference on Information Technology: New Generations, Las Vegas, NV, 2013, pp. 387-391.doi: 10.1109/ITNG.2013.135

[59]. Chebbi, C. (2018). Advanced Infrastructure Penetration Testing: Defend your systems from methodized and proficient attackers. Packt Publishing.

[60]. Y. Stefinko, A. Piskozub and R. Banakh, "Manual and automated penetration testing. Benefits and drawbacks. Modern tendency," 2016 13th International Conference on Modern Problems of Radio Engineering, Telecommunications and Computer Science (TCSET), Lviv, 2016, pp. 488-491, doi: 10.1109/TCSET.2016.7452095.

[61]. Lobo, S. (2018, September 20). 6 artificial intelligence cybersecurity tools you need to know. Packt Hub. https://hub.packtpub.com/6-artificial-intelligence-cybersecurity-tools-you-need-to-know/

[62]. Nica, C., & Tănase, T. (2020). Using Weaponized Machine Learning in Cyber Offensive Operations. International Conference KNOWLEDGE-BASED ORGANIZATION, 26(1), 94–99. https://doi.org/10.2478/kbo-2020-0014

[63]. Buchanan, B., Bansemer, J., Cary, D., Lucas, J., & Musser, M. (2020). Automating Cyber Attacks. Center for Security and Emerging Technology, 13–32. https://doi.org/10.51593/2020ca002

[64]. Truong, T. C., Diep, Q. B., & Zelinka, I. (2020). Artificial Intelligence in the Cyber Domain: Offense and Defense. Symmetry, 12(3), 410. https://doi.org/10.3390/sym12030410

[65]. Rege, M., & Mbah, R. B. K. (2018). Machine learning for cyber defense and attack. Data Analytics 2018, 83.

[66]. Enoch, S. Y., Huang, Z., Moon, C. Y., Lee, D., Ahn, M. K., & Kim, D. S. (2020). HARMer: Cyber-Attacks Automation and Evaluation. IEEE Access, 8, 129397–129414. https://doi.org/10.1109/access.2020.3009748

[67]. Gadhgadhi, R., Nguyen, K.-K., & Cheriet, M. (2012). Automated intrusion attack with permanent control: Analysis and countermeasures. 2012 11th International Conference on Information Science, Signal Processing and Their Applications (ISSPA), 1440–1441. https://doi.org/10.1109/isspa.2012.6310530

[68]. Hassan, N. A. (2019). Ransomware Revealed: A Beginner's Guide to Protecting and Recovering from Ransomware Attacks (1st ed.). Apress.

[69]. Kevin Savage, Coogan, P., & Lau, H. (2015). Symantec SECURITY RESPONSE The evolution of ransomware. 57. http://www.symantec.com/content/en/us/enterprise/media/security_response/whitepapers/the-evolution-of-ransomware.pdf

[70]. FireEye Inc. (2018). How Government Agencies are Facing Cyber Security Challenges. https://www.fireeye.com/content/dam/fireeye-www/solutions/pdfs/how-govt-agencies-are-facing-cyber-security-challenges.pdf

[71]. Talamantes, J. (n.d.). What is Red Teaming and Why Do I Need It? Redteamsecure. https://www.redteamsecure.com/blog/what-is-red-teaming-and-why-do-i-need-it-2

[72]. Red Team Improving Organizational Resilience [online]. www.deloitte.com.

[73]. Deo, P., Raj, G., Perumal, R., & Subramoni, S. COVID-19 Impact on Cyber Security & Ways to Confront the Risk | TCS. www.tcs.com. https://www.tcs.com/perspectives/articles/how-covid-19-is-dramatically-changing-cybersecurity

[74]. Gouget, A. (2020, August 5). Quantum Computing and the evolving cybersecurity threat. Security Boulevard. https://securityboulevard.com/2020/08/quantum-computing-and-the-evolving-cybersecurity-threat/