Autonomous Penetration Testing in Cybersecurity

Course: Topics in Cybersecurity

CS 490DJ

Project Research Paper

Dec 6, 2024

Prepared by:

Devansh Someshwar - 200455072,

Nisarg Dolasiya - 200497867,

Waseera Ishtiaq - 200488856

Table of Contents

Introduction

In a world where technology is steadily growing and developing, the concept of cybersecurity has not remained simple as it used to be. In their article published on the American Public University System's website, "What is Cybersecurity? The Realities of the Digital Age," the author explains several factors that have led to this increase in complexity.

Malicious attacks and the actors who conduct them have also been on the rise because of technological advancement. It is important to note that cybercriminals are constantly coming up with new strategies on how to penetrate through security systems and some of the techniques that they use are ransomware, phishing and zero-day vulnerabilities which are hard to avoid. These attackers can be part of sophisticated and well-resourced groups that are able to plan and execute sophisticated attacks on specific systems. One more contributing element is the increasing attack surface caused by the emergence of the Internet of Things (IOT). Every new device connected to the network can potentially increase the threat if it is not properly secured, putting further pressure on the cybersecurity experts.

Technological advancements in areas that are currently emerging and are set to emerge in the future such as cloud computing, artificial intelligence, and machine learning present new security risks that must be addressed. Although these technologies are highly beneficial, they also lead to new vulnerabilities that emerge as threats to organizational security. The rate at which data is being produced by individuals and organizations also contributes to the difficulties being experienced when attempting to protect sensitive data. Securing and protecting data becomes a challenge especially when dealing with large volumes of data moving across different platforms hence the need for efficient security solutions.

Global connectivity has also raised the level of complexity of cybersecurity. This is because the current networks are interconnected making it quite easy for a vulnerability in one system to spread to other systems within the network and vice versa hence posing a risk of spreading cyber threats across different networks and impacting numerous organizations and industries. This interconnectedness demands a systematic approach to security that looks at the bigger picture of the network. Another challenge that is evident is regulatory compliance. It is important for organizations to operate within the boundaries of numerous cybersecurity laws and guidelines including the GDPR, HIPAA, and the PCI DSS. This makes it imperative to constantly monitor and adjust to the ever-shifting legal requirements which can be quite cumbersome.

People are still the "weak link" in the cybersecurity chain. This is because social engineering attacks aim at preying on people's character to compass gains that they would not otherwise have into systems and this is why awareness and training of the employees is considered as fundamental to any organization's security measures.

All these factors therefore imply that cybersecurity is no longer a stagnant field but rather a dynamic and complicated process that has to be well handled. Old security measures are ineffective given the current threat landscape and there is a clear requirement for new approaches.

In view of this, penetration testing or commonly referred to as pentesting comes into play as an important technique to identify and mitigate vulnerabilities found on the network. Pentesting is to find out the vulnerabilities by launching cyber-attacks on a targeted computer, network or a web application. This process is performed by ethical hackers and the same techniques and tools are used by the ethical hackers to break into an organization, yet the intention is not to steal but

to make the organization’s security stronger. Penetration testing is not a one time activity. Instead, it is a process that an organization must undertake regularly. The frequency of the tests depends on risk assessments and the organizational structure of the company.

Pentesting is a powerful technique that helps in identifying the vulnerabilities before they are exploited by the attackers. It aids in alerting the organisation about the weaknesses in its security posture that may not even get detected through regular security audits and automated alerts. It offers a systematic assessment of the current security mechanisms to check if they can hold off the recent emerging threats. In particular, penetration testing is critical in avoiding data breaches that cause financial loss, damage to the image of the organization, and legal penalties. Thus, identifying vulnerabilities will reduce the chances of intrusions into sensitive information systems. Pentesting is also crucial for organizations because many industries have legal requirements that stipulate that security audits have to be conducted frequently and one of them is the Payment Card Industry Data Security Standard (PCI DSS) and the Health Insurance Portability and Accountability Act (HIPAA). Also, the information collected from penetration testing improves an organization's incident response where they can stimulate their systems with real world attacks and test their response procedures. This helps in identifying the areas of improvement within their network.

In conclusion, as the threat landscape expands and the nature of threats becomes more complex, penetration testing has emerged as an important security measure. It allows to identify and remediate vulnerabilities thus enhancing an organization's security posture and preserving effectiveness of intellectual property and other valuable data.

Background: The Evolution of Pentesting

Penetration testing as a field has its roots in the early years of interconnected computing systems where specific groups that were called as “tiger teams” performed the task of challenging the security mechanisms of government and 1970s military and networks. These were very early unstructured tests and very heavily dependent on humans’ expertise and creativity. As described by Infosec Institute and Christian Espinosa, the U. S. Department of Defense and other organizations such as RAND Corporation played an instrumental role in defining what would be termed as penetration testing engagement. These early tests were meant to identify weaknesses that could have been exploited by the adversaries, thus laying the groundwork for what is today referred to as ethical hacking much before the term gained popularity.

In the 1980s and the early 1990s, penetration testing grew out of the government and military domains as a function of technology advancement. The development of the internet, the emergence of sector e-commerce as and the security growth measure. Early practitioners relied on a combination of manual techniques, custom scripts, and emerging security tools. Over time, guidelines and the generally accepted best practices were defined and by the early 2000s, the frameworks and certifications from EC-Council and others became available. These defined approaches provided for consistency, repeatability and adherence to a well-defined set of phases and marked the beginning of modern-day penetration testing.

Today’s pen testing is much different from the early manual form of ethical hacking that was seen in the past. Currently, security professionals employ a number of effective automated scanning tools, vulnerability management platforms as well as exploitation frameworks that

enhance the discovery and testing processes. There is an increasing use of machine learning and artificial intelligence-based solutions to assist the testers in analyzing large data sets and deal with hard problems involving context-specific security issues. It has to be noted that despite the overall trend towards automation and use of tools, the human factor is still important. The experienced pen tester's ability to craft chains of vulnerabilities, to understand the nuances of system responses, and to think like a malicious attacker makes today's pen testing an effective balance of technology and experience.

The Traditional Pen testing Workflow

The basic pen testing workflow has not changed much and is still divided into several well-defined stages even though the technologies and methods used have evolved significantly. Resources from the blog "Understanding the Five Phases of the Penetration" on EC-Council outline these phases as follows:

1. Reconnaissance:

In this initial phase, testers try to obtain as much information as possible about the target system or network. This can include passive information gathering from the public domain, identification of domain registrations, social media profile, technical documents etc. The objective is to identify the target's environment, the technologies that are used in it, the possible points of entry, and the overall security status of the target.

2. Scanning:

The next phase is the active scanning after the reconnaissance. This involves the use of tools in identifying open ports, the running services, the network architecture and the system details. Various scanning tools and techniques enable the identification of weaknesses, the assessment of the attack surface and the identification of services or applications that may be vulnerable to attacks.

3. Vulnerability Assessment:

The third penetration testing phase is vulnerability assessment, in which the tester uses all the data gathered in the reconnaissance and scanning phases to identify potential vulnerabilities and to check how they can be exploited. Like scanning, vulnerability assessment can be effective on its own, but it is even more so when integrated with the other penetration testing phases. When assessing the risk of the vulnerabilities identified in this phase, penetration testers can turn to numerous sources. One such is the National Vulnerability Database (NVD), which is a database of vulnerability management data that is sponsored and compiled by the US government and based on the CVE database that lists the common software vulnerabilities. The NVD assigns scores to the known vulnerabilities based on the impact and the exploitability of the vulnerability with the help of the Common Vulnerability Scoring System (CVSS).

4. Exploitation:

With the information gathered from the previous steps, the testers try to exploit the vulnerabilities or the misconfiguration of the system. They may employ Borgs such as Metasploit or employ their own handcrafted methods to achieve goals that include gaining

access, escalating privileges and/or extracting key information and this is done in a way that resembles a real attacker but in a controlled environment.

5. Reporting:

After successfully demonstrating vulnerabilities, testers methodically document their work including the process, tools employed, and the effect of each vulnerability discovered. A long report outlines the work done, the risks found, and the solutions to the risks in a detailed manner. During some of the engagements, the process may also include maintaining access or testing for persistence and/or cleaning up all test objects to ensure that all systems are restored to their pre-test conditions.

These phases have stayed as the foundation of reliable penetration testing even as it has evolved over time. State of the art practices include the use of automated scanning, continuous testing, Artificial intelligence and data analysis and all these build on this traditional approach. This systematic and phased approach not only guarantees that no ground is left uncovered but also adheres to best practices and industry standards that are helpful in a logical approach that can be followed by the testers as well as the clients to know the scope, process and findings of a penetration test.

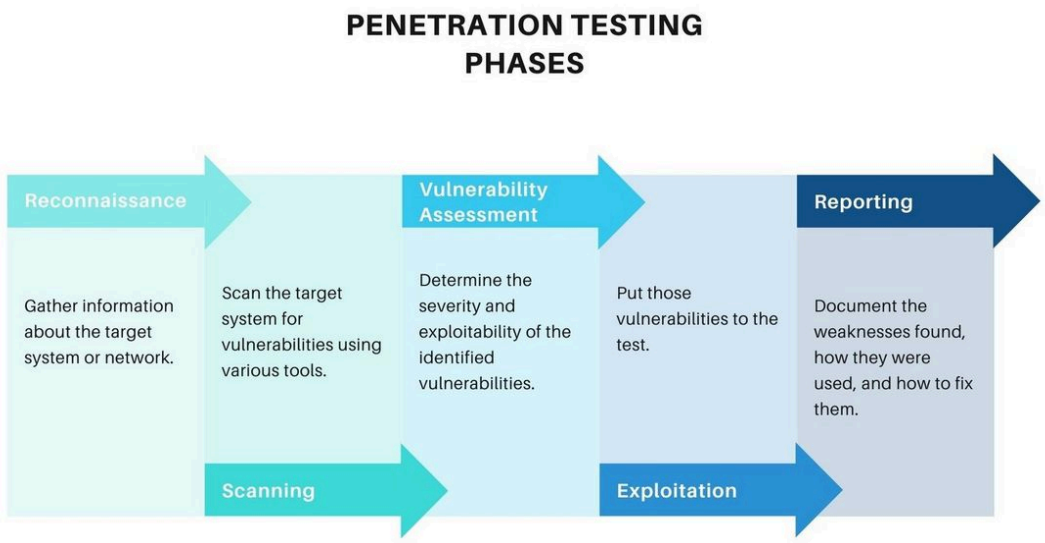


Figure 1: Penetration Testing Phases. Source: EPAM SolutionsHub. (2024). Uncovering the Invisible: A Guide to Essential Types of Penetration Testing, from [*https://solutionshub.epam.com/blog/post/types-of-penetration-testing*](https://solutionshub.epam.com/blog/post/types-of-penetration-testing)

There are several tools that are currently used in the penetration testing process and they are used in the different aspects of the assessment. Metasploit, Nmap, Nessus, and Burp Suite are some of the most often mentioned tools in the context of basic as well as advanced pen testing tool sets. Security professionals from Medium's cybersecurity community agree that these tools are indispensable in today's security landscape.

Challenges: Manual pentesting is slow, resource-heavy, and cannot scale to modern infrastructures

Penetration testing is a process which is performed by security experts, manually, who try to identify the weaknesses in systems, applications and networks as if it is a real attack. While the automated tools use a set of rules and scripts to penetrate through the defenses, the manual testers use their experience, brain and flexibility to find the loopholes that may have been overlooked by the former.

Nevertheless, this depth comes with significant drawbacks as highlighted in the article published on Yellow Systems' blog, "Automated vs. Manual Penetration Testing: Which One Is Better for Your Business?" Manual testing is a time consuming process, it is very costly and it cannot keep up with the changes in the infrastructure since there are so many assets to be tested (Yellow Systems, 2021). This is where manual penetration testing faces major obstacles in trying to meet the complexities of the current infrastructures. The main disadvantage of manual pentesting is the labour-intensive and time-consuming process that it is. Security specialists with the required skills have to dissect systems, a process that could take weeks or even months. This is due to the fact that there are several steps that have to be carried out in the process of identifying vulnerabilities in large complex networks when conducting manual penetration testing. Furthermore, it is difficult for manual pentesting to meet the requirement of scalable testing for the large and dynamic modern IT environments. Large organizations have large networks of devices, applications, and services that are constantly in development.

According to a blog "Manual Penetration Testing vs Automated Penetration Testing – Which One is Better?" written by *Igor Kantor* on Iterasec, such environments are complex and large,

and it is impossible to perform a complete test of all the vulnerabilities using a manual approach only. The number of assets to be assessed also exceeds the capabilities of manual methods, thus creating security risks. It also points out that while manual tests are capable of revealing the fine details of an organization, their inefficiency in adapting to the dynamic nature of today's IT environments means that there is still a need for more effective strategies (Iterasec, 2024). In the long run, it is seen that although manual penetration testing can provide detailed information regarding the security of an organization, these constraints strengthen the case for incorporating more automated and dynamic strategies to deal with the contemporary threats.

Another important disadvantage of manual pentesting is the resource dependence. It employs expert professionals who possess the knowledge of the current tactics and strategies of penetration testing. It is rather difficult to find and keep such specialists, which is why the companies have to pay for their services. Moreover, the conventional testing involves the use of tools, which can be quite costly and also consume a lot of time and resources, which may be constrained for many organizations particularly those with limited cybersecurity budgets. An article written by *James Baucom* published on Tangible Security states that the use of experienced personnel and time is a major challenge in conducting manual penetration testing as it is very expensive and cannot be easily extended (Tangible Security, 2024).

Gap: Existing tools lack sufficient automation and fail to prioritize vulnerabilities effectively

While the use of automated penetration testing solutions can help in reducing the time and effort required for manual assessments, most of the current tools are still not automated enough to meet the current complex environment needs. Many of these tools rely on pattern matching to identify known vulnerabilities but they cannot adapt to the continuously changing environment or the specific characteristics of any given network. Hence, large data sets are produced, and security teams have to go through long lists of outcomes instead of getting a list of vital and prioritised issues. This approach however can hide the most critical vulnerabilities thus giving organizations a wrong impression of where to direct their already scarce resources.

In addition, as noted by blogs published on Yellow Systems and Tangible Security, while the automated tools are better than the manual based approaches, they are not able to identify the criticality of the vulnerabilities properly, and hence security teams have to make heuristics decisions about which of the issues need to be fixed first. Without the ability to rank the vulnerabilities according to their potential to be exploited and the potential damage that could be caused by such exploitation, such tools only increase the complexity. Iterasec's article insights, highlighting the fact that even when automation is introduced into the process, it has to come with a better, contextual understanding of the problems to be solved and risks to be mitigated, to prevent overwhelming the team with noise and low-risk findings and this can be achieved through incorporating Artificial Intelligence. Thus, organizations should be able to convert a raw vulnerability feed into a meaningful, actionable roadmap when they incorporate advanced analytics, artificial intelligence, and risk-based prioritization into their workflows.

Together with the limitations of manual penetration testing discussed above, these shortcomings of current automated tools prove that there is a need for a more balanced, informed, and dynamic strategy. Organisations are now in a place where they need solutions that can combine the expertise of humans with the automation process of identifying, contextualizing and managing the most vital security risks in the dynamic environments of the modern digital world. In this regard, the application of artificial intelligence to these tools can be of great benefit to organizations in as far as the identification of risks, vulnerability prioritization, exploitation, and reporting.

Objective: How AI can automate pentesting tasks

The main goal of this project is to determine how AI can be utilized in a way that goes beyond the standard automation and improves the effectiveness of pen testing. Conventional manual pen testing although more comprehensive is very time consuming, labor intensive and inefficient for large networks as today. The current automated solutions help to some extent with increasing the efficiency, but they cannot easily change to meet the new threats and cannot always determine the most critical vulnerabilities that should be fixed first, depending on the organization's security risk profile. In this paper, we want to show that the integration of AI can enhance the accuracy, scalability, and operational efficiency of pen testing, and at the same time, reduce the manpower needed to address modern day cyber threats. According to a blog by Anupama Mukherjee on ThreatIntelligence.com, automation only performs routine tasks well, while the AI-based approaches are capable of learning from the data, identifying the new patterns of the attack, and adjusting their strategy in the real-time to reveal the vulnerabilities that have not been found before.

The pen testing solutions that incorporate AI, as explained by the article "The Impact of AI on Penetration Testing " on GetOppos, employ machine learning and data analytics to recognize, assess and even forecast the possible tactics that a malicious hacker may employ in the process of penetration testing than a team of human testers would take. Instead of using only the signatures that have been predefined or scanning for the vulnerabilities that are well known, these tools are able to learn from the environment and adjust the strategies used in the process of simulating advanced attacks in real time. Jamison Utter's posts on LinkedIn demonstrates how generative AI models can not only identify the flaws, but also have capabilities to identify new attack scenarios. Such techniques eases the burden of time-consuming recon, effectively

identifying the most vital problems, and regularly updating the approach during operations as the environment and threats change.

The work of the pen testers can be realigned from the routine tasks and basic level scanning to more advanced level tasks including the analysis of complex vulnerabilities and the optimization of incident response plans, effectively increasing the output of security teams. In addition, the increased efficiency enables the conduct of assessments more frequently thus ensuring that organizations have a constant view of their security posture instead of having to wait for a one-time or periodic assessment.

In conclusion, integrating AI into the pentesting process increases the precision in vulnerability detection, decreases time to identify vital weaknesses, adapt to the size and complexity of networks, and relieve the shortages of qualified resources while at the same time ensuring that no aspect of the assessment is overlooked.

In the modern world where digital processes are involved in managing almost on every organizations business has process, increased the significantly pressure regarding the frequency and complexity of cyber threats. Cyber criminals are always a step ahead in their planning and strategies, leveraging modern complex structures, diverse networks, and regularly deployed technologies. While companies are increasing their online presence, they are not only exposing themselves to a larger number of risk points for the attackers but also to increasing the number of compliance and customer demand for data privacy and security.

These challenges are further exacerbated by the well-documented shortage of skilled cybersecurity personnel. It is thus evident that the demand for the experienced security personnel is high while the supply is low hence it is difficult to attract, develop and retain the required

talent to meet the ever evolving threats. It can be said that security teams are overwhelmed by the vast number of alerts, the need to analyze vast amounts of vulnerability data, and still little ability to determine which of these threats are most likely to materialize.

This is where automation, and more importantly, artificial intelligence-driven automation, step in as a crucial solution. Instead of relying on the labor-intensive, and often inefficient, manual processes or basic tools that produce large quantities of data with no clear structure and priority, real intelligence is created from raw data. With the ability to contextualise and rank vulnerabilities based on priorities, AI-enabled pentesting tools help companies to manage their security risks in a more efficient manner. Security teams can thus deal with the threats that appear in the course of operation, effectively mitigate the risks that pose the biggest threat to the organisation, and improve the overall security posture. In other words, automation is not anymore just a tool to increase productivity, it is a vital approach that enables organisations to compete in an environment of increased cyber risk and restricted staff.

Role of AI in Cybersecurity

AI has been very effective in enhancing different aspects of the cybersecurity domain through automating complex analyses, detecting patterns a human analyst might miss, and handling the dynamics of threats at faster rates. For example, in the case of malware, AI-based systems use machine learning algorithms that have been trained using a large number of malicious files. This makes them able to identify new strains of malware by picking out the differences in code signatures, behaviors or patterns. Consequently, the security teams can shift from the reactive model of patching when a threat is well known to proactive defense, where the threats are identified before they cause harm.

In the same manner, threat intelligence has been changed by the capabilities of AI to filter through large sets of security data from different sources. The article “What Is the Role of AI in Threat Detection?” published on Palo Alto Networks states that AI-based threat detection consolidates logs, network traffic, and endpoint telemetry data with the ability to identify anomalies, patterns and pre-emptive attacks. AI can link IOCs with known adversaries' TTPs and present the findings in a prioritized manner much faster and with less chances of error. This enables organizations to effectively predict the actions of the adversaries, optimize their defense posture, and reallocate their security resources in a dynamic manner.

In the area of intrusion detection, the AI models are especially good at identifying the differences from the set of normal network activity or patterns. This is contrary to the traditional rule-based IDS that only works with known attack signatures and may not be effective against unknown or rapidly evolving threats. By using AI, these systems can constantly determine what is considered normal for a particular environment, and any deviations will be raised as potential issues in real

time. This reduces false positives and at the same time facilitates faster response times. Consequently, the modern intrusion detection systems that incorporate AI are not only able to identify known threats more efficiently but also zero-day threats and highly advanced, stealthy malware.

However, AI has the potential to change the face of penetration testing. The process of identifying gaps in an organisation's security structure and identifying vulnerabilities aids in evolving the traditional time-consuming tasks like manual reconnaissance and scanning. As pointed out by Onkar Mhaskar on LinkedIn when discussing AI in cybersecurity, including AI into the pen testing process enables fast assessment of system configuration, application logic, and network topology. Such anomalies may include identification of unusual patterns or configuration that could be deemed as hidden vulnerabilities or misconfigurations which can be efficiently directed to the most critical components by the human analysts.

In general, AI enhances both the preventive as well as the detective measures of cybersecurity. They make malware detection, threat intelligence and intrusion detection from being static and reactive to dynamic and adaptive. As the technology advances, it will also enhance penetration testing to perform routine tasks faster, to identify complacent attack vectors and to help security teams to keep up with the ever evolving cyber threats.

However, these tools also have the following important common drawbacks. They do not provide true end to end automation for all the activities involved in the pen test. Although each has its strength at a certain phase of the penetration test including the scanning, exploitation or vulnerability detection there is no tool that can perform reconnaissance, analysis, exploitation and reporting in a continuous manner as an integrated platform with artificial intelligence.

Furthermore, very little AI/ML integration coupled with static rules means that these tools can only detect known issues or act in a predefined manner. They cannot change their approach based on the dynamic environment, cannot rank vulnerabilities according to the rapidly changing threat landscape, or improve their performance with the help of previous engagements' data without the help of many humans. This is where the role of Artificial intelligence comes into play to leverage these tools.

Significance of AI in Penetration Testing

In the modern world where digital processes are involved in managing almost on every organizations business has process, increased the significantly pressure regarding the frequency and complexity of cyber threats. Cyber criminals are always a step ahead in their planning and strategies, leveraging modern complex structures, diverse networks, and regularly deployed technologies. While companies are increasing their online presence, they are not only exposing themselves to a larger number of risk points for the attackers but also to increasing the number of compliance and customer demand for data privacy and security.

These challenges are further exacerbated by the well-documented shortage of skilled cybersecurity personnel. It is thus evident that the demand for the experienced security personnel is high while the supply is low hence it is difficult to attract, develop and retain the required talent to meet the ever evolving threats. It can be said that security teams are overwhelmed by the vast number of alerts, the need to analyze vast amounts of vulnerability data, and still little ability to determine which of these threats are most likely to materialize.

This is where automation, and more importantly, artificial intelligence-driven automation, step in as a crucial solution. Instead of relying on the labor-intensive, and often inefficient, manual processes or basic tools that produce large quantities of data with no clear structure and priority, real intelligence is created from raw data. With the ability to contextualise and rank vulnerabilities based on priorities, AI-enabled pentesting tools help companies to manage their security risks in a more efficient manner. Security teams can thus deal with the threats that appear in the course of operation, effectively mitigate the risks that pose the biggest threat to the organisation, and improve the overall security posture. In other words, automation is not anymore just a tool to increase productivity, it is a vital approach that enables organisations to compete in an environment of increased cyber risk and restricted staff.

Why AI-Driven Pentesting?

As the organizations adopt the cloud computing, Internet of Things (IoT) and other technologies, their attack vectors are also increasing at a rapid pace. The conventional penetration testing approaches are mainly dependent on manual labor and cannot be matched with the current rate of expansion. IoT devices bring in billions of new endpoints that need to be evaluated while cloud environments are agile and continuously dynamically change their configuration, services and applications. This complexity can even challenge experienced security teams as they work to understand, prioritise and address vulnerabilities throughout large numbers of systems.

At the same time, the cybersecurity industry still has a deficit of qualified experts. When workload increases and the environment becomes more threats prone, many companies have no choice but to perform security assessment infrequently and with less detail than is needed. This is

where AI-based pen testing tools come in, they take over the time-consuming parts of the process that were previously done by humans. Instead of assigning the tasks of system mapping, identification of weaknesses and simulate the attacks by human testers, these smart platforms use machine learning and data analysis to perform fast and efficient scans of large number of systems, and adjust their scans according to the dynamic conditions to pinpoint the most crucial concerns that need to be addressed immediately.

Thus, integrating AI into pen testing workflows, organizations are able to shift from the traditional, time-consuming, and reactive approach to security to the proactive and scalable ones. Keshav Malik blog “AI in Cybersecurity: Benefits and Challenges” published on Get Astra’s (2024) state that the application of machine learning improves the efficiency of security operations by effectively identifying trends, filtering out the irrelevant and efficiently pinpointing threats that pose a severe risk before the adversaries’ act on them. Likewise, David Lefever blog “AI Pen Testing: Is Automated Penetration Testing Right for You?” on Centric Consulting (2024) argues that pen testing approaches enhanced with AI can offer continuous assurance as opposed to conventional methods that are slow to catch up with the technological changes. Onkar Mhaskar’s article on AI in penetration testing also supports this concept where he explains that generative AI models can create new possible attack vectors and detect variations that other tools might have failed to identify in the past. Thus, helping organizations save time and effort by focusing on the key issues that require their attention and assistance of their limited human resources.

Hence, AI-based pen testing is suitable for meeting the main challenges of today's cybersecurity teams: increasing attack surface, the shortage of skilled workers, and the requirement for active and effective solutions. This paper demonstrates how AI can help in enhancing the pen testing process by performing or even assuming the critical part of the process and thus enabling the organization to effectively enhance their defense, navigate through the increasing complexity, and maintain a sound security posture in the dynamic environment.

The Solution part for the proposed problem

In this section, the research project's solution is proposed, with an emphasis on the development of a conceptual model for autonomous penetration testing. The solution is aimed to solve the current problem of manual penetration testing which is slow, resource heavy and less scalable. Pentesters do use automated tools and scripts to penetrate through the defenses, but not fully leverage the power of today's AI models.

The solution focuses on the key stages of pentesting:

- Reconnaissance & Scanning
- Vulnerability assessment
- Exploit simulation
- Reporting and Recommendation.

The designed solution is a modular framework that will overcome the inefficiencies of manual penetration testing by integration of artificial intelligence models into the workflow. The stages mentioned above use industry standard tools such as Nmap, Nessus, Burp suite, Metasploit and AI models for prioritization and scoring severity. This modular approach will fill the gaps of scalability, flexibility and adaptability in the current solution.

Background on Types of Machine Learning.

Classification model -

Classification refers to a supervised machine learning method in which the model categorizes the data into appropriate classes. In a supervised learning method, models learn by examples. Therefore, classification algorithms use input training data to predict how likely future data will

belong to one of the already specified classes.For example, dividing emails into different categories such as- spam, junk etc.

Regression model -

Regression model refers to a supervised machine learning method used to predict continuous data.The set of algorithms defined learns the relationship between independent variables and a dependent or resultant variable. A model can thus be used to make a prediction of the output for new data or missing values. For example - Predicting weather patterns.

Reinforcement Learning model -

While supervised learning learns from examples, reinforcement learning learns from experience. In other words these models are trained by trying different actions and based on outcome rewards and penalties are assigned.

Workflow Diagram

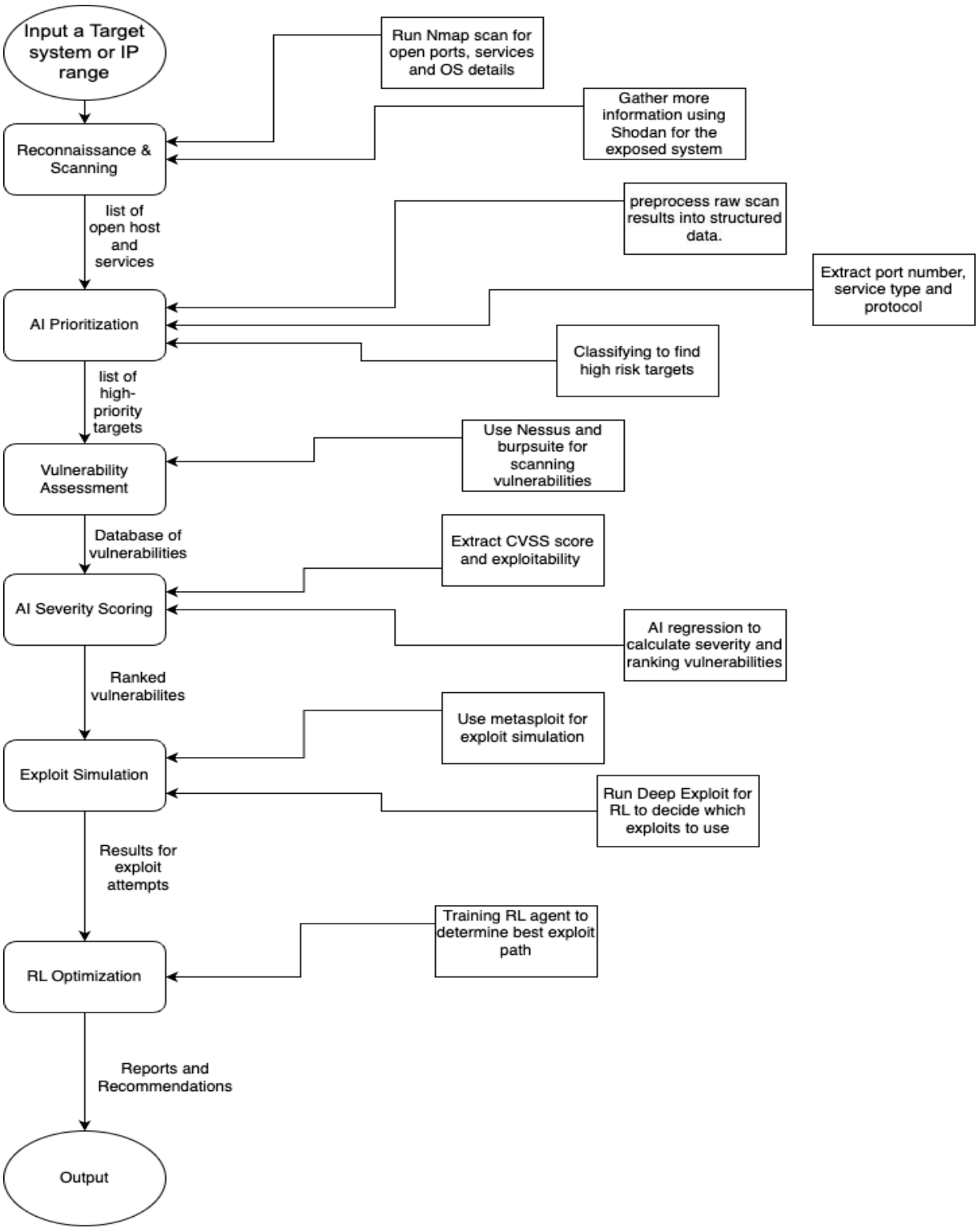


Figure 2 - Made with draw.io

Conceptual Solution and Workflow Steps

The section includes a conceptual solution and workflow setup which is designed as a modular system where each part of the solution operates independently and contributes to the overall automation of pentesting. At each stage of the solution, there is a combination of pentesting tools and machine learning models. The pentesting tools which are utilized to generate scan results or dataset which is fed into trained models for assessing the vulnerabilities, severity and exploitation paths.

1. Reconnaissance and Scanning

Reconnaissance and scanning is the initial step for penetration testing workflow. In this step, information is gathered about the target systems and identifying open ports, active hosts and services that are running. The data collected from scanning is further preprocessed and converted to structured data. These results are then used to find high risk targets for further analysis.

Tools:

- **Nmap (network mapper):** nmap is a tool which scans a network and identifies any open ports, services and details about the operating system.

  Example of a Scan:

- **Shodan API:** Shodan is a search engine which contains information about internet connected devices and the API is used here to get metadata about the target IP, some of which includes geolocation, vulnerabilities and configurations.

  Example of API query to public database:

AI integration:

Using AI enhances the effectiveness of the intel from the above scans and automatically finds high-risk targets.

Data Collection:

- Data collected from the scans is preprocessed into a structured format,

  a CSV file. Since, classification models use supervised learning.
- Information is extracted such as
  - Port numbers
  - Service types
  - Protocols

  along with the known vulnerabilities from shodan.

Here, a classification model is used which is already trained on historical vulnerability data. A classification model is a machine learning model that categorizes data into predefined classes, labelling them as High risk, Medium risk, low risk or no risk. This model can forecast future data and classify them.

Workflow Steps

1. Input - A target system Ip or a subnet.
2. Run Nmap scans on the target system and get information on open ports and services.
3. Use shodan to get metadata about the target IPs
4. Process the scan results from both sources and use the classification model to label targets based on risk level.

Example code:

```
# referred from https://www.datacamp.com/tutorial/random-forests-classifier-python

import nmap  #importing nmap python library
from sklearn.ensemble import RandomForestClassifier   #importing  randomForest machine learning algorithm.

#nmap scan
scanner = nmap.PortScanner()
target = "192.168.100.0/24"    #target system is a subnet of IP
scanner.scan(target, arguments="-sV")         # in this example, using -sV flag scan the verion.

#Training AI model to make predictions
features = [key_features]
labels = [0,1,2,3]
# label represents risk level
# 3 - High risk
# 2 - Med risk
# 1 - Low risk
# 0 - No risk

rf =  RandomForestClassifier()
rf.fit(features, labels)
hpt = rf.predict(features)
print("High-Risk Targets:", hpt)
```

Outcome

Outcome - High-priority targets which will be used as inputs for vulnerability assessment

2. Vulnerability Assessment

Vulnerability assessment is a phase where all the data gathered from reconnaissance and scanning phases is used to find potential vulnerabilities and to check how they can be exploited. The goal is to find security weaknesses of the high-priority targets. Pentesters use tools such as Nessus and Burpsuite inorder to find misconfigurations and CVE based vulnerabilities. Nessus is efficient in creating a baseline of vulnerabilities, although the user is left to manage the priority of the issues as well as the deeper exploitation testing. Thus, the use of a Regression model can help determine the severity of vulnerabilities and prioritize high risk targets.

Tools:

- **Nessus** is a vulnerability scanner that can effectively scan for security holes, mistakes, and policy breaches impact systems. Nessus' vulnerability database is constantly updated, making it possible to quickly identify the most critical issues.

- **Burp Suite** is used for performing web application security testing such as automated scanning, proxy interception, and identifying vulnerabilities like SQL injection, cross site scripting. Burp Suite has a scanner which is able to identify standard vulnerabilities for instance those based on common patterns, but lacks an AI-based system that is capable of modifying its test strategies based on the results of prior tests.

AI Integration:

The output scan results from nessus and burp suite are fed to a regression model, which is a type of supervised machine learning method. It can predict continuous data.

It takes in the input such as CVSS score and exploitability from the vulnerability database and its set algorithms can learn the relationship between independent and dependent variables in the provided dataset. The trained model can use its regression algorithm to predict future data or the values that are missing.

It scores the severity of the vulnerabilities and ranks them accordingly.

Workflow Steps

1. Input the data, list of high priority targets, collected from reconnaissance.
2. Use Nessus to scan and identify CVE based vulnerabilities.

3. Burp Suite identifying for finding application vulnerabilities.
4. Get the output from both the scans and feed it to the regression model.

Example Code -

```
#source https://www.geeksforgeeks.org/python-linear-regression-using-sklearn/

from sklearn.linear_model import LinearRegression #importing regression model
# Train Regression Model
key_features = [Vulnerability Data]
labels = [Severity scores]
LR = LinearRegression()
LR.fit(key_features, labels)

severity_pred = LR.predict(key_features) # Predict Severity
```

Outcome - AI severity scoring will output a list of ranked vulnerabilities.

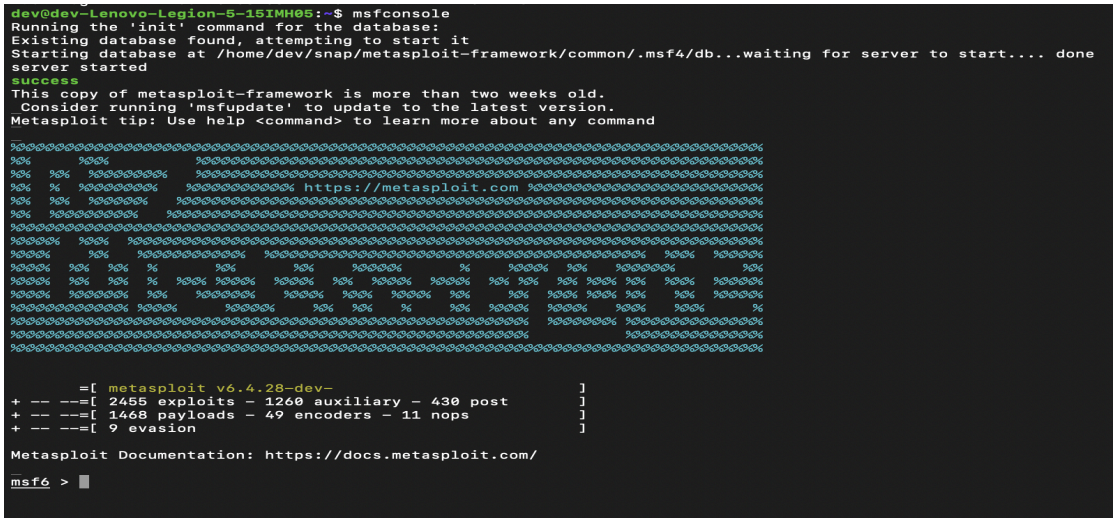
3. Exploit Simulation

The information gathered from vulnerability assessment is used to exploit the target. This phase uses the discovered vulnerabilities and automates the exploitation by using tools such as Metasploit and Deep exploit. Here incorporating a reinforcement learning model can dynamically identify an optimal path for an exploit to execute.

Tools

- **Metasploit** is an exploitation framework, and that offers a large database of exploits and payloads and as such, the penetration testers can emulate real life conditions with the aim of evaluating the robustness of a system. Although Metasploit can be used for translating the identified vulnerabilities into real attacks, the work will depend on the information provided by other scanning and enumeration tools. Thus, it does not entirely encompass the entire pen testing process.

  Metasploit terminal -



AI Integration:

Use of RL (reinforcement learning) agents, this model learns from past experience by trying different actions, based on outcome rewards. This becomes very useful during exploiting targets. It takes in data from the metasploit which runs an exploit simulation and gives reports on successful/failed attacks. The RL agent uses this information to find optimized exploitation paths for the attempts made. This feature becomes very useful during pentesting.

Workflow Steps

1. Input - list of vulnerabilities from previous phase.
2. Use metasploit to run the exploits
3. Outcome - Attempt was either successful or failed and what was the impact on the system.

Outcome - The reinforcement learning model will output the optimized exploitation path which will be used for generating a final report with recommendation.

4. Reporting and Recommendation:

Based on the output from RL Agent the automated report generation takes place. It converts the output data from RL optimization into recommendations and actionable insights. It utilizes advanced Natural language processing tools (NLP) such as GPT-3 and transformers to create reports.

These reports focus on 3 core components - an executive summary for most critical vulnerabilities, their impacts on the organizations and an action plan for immediate and future advancements.

In other words it translates complex technical findings into valuable insights.

5. Challenges and Proposed Improvements

A major problem with the integration of the new AI-driven capabilities into a tool like Nmap is that it is difficult to obtain a large enough dataset that is diverse enough. While in many cases of machine learning, data can be collected rather easily, collecting real world penetration testing data poses legal and ethical constraints as it is illegal to perform scans on external networks since such practices are prohibited by various laws including the US Computer Fraud and Abuse Act (CFAA), the Computer Misuse Act of the UK or any other legislation in other countries.

Synthetic data collection: The test-labs of organizations or emulated systems.. The outcome is a dataset which might not depict real-life situations, thus making the AI model less efficient and unable to generalize.

Also, developing the kind of extensive, varied data that is necessary for solid AI training is very expensive, in terms of time, effort and experience. Developing a training network that includes a variety of network configurations, active services and realistic vulnerabilities is not an easy task.

Security professionals have to periodically deploy, configure and set up test networks, set up new software versions, introduce new vulnerabilities and identify patterns. Also, it is important to mention that data labeling, identifying which scans are related to which vulnerabilities, the severity of the latter and their real-world exploitations is a time-consuming process that may require the input of security experts.

Proposed Improvements:

To improve the efficiency and the potential of AI-enabled penetration testing technologies for the future, the following recommendations are suggested:

1. Use Real-World Data from the Public Sources, for example, Shodan and Censys:

The integration of real-world data from sources such as Shodan and Censys can greatly improve the effectiveness of the AI-based pen testing tools. These platforms are always active in scanning and indexing of the internet connected devices which offer a wealth of information on network configuration, open ports and the services in operation. This real-life data helps the AI models to create realistic attack scenarios and to detect vulnerable services based on the real-world network footprint. As per the article “Next Level Reconnaissance with Shodan and Censys” on Breakpoint Labs, the easy and frequent access to various data sources make it easier for the AI to understand the current threat environment.

2. Use Anomaly Detection Models for Zero-Day Vulnerabilities:

The use of sophisticated anomaly detection models facilitates a process of detection of zero-day threats, which can be defined as the newest security risks that may go undetected by the standard tools and applications. The AI-based models can set the standard of the normal operation of the

network and raise an alarm when there are variations which could be new threats. Andy Schneider “Anomaly detection and the xz-utils zero-day: A Composite Alert demonstration story” published on Lacework proves how anomaly detection can help in the identification of the unusual activities thus improving the detection of sophisticated and covert threats.

3. Integrate AI-Driven Pen testing into CI/CD Pipelines for Automated Security Testing:

The integration of AI-based pen testing tools into the Continuous Integration/Continuous Deployment (CI/CD) as a method of implementing continuous and automatic security testing across the entire software development process. This integration enables real-time identification and fixing of vulnerabilities as the code is being written and deployed and this is in conformity with the agile development practices. According to Evan Malamis blog “How PenTesting Curbs Security Leaks in the CI/CD Pipeline “on ConvergeTP, the integration of pen testing into the CI/CD workflows helps in establishing a strong and effective security posture that supports frequent testing and deployment without affecting the security concerns.

The implementation of the suggested advancements such as using real-world data sources, incorporating ADC anomaly detection, and embedding AI-based pen testing into the CI/CD process is expected to enhance the efficiency and scalability of penetration testing. The above enhancements can make it possible for organizations to counter current and future cyber risks while effectively managing security resources and ensuring that they are always on the offensive in the current dynamic environment.

Conclusion

The conceptual model is capable enough of automating the due diligence part of penetration testing and the types of models introduced are sufficient to speed up the pentesting lifecycle and this solution can be reiterated again and again, staying up to date and to find new undiscovered vulnerabilities. Incorporating well trained models makes the framework scalable, flexible and adaptable. The team worked together to research about type of machine learning and determining the most appropriate solutions for each phase of pentesting. The concept can lay the foundation of well trained AI which can be used in future to detect zero day vulnerabilities before they are exploited by threat actors. AI-driven pentesting models can have significant impact on real-world cybersecurity practices.

References

- What Is Cybersecurity? The Realities of the Digital Age | American Public University.
- Penetration Testing and Why it is Important | GetOppos
- Automated vs. Manual Penetration Testing: Finding the Right Balance for Cybersecurity | Yellow
- Manual Penetration Testing vs Automated Penetration Testing
- Manual versus Automated Penetration Testing - Tangible Security
- The Impact of AI on Penetration Testing | Oppos
- https://www.linkedin.com/pulse/exploring-role-generative-ai-penetration-testing-jamison-utter-mrsdc/
- AI Pen Testing: Is Automated Penetration Testing Right for You?
- Can AI Enhance Penetration Testing?
- The history of penetration testing | Infosec
- Penetration Testing History - Christian Espinosa
- Understanding the Five Phases of the Penetration Testing Process
- https://www.linkedin.com/pulse/penetration-testing-using-ai-onkar-mhaskar-fmhrf/
- What Is the Role of AI in Threat Detection? - Palo Alto Networks.
- AI in Cybersecurity: Benefits and Challenges
- Top 5 Penetration Testing Tools and Techniques Used by Experts | by John Nathan | Medium
- Pentest Tools: The Complete List with Benefits, Limitations and Best Practices
- Deep Exploit: Fully Automatic Penetration Test Tool Using Reinforcement Learning - HackMD

- What is Classification in Machine Learning? | Simplilearn
- Machine Learning Regression Explained - Seldon
- Next Level Reconnaissance with Shodan and Censys - BreakPoint Labs
- How Pen Testing Curbs Security Leaks in the CI/CD Pipeline - Converge Technology Solutions
- Anomaly detection and the xz-utils zero-day: A Composite Alert demonstration story | Lacework

  AI and Zero-Day Attack Detection: Anticipating Unknown Threats | by Megasis Network
- https://www.datacamp.com/blog/classification-machine-learning
- Reinforcement learning - GeeksforGeeks
- Regression in Machine Learning: Definition and Examples of Different Models.
- Automating Network Scanning with Python and Nmap | by Amal Tom Parakkaden | Medium
- Tools :
    - Shodan API
    - Nmap Network Scanning - Official Guide
    - Metasploit
    - Nessus - Install Tenable Nessus
    - draw.io
    - Burpsuite - https://portswigger.net/burp/documentation/desktop/getting-started/download-and-install