# Introduction

## Background of the Course Topic

* + - Challenges/ complexities of Penetration Testing [1]
      * high levels of expertise [1]
    - Explores the automation group of AI, RL, and ML
      * automation is necessary in penetration testing due to the [2]
        + growing complexity of networks
        + Traditional manual penetration testing methods are time-consuming, resource-intensive, and may not keep pace with the dynamic nature of cyber threats.
        + “incorporating machine learning in an automated PT system will reduce recurrent human errors due to tiredness, omission, and pressure.”
        + “reduce time and resources’
        + “relieve network congestion and downtime by working after normal working hours”
      * Automated PT “uses advanced algorithms, machine learning, and AI to scan systems for vulnerabilities.”[4]
    - Definitions
      * Penetration Testing - Rev2, 4
        + “Penetration testing (PT) is an offensive approach aiming to evaluate the security of digital assets (network, website, application, database, database) by trying to actively identify its vulnerabilities and then exploit them in the same way as a genuine attacker.”[2]
        + “method of evaluating the security of computer systems or networks by simulating an attack by a malicious individual.”[4]
        + “goal of penetration testing is to provide organizations with actionable information about their security posture, enabling them to identify and prioritize areas of risk and improve their overall security”[4]
      * RL – Rev4
        + “Reinforcement Learning (RL) is a subfield of Machine Learning (ML) that deals with the problem of an agent learning to interact with its environment in order to maximize a reward signal through trial and error.”
        + RL is often referred to as Q or Deep-Q learning
        + Q – quality
        + “reinforcement learning to enable an agent to learn in an interactive environment by trial and error approach using feedback from its own actions and experiences”[5]

## Importance and Relevance of the Course Topic

* + - A
  + Outline Pentesting Steps
    - Pentesting Step 1 - Gathering Information (Module 1)
    - Pentesting Step 2 - Scanning (Module 2)
    - Pentesting Step 3 - Exploiting (Module 3)
    - Post-Breach/Exploit, Pentesting Steps 4,5, and 6 (Module 4)

# Literature Review

## Research Paper 1: Introduction of Topic

### Details

* + - * Getting pwn’d by AI: Penetration Testing with Large Language Models
      * Andreas Happe, Jürgen Cito
      * [link](https://arxiv.org/pdf/2308.00121.pdf)

### Introduction

* + - * Necessity of for Sparring Partners
        + partners offer alternative ideas/approaches
        + “A good sparring partner should be able to cover the different tactics, techniques, and procedures covered by ATT&CK.”
        + recent interview emphasized need for human sparring partner
        + critical shortage of skilled security professionals

ISC2 Cybersecurity Workforce Study 2022

* + - * Use of Large Language Models (LLMs), in penetration testing
        + LLM is a neural network that is highly trained with an extensive volume of data using self-supervised learning techniques.
        + Emergence of Large Language Models (LLMs)

*The first thought many people have when they hear AI in relation to society today, they think of LLMs*

ie ChatGPT, GPT3.5, AutoGPT

*While the traditional GPT prompts, ChatGPT, GPT3.5, are started by a user initiating the conversation with a prompt… AutoGPT automates the initial query/prompts*

* + - * + Using AI to train “novice penetration testers”
        + Outsource sparring partnerships to AI [SEGUE]

### Summary

* + - * explores Outsource sparring partnerships to AI
        + “empower existing human security testers”
        + “counteract the lack of sufficiently educated security professionals”

#### High-Level Guidance (Task Planning)

* + - * + planning phase

tasks that involve strategic planning and decision-making

* + - * + Examples:

design overall pentest plan,

determine tactics/techniques, and

identify potential vulnerabilities

#### Low-Level Guidance (Vulnerability Hunting)

* + - * + Execution phase

More detailed/ specific actions

provide step-by-step assistance during execution

often after the strategic planning has completed

Assumed Pentesters have already selected a specific tactic and just require specific TTPs

* + - * + Examples:

Identifying/targeting vulnerabilities within a specific system

executing commands/ specific exploits

information on how to escalate privileges or perform specific attacks

### Methodologies

* + - * referred to MITRE ATT&CK to identify various tactics, techniques, and procedures (TTPs) that are commonly used by adversaries
        + Authors “assume that a sparring partner for penetration testing should cover the whole TTP spectrum”

#### High-Level Guidance (Task Planning)

* + - * + Expectation

LLM “should be able to select suitable tactics and corresponding techniques.”

* + - * + Experiment

used to provide guidance/ insights at a strategic level.

Ie suggest a penetration testing plan for a target organization/ recommend tactics to use

prompted the “LLM to help design penetration tests for both generic scenarios as well as for a concrete target organization”

generic

*To achieve this, the authors* *instructed AgentGPT, an autonomous AI utilizing the GPT3.5 architecture, with the task of "Becoming a domain admin in an Active Directory."*

Concrete

AutoGPT is tasked with creating a penetration testing plan to test an approved organizations external-facing systems, networks, and digital assets from the perspective of a potential external attacker.

AutoGPT also continued on to crawl the company's website and identified potential targets for phishing attacks, (users and their email addresses)

* + - * + Results

AutoGPT returns a comprehensive plan which included common/standard methods :

Network vulnerability scanning.

OSINT (Open-Source Intelligence)

user enumeration

Phishing attacks against identified users.

“declined to perform any “real” network security scan or perform phishing operations due to its ethical filters.”

“Both answers were realistic, and feasible, and would give a penetration tester good feedback about potential attack vectors.”

#### Low-Level Guidance (Vulnerability Hunting)

* + - * + Expectation

LLM “given an employed tactic, it should be able to derive feasible techniques and procedures”

The main goal is to achieve privilege escalation and gain root access on the virtual machine.

* + - * + Experiment

analyze the state of a vulnerable virtual machine, suggest specific commands or actions to exploit vulnerabilities

Scenario: Assumed Pentesters have already gained some level of access, and require guidance to escalate to root

Authors create A deliberately vulnerable Linux virtual machine

set up a connection between GPT3.5 and a VM (VM was intentionally designed to have security vulnerabilities)

“Python script that uses SSH to connect to a deliberately vulnerable *lin.security* Linux virtual machine”

Asked the LLM to analyze the VM's state, generate commands or actions, and potentially control or influence the VM's behavior.

The script operates in an infinite loop.

Within the loop, it instructs GPT3.5 (or LLM) to “imagine being a low-privilege user” and to suggest Linux shell commands.

The suggested commands are executed over SSH on the vulnerable virtual machine.

The command and resulting output from the vm are then sent to GPT and it is asked to analyze the provided command and its output and identify any potential security vulnerabilities in the system

IF GPT identifies a vuln, it is instructed to provide steps on how to exploit the vulnerability

To bypass the ethical filters, the authors utilize prompt engineering by asking for “verification commands” instead of “exploitation examples”

Note: all command, output, and subsequent actions were meticulously documented.

Script is ran multiple times

* + - * + Results

successfully obtained root privileges

Identified and exploited security vulnerabilities, including the use of sudo commands, GTFObins, and retrieval of /etc/passwd for privilege escalation.

“GTFObins are benign system commands that when called through sudo, can be abused to gain a root shell.”

Recorded and analyzed command executions and their outputs to determine vulnerability grounding.

### Main Findings

#### General

* + - * + GPT3.5 displayed signs of understanding causal relationships between actions and outcomes during the experiments.

*the suggestions provided by the LLM seemed to* *exhibit a degree of logical thinking or understanding of cause-and-effect relationships.*

* + - * + GPT3.5 suggestions consistently followed logically from the obtained data

LLM Was able to make suggestions based on its broader understanding of common vulnerabilities in Linux environments, even without specific information about the target system's configuration or weaknesses.

* + - * + However, the suggested system commands were clearly based on pattern-matching

*While there were instances where the LLM's suggestions might have seemed rational, the authors emphasize that these suggestions were still primarily driven by pattern recognition and did not reflect a deep understanding of the subject matter. They are acknowledging that the LLM's responses might appear reasonable on the surface but are ultimately based on data patterns rather than true comprehension.*

*despite the appearance of reasoning,… suggestions were primarily generated based on patterns it had learned from its training data, rather than a genuine comprehension of the Linux systems.*

* + - * + *The outcomes provided included "realistic, feasible, and commonly employed" exploit options*

#### Hallucinations

* + - * + Invention of “facts that seem statistically plausible.”
        + Can be reduced by “using external knowledge and automated feedback” such as with AutoGPT
        + Happened rarely and were easily identifiable

*These hallucinations were infrequent and easily identifiable, with the most common example being the suggestion to execute "exploit.sh."*

#### Reproducibility

* + - * + On a small scale, the performance of LLMs are unstable and inconsistent

variations in the commands generated and vulnerabilities identified during these runs

Variation in single runs related to focus on specific aspects

likened to the experience of humans who become overly focused on particular a element and end up “going down a rabbit hole”

* + - * + longer runs, or aggregated results from multiple runs led to more consistent outcomes
        + outcomes/ results produced were less predictable/ consistent compared to traditional enumeration tools like linpeas.sh.

“Compared to tools such as linpeas.sh [30], LLMs seem to be less deterministic”

#### Ethical Moderation in LLMs

* + - * + LLMs are limited by their ethics filters

safety measures against malicious prompts

* + - * + Prompt Engineering

Slight variations in prompts helped reduce ethical concerns.

ie asking for “verification commands” instead of “exploitation examples”,

ie instructing the AI not to ask questions or make judgments

these prompts helped reduce ethical denials and prove that these ethical filters can be bypassed

can be used by both legitimate security professionals and malicious actors.

* + - * + switching from OpenAI to one of the
        + locally running LLMs would remove all server-side ethics checks.

#### Future:

* + - * + propose enhancing the collaboration between penetration testers and AI by integrating high- and low-level tasks into a single AI system.
        + suggest evaluating locally run AI models for cost savings and customization.
        + Improving AI memory and context retention could reduce errors
        + automated prompt generation could enhance query effectiveness.
        + emphasize the need for cybersecurity professionals to prepare for AI-driven attacks

### Relevance to Your Course Content

* + - * Focuses on Topic Introduction

#### Provide as an introduction to AI in pentesting

* + - * + Focuses on one of the more popular AI models people are familiar with: the integration of AI, specifically LLMs, into penetration testing.
        + Shows some basic prompt-response techniques that can be leveraged to assist in high-level planning and low-level vulnerability identification
        + This integration also shows how AI can streamline various aspects of penetration testing

#### Real-world Application:

* + - * + The paper shows practical applications of AI in penetration testing scenarios.
        + Real-world scenarios are important for allowing students to see AI's transformative power through real examples. (interesting, engaging, etc)

#### Introduction to Ethical Dilemmas:

* + - * + Begins introducing the ethical challenges/ dilemmas that arise when AI is used in penetration testing

Remember that many hackers are not concerned with the ethics

When is prompt engineering okay/when is it not?

* + - * + Explain that while AI can enhance security, it also brings about unique ethical concerns.

## Research Paper 2: Pentesting Step 1 - Gathering Information (Module 1)

### Details

* + - * Reinforcement Learning for Intelligent Penetration Testing
      * Mohamed C. Ghanem, Thomas M. Chen
      * [link](https://ieeexplore.ieee.org/document/8611595)

### Introduction

* + - * Focuses on Pentesting Step 1 - Gathering Information
      * *Introduce Reinforcement Learning as a subset of AI*
      * automation is necessary in penetration testing due to the
        + growing complexity of networks
        + Traditional manual penetration testing methods are time-consuming, resource-intensive, and may not keep pace with the dynamic nature of cyber threats.
        + “incorporating machine learning in an automated PT system will reduce recurrent human errors due to tiredness, omission, and pressure.”
        + “reduce time and resources’
        + “relieve network congestion and downtime by working after normal working hours”
      * Current Automation is limited
        + Optimization is a major challenge in designing Automated Systems

“ ensure that all existing threats are checked systematically and efficiently. The system should not take excessive time by processing irrelevant tasks and at the same time ensure that no threat is overlooked.”

current automation remains either limited to specific tasks or inadequately optimized

Existing systems often fall short, especially on large-scale assets, compared to human experts.

* + - * + critical issues

vast volume of data generated during comprehensive testing,

underutilization of information,

rapid emergence of new threats,

complex and evolving attack paths

* + - * paper objective: use ML techniques to develop an Intelligent Automated Penetration Testing System (IAPTS) to automate and improve the penetration testing processes.
        + “The ultimate goal is a system capable of imitating human PT experts in performing an intelligent and automated pen test”

### Summary

* + - * Complexities of PT
        + humans themselves find challenging to master fully.
        + *The authors note that blind automation (complete automation without any human intervention) is impractical, especially during the initial phases which often yield incomplete results that result in uncertainty and often necessitate continuously revisiting tasks and changing approaches.* *However, by using machine learning, specifically reinforcement learning, to automate these phases intelligently, the automation can be more closely resemble a human expert’s decision-making processes.*
      * Challenges of Automation
        + While autonomous systems are not a new idea, and are regularly utilized in the industry, these systems require a lot of hands-on guiding and require a lot of time and resources, and as a result are limited to smaller networks.
      * Automation will improve PT
        + reducing the cost of manual repeated/methodical testing

*As “PT should be repeated and performed on a regular basis to ensure continuous security”, Automation can help in various ways, such as…*

* + - * + more efficient/targeted, which will reduce strain on tested assets
        + reducing the time by streamlining/automating repetitive tasks
        + improve adaptability in relation to ever-evolving landscape of cyber threats
        + Improved/ unconventional techniques
      * Use Reinforcement Learning
        + Reinforcement Learning: “reinforcement learning (RL) is concerned with goal-directed learning and decision making which accurately reflect the PT context”

*Instead of using predetermined rules, as seen in analyst-driven solutions, RL…*

* + - * + learns through the consequences of its interactions and focuses on long-term instead of short term
        + [segue] *a crucial step in applying RL to address these PT challenges.. converted into a formal computational model known as POMDP*

### Methodologies

* + - * Framework: POMDP
        + Took the real-world PT challenges and scenarios and have translated them into a mathematical framework known as Partially Observable Markov Decision Processes (POMDPs).
        + Model:

incorporates elements such as state observations, actions, rewards, and transition dynamics

RL agent makes decisions (actions) based on its observations and aims to maximize cumulative rewards.

RL agent learns strategies to solve complex problems

Stores solutions in memory for similar cases in the future

* + - * + Method: Policy Search

determining the optimal sequences of actions that lead to the highest cumulative rewards over time

the general approach they used to find optimal strategies within the POMDP framework

* + - * + Algorithms

By combining these algorithms with our POMDP model, we aimed to address the challenges of PT systematically and autonomously.

PERSEUS [2]

“randomized point-based value iteration” algorithm designed specifically for solving POMDPs

“operates on a large set of beliefs which are gathered by simulating random interactions”[POMDP]

performs “backup stages” to this set of beliefs

gradually improves understanding of the situation

performs rounds of calculations where ensures “the value of each point in the belief set is improved (or at least does not decrease)”[POMDP]

PEGASUS [2]

policy search method

*This method is used to determine optimal sequences of actions, known as policies, that lead to the highest cumulative rewards over time*

transforms into an equivalent deterministic

each state/action has one possible outcome

calculates an estimated value for all policies

aka simulates policies to figure out which one is likely the best

“Pegasus method (Ng & Jordan, 2000) estimates the value of a policy by simulating a (bounded) number of trajectories from the POMDP using a fixed random seed, and then takes steps in the policy space in order to maximize this value.”[POMDP]

estimates the value of policies and seeks to find a policy with a high estimated value, simplifying the search for optimal decision policies.

Best for solving large POMDPs

*effective when dealing with PT problems, as it offers* *“polynomial rather than exponential”* *time complexity, making it suitable for large-scale PT scenarios.*

Other Reinforcement Learning (RL) Algorithms:

genetic algorithms (GA)

temporal difference (TD) methods.

Used for benchmarking and testing their system's performance.

* + - * system design and its operative modes.
        + Proposed system: Intelligent Automated Penetration Testing System (IAPTS)
        + clustering and data processing Set up manually at the beginning of development, with hopes that as the system evolves and learns, it may become more autonomous and require fewer manual interventions.
        + *In the beginning, the learning process is done by humans - They teach the system and provide it with knowledge and expertise. However, over time, the system will develop modules or components that can do this learning automatically. These modules will be added to the system gradually. So, initially, people teach the system, but eventually, the system will learn and improve on its own.*
        + Modes:

Fully Autonomous (Level 4): system operates entirely on its own

Partially Autonomous (Level 3): the system operates independently but under supervision

Decision-Making Assisting Mode (Level 2): system works alongside a human expert, “processing the data on a real-time basis and suggesting better alternatives.”

Learning Mode (Level 1): a human performs PT while the system observes and learns

* + - * Testing
        + primary goal was not to evaluate performance in terms of capabilities/effectiveness, but to demonstrate that RL can be suitably and effectively applied to PT and to evaluate how well IAPTS learned/adapted its policies over time.

early stages of validating potential/refining its learning

* + - * + Simple Simulation:

tested IAPTS within a controlled environment

set up a simulated network consisting of seven machines (M0 to M6) to mimic real-world PT scenarios

provide insights into how IAPTS would perform in real-world penetration testing scenarios by simulating various conditions and measuring its performance metrics/identify weaknesses

execution times

how long IAPTS would take to complete various penetration testing tasks.

Used this to “calculate an approximate PT time”

* + - * + Experience replay: simulated scenarios where the same network underwent updates and upgrades

evaluate how well it learned policies/ adapted to changes in the network.

### Main Findings

* + - * in the Experience Replay, they found that the knowledge from the previous test was not only learned and stored successfully, but the policies were effectively reused in the majority of instances.
      * RL can significantly improve the performance of automated PT
        + when compared to humans and Blind Automation, their system significantly reduces the time required to complete PT tasks

*IAPTS outperforms both the manual approach, which relies on human expertise, and the blind automation approach, where tasks are automated but lack intelligent decision-making…*

*more efficient and effective*

*saves time and resources*

* + - * + generated alternative attack suggestions that human experts might overlook
      * RL-generated attack policies are relevant and accurate.
        + the RL-based system produced highly relevant attack policies when targeting the most secure machine in the network, Machine M2.
        + the authors found that that these policies were highly plausible or realistic in terms of how an actual attacker might approach and carry out an attack on the target system or network
      * Designed with Flexibility in mind
        + intentionally designed in a way that makes it easy to add additional features and functionalities in the future.
        + Hope to one day include features like:

Pivoting: using compromised systems to launch further attacks within a network

Security Clustering (Isolation): segregating or isolating certain parts of the network for security reasons.

Preprocessing: preparing/ optimizing data or information

Memory Management Modules: components to handle how data is stored/ accessed

### Relevance to Your Course Content

* + - * Introduction to PT Challenges:
        + Especially regarding PT Step 1
        + introduces the challenges associated with penetration testing
        + Gives a great overview of what Penetration Testing is/does/advantages/disadvantages/etc
      * Automation as a Solution: Discuss how the paper
        + highlights automation as a potential solution to address PT challenges. Emphasize that

PT is traditionally labor-intensive/time-consuming, automation offers a way to improve this

* + - * + Automation is important to addressing the growing complexity of threats.
        + providing insights into the benefits/challenge
      * cutting-edge techniques
        + introduces advanced techniques like Reinforcement Learning and POMDPs
        + Reinforcement learning (RL) is a subset of artificial intelligence (AI)
        + crucial for staying up-to-date with the evolving field of cybersecurity
        + hints at future research directions/ shows that automated PT is evolving, and there are ongoing developments to look forward to.
        + broaden perspective on the possibilities in the field.

## Research Paper 3: Pentesting Step 2 - Scanning (Module 2)

### Details

* + - * Penetration Testing Procedure using Machine Learning
      * Reevan Seelen Jagamogan, Saiful Adli Ismail, Noor Hafizah Hassan, Hafiza Abas
      * [link](https://ieeexplore.ieee.org/document/9870951)

### Introduction

* + - * Focuses on Pentesting Step 2 - Scanning
      * paper's objective:
        + To evaluate GyoiThon's effectiveness in identifying vulnerabilities.
        + examining how efficient/capable/successful it works in identifying vulnerabilities and potential security weaknesses during a PT
      * GyoiThon
        + Employs ML, specifically Naïve Bayes algorithm
        + Leverages other PT tools to enhance its capabilities.
        + automates the process of gathering data from target URLs

Known Vulnerabilities:

detect vulnerabilities documented and reported in sources like the National Vulnerability Database (NVD).

Software Identification:

determine and report the various software components that make up a web server's environment

Configuration Weaknesses:

instances of comments in HTML or JavaScript code that may reveal sensitive information or other insights that pose security risks

Authentication Issues:

Can specifically focus on login pages to detect and flag issues or vulnerabilities associated with the authentication process.

General Web Application Vulnerabilities:

Can identify a range of common web application vulnerabilities SQL Injection, XSS, CSRF, security misconfigurations, etc

Can detect general web application vulnerabilities that are widely recognized and frequently targeted by attackers

### Summary

* + - * RQ: ’How effective is the GyoiThon tool in detecting vulnerabilities?’.
      * hypothesis:
        + if a penetration tool incorporates a ML algorithm, it would be more effective in searching for and identifying vulnerabilities compared to tools that do not use Machine Learning.
      * Emphasis on the comparison between GyoiThon's default and Machine Learning modes.
      * Explores previous ML PT methods
        + HARMer Tool:

mimics realistic attacks in order to assess a system's ability to defend against them

“was determined to be useful for penetration testers”

* + - * + FUSE Tool:

detects vulnerabilities in web applications

designed to address security weaknesses related to file uploads in web applications

“from the evaluated web apps the author was able to find 30 undisclosed UEFU vulnerabilities and 15 Common Vulnerabilities and Exposures (CVEs)”

* + - * + Intelligent Automated Penetration Testing System (IAPTS):

Mentioned in previous lit review

leverages RL and various algorithms, including PEGASUS and PERSEUS, within a POMDP model, to automate and optimize penetration testing procedures.

* + - * Executed a “penetration testing procedure” using GyoiThon
      * Highlighting the importance of the study in the context of AI-driven penetration testing.

### Methodologies

* + - * exploratory methodology
        + focus on a relatively new / less-studied area of interest.
        + The only instance of a comprehensive PT study of GT was conducted by the developer
      * Provide isolated environment using the Kali Linux OS inside VirtualBox
      * Used GT to detect vulnerabilities in unencrypted HTTP data exchanged over Port 80, and in encrypted HTTP data exchanged over Port 443
      * target websites were set up on a server provided by OWASP and accessed via a locally hosted environment
      * Executed PT using two modes
        + Default Mode: This mode includes steps like gathering HTTP responses, identifying product/version information, assessing vulnerabilities using Common Vulnerabilities and Exposures (CVE) numbers, examining unnecessary HTML/javascript comments, analyzing debug messages, and assessing login pages.
        + Machine Learning (ML) Mode: In addition to the steps in the Default Mode, the ML Mode incorporates the use of the Naïve Bayes algorithm for product/version identification.
        + “GyoiThon [was] not tested to its full potential…[it] has 9 modes to be utilized for penetration testing, but only 2 of them are tested for this penetration testing process”

“only the Default and Machine Learning (ML) modes were evaluated and compared to see whether the ML algorithm (Na¨ıve Bayes) used is more effective than using normal penetration testing tools.”

### Main Findings

* + - * Port 80 had more vulns in both default and ML modes
        + It is expected that port 80 had more vuln as port 80 is used for HTTP, which is less secure compared to HTTPS (Port 443). The lack of encryption in HTTP makes it more susceptible to vulnerabilities and attacks.
      * ML Mode detected more vulnerabilities, particularly in Port 80, compared to the Default Mode.
        + “Since the Machine Learning mode of GyoiThon managed to detect three more vulnerabilities that the Default mode could not detect, the hypothesis stating that the Machine Learning algorithm works effectively has been proven to be true.”
      * Unable to identify vulnerabilities that aren’t recorded in the NVD, as GyoiThon may not have the information it needs to recognize and flag it during the scanning process
        + *This highlights a limitation of the tool's reliance on existing vulnerability data sources; it may miss vulnerabilities that haven't been previously documented in such databases.*
        + This highlights the tool's reliance on secondary sources for identifying certain vulnerabilities
      * while GyoiThon is a valuable tool with potential, it has room for improvement.
        + Need to test against real sites
        + Need to test all 9 modes

### Relevance to Your Course Content

* + - * Explores AI in Penetration Testing:
        + explores the application of AI, specifically Machine Learning, in penetration testing using GyoiThon.
        + delves into how AI-driven tools can identify vulnerabilities, which aligns with the central theme of my course.
      * Comparative Analysis:
        + The experiment is acomparative analysis between default penetration testing methods and those augmented with AI
        + this shows the effectiveness of AI-driven approaches.
        + This can serve as a case study for the course to showcase the advantages of AI in cybersecurity.
      * Discusses Common Vulnerabilities:
        + discusses common vulnerabilities found in web applications, which is essential knowledge for any cybersecurity professional.

## Research Paper 4: Pentesting Step 3 - Exploiting (Module 3)

### Details

* + - * Vulnerability Exploitation Using Reinforcement Learning
      * Anas AlMajali; et al
      * [link](https://ieeexplore.ieee.org/document/10185700)

### Introduction

* + - * Focuses on Pentesting Step 3 - Exploiting
      * Current Automated Exploits
        + existing tools are time-consuming/Resource-Intensive as they adopt a brute-force approach of trying every possible payload
        + authors created an AI agent that has been trained using reinforcement learning techniques and trained it to exploit specific vulnerabilities

“designed to be trained on various vulnerabilities and operating systems”

* + - * + stores successful exploits as a record of different states (combinations of operating systems and vulnerabilities) and the corresponding payloads that have a high likelihood of success in those states
        + in the future, it references this table and intelligently performs the exploitation
        + lack customization options.
      * Objective: develop an intelligent agent for automating exploitation in penetration testing.
      * Metasploit
        + a well-known penetration testing tool
        + has wide range of payloads for various purposes
      * Hope to lay the foundation towards creating a “general agent”
      * “continuously improve its exploitation strategy over time”

### Summary

* + - * RL Agent:
        + The authors use ML to create a reinforcement learning (RL) agent which makes decisions by interacting with an environment.
        + Primary goal of this RL is to exploit vulnerabilities
        + Hopes to eventually develop a "general agent" capable of handling various vulnerabilities and operating systems.

This vision extends beyond the specific context, but the paper offers insights into the potential for versatile cybersecurity tools.

### Methodologies

* + - * Train
        + Environment

cross referencing the specific operating system of the target machine and the specific vulnerability being targeted

The combination of an operating system and a vulnerability are referred to as "states"

The state represents the current situation or configuration of the environment at any given moment

* + - * + Learning

By representing the environment as states consisting of operating system-vulnerability combinations, the RL agent can make informed decisions about which actions to take.

These states allow the RL agent to adapt its decision-making based on the specific context it encounters

Use a well-known reinforcement learning technique, Q-learning, to help their agent learn which actions to take in different situations to maximize rewards.

“The RL algorithm is utilized to train the agent to identify the most appropriate payload for a given operating system and vulnerability.”

* + - * + Reward System

Use a point based reward system with +100 reward for success, and -10 for failure,

Used to motivate the agent's decision-making

“This often results in the agent displaying more conservative behavior and being less likely to take risks that could potentially yield negative outcomes.”

* + - * + Parameter Tuning:

Uses parameter tuning efforts, such as epsilon (exploration rate), alpha (learning rate), and gamma (discount factor), and their impact on performance.

α (Learning Rate): Determines how much weight new information should have compared to old information. High α values prioritize new information, while low values favor old information.

γ (Discount Factor): Balances immediate and future rewards in decision-making.

ε (Exploration Rate): Represents the probability of the agent taking a random action instead of the one with the highest expected reward. High ε values encourage exploration, while low values favor exploitation.

selects an action> applies to environment > observes result and state > receives reward a reward > decrease epsilon and repeat.

“gradually decreasing epsilon after each iteration to focus on exploiting the best known actions.”

the agent uses epsilon greedy strategy to balance between exploring new actions (randomly) and exploiting its current knowledge (choosing what it thinks is the best action) when making decisions. The value of epsilon determines how often it explores (randomly) versus how often it exploits (chooses what it thinks is the best action). A higher epsilon value encourages more exploration, while a lower value encourages more exploitation.

* + - * + Actual Experiment

Learning Phase - agent is learning learning and improving its exploitation skills

across 7 trials, the agent average 2.5 hours to execute 500 attempts to exploiting vulnerabilities

500 was a “deliberate choice”

the agent is learning and fine-tuning its strategies through repeated attempts - exploring different actions and learning from its successes and failures

a computationally intensive process as it's actively learning and adapting

* + - * Exploit
        + payload Selection:

agent selects payloads from the Metasploit framework based on its learned strategies.

evaluates payloads from the Metasploit used to deliver an exploit and compromise a system

payloads from the Metasploit framework. In this context, "payloads" refer to pieces of code or scripts used to deliver an exploit and compromise a system.

* + - * + Successful Exploitation:

agent success is measured by its ability to establish a reverse shell session on the target system

* + - * + Actual Experiment

Deployment Phase - After the training phase, the agent was considered trained and ready to be used in a real-world scenario

agent deployed on multiple vulnerable machines that have a "remote code execution," (execute malicious code on a target system from a remote location) vulnerability that exists in Apache CouchDB, specifically in Version 3.1.0 of the software

it was able to exploit these vulnerable machines in “mere seconds.”

This indicates that the training phase prepared the agent well, and it could execute its learned strategies very quickly and effectively on real-world systems.

* + - * a

### Main Findings

* + - * Success rate
        + as the agent undergoes training, it is beneficial to slowly shift from trying out new, exploratory actions to relying more on actions that it has learned are effective, which leads to higher success rates in accomplishing its tasks
        + striking a balance between learning and randomness
      * showcase the efficacy of the RL agent in automating exploitation tasks.
        + the agent successfully exploited vulnerabilities in Apache CouchDB version 3.1.0, achieving the primary goal of establishing reverse shell sessions in a 8.26 seconds on average

*not only achieving the goal, but doing so in 8.26 seconds…*

* + - * The paper shows that the use of RL algorithms to enhance PT can be a reliable method to verify exploitable vulnerabilities.
        + Focuses on maximizing overall rewards
      * More efficient/effective than other current tools
        + RL agent's approach is more efficient/effective than current existing tools that rely on exhaustive testing

*Wihle RL agent's training phase takes time and resources but contrasts this with the deployment phase, where it exploits vulnerabilities within seconds.*

* + - * + By selecting optimal parameters α (Learning Rate)/ ε (Exploration Rate) RL has an average success rate of 83.64%

Success Rate = (Number of Successful Runs / Total Number of Runs)

* + - * + This ability to fine tune results in an adaptability, versus a one-size-fits-all approach

### Relevance to Your Course Content

* + - * cutting-edge techniques
        + Explores the innovative use of RL algorithms for vulnerability exploitation
        + shows another way that RL can be used in the field

Demonstrates how RL can offer multiple, unique approach to solving complex security challenges.

* + - * + Versatile/broad of a topic/ encompass many different methods

Can be adapted to encompass a wide range of applications within the field

* + - * + crucial for staying up-to-date with the evolving field of cybersecurity
        + broaden perspective on the possibilities in the field.
      * Adaptability
        + provides insights into the impact of parameters tuning like learning rate and exploration rate on the performance of RL algorithms.
        + can show the benefits of fine-tuning.optimizing security tools
        + keep up with changing conditions and evolving threats
      * Importance in PT
        + RL consistently selects the best actions to maximize rewards

importance of developing similar tools that prioritize effectiveness in penetration testing

* + - * + *This adaptability and … highlights the necessity for the development of similar tools*

## Research Paper 5: Post-Breach/Exploit, Pentesting Steps 4,5, and 6 (Module 4)

### Details

* + - * Automated Post-Breach Penetration Testing through Reinforcement Learning
      * Sujita Chaudhary, Austin O’Brien, Shengjie Xu
      * [link](https://ieeexplore.ieee.org/document/9162301)

### Introduction

* + - * Focuses on Post-Breach/Exploit, Pentesting Steps 4,5, and 6
        + Post-exploitation typically refers to the phase that occurs after an attacker has successfully gained unauthorized access to a system or network. In this phase, the attacker (or in this case, the agent) explores the compromised system further, often seeking to achieve specific objectives like accessing sensitive data or maintaining persistence.
      * Objective
        + apply machine learning in the post-exploitation phase of penetration testing.
        + train AI agent in diverse network environments to make it more flexible and suitable for a wide range of applications
      * “At the post-exploitation phase of pentest, there is still a huge gap to cover with machine learning algorithms”
      * Reasons for selecting DQ
        + Due to recent success of deep Q-learning methods in recent years, the authors opted it as their training model
        + it performs well in scenarios where the agent doesn't receive rewards frequently

### Summary

* + - * RL
        + Train agent using RL
      * Deep Q
        + Instead of having a detailed model of the environment, DQ learns directly from interactions with the environment.

Uses neural networks - takes current state as input and, produces estimated cumulative reward values for taking various actions (Q-values) for all possible actions in that state.

Q-value: number associated with Each action .

numbers represent how good/valuable each action is (higher = better)

The agent's then uses its decision-making process to compare these Q-values and selects the action associated with the highest estimated Q-value

Receives ‘rewards’ which it uses as feedback to adjust its Q-value estimates through a process called "Q-learning."

Thus Q-values are refined over time, which improves its decision-making ability. It learns which actions are more likely to lead to higher rewards in different states, all without explicitly modeling the environment.

### Methodologies

* + - * Training
        + Library Selection

specialized modular software called (TensorFlow Agents) TF\_Agents, to offer flexibility and extensibility

TF\_Agents leverages TensorFlow's capabilities for deep learning

* + - * + Environment

Simulate PT using various VMs, including windows and linux OS

They use environments designed to mimic real-world cybersecurity scenarios-> computer networks that have been specifically altered/adapted for the purpose of cybersecurity training and competitions.

To simulate target data, they use password files, shadow files, system configurations, etc

Offer feedback in the form of a rewards system, “scores”

Agents actions are limited to a predefined list of terminal commands, but some specific details within thecommands can vary based on what the agent observes in the env

* + - * + Set Baseline

Write/test a python script to find files

specify which files are considered rewarding

will compare how well their agent can perform tasks like finding and exploiting files compared to what their script achieved

* + - * + collecting data
        + estimate Q-values
        + iterative/improve

balance of exploitation/exploration

using a softmax function

temperature/randomness parameter

agent starts with higher level of randomness in action selection (exploration)

As it learns, it reduces randomness and favors actions with higher Q-values (exploitation)

Compare to baseline

measure the accuracy of their agent's predictions,

compare its performance to a randomized policy

assess effectiveness in locating and exploiting specific files

### Main Findings

* + - * Application of RL in Post-Exploitation
        + Ability to use RL, specifically DQ, in the post-exploitation phase
        + explores *how* an agent can be trained to perform tasks in a compromised network environment
        + “Conceptual Feasibility”
      * Importance of RL in Post-Exploitation
      * Sets foundation for Further Research
        + sets the stage for further research and development in the field of AI post-exploitation PT
        + provides a conceptual framework for future researchers to build on/ refine

### Relevance to Your Course Content

* + - * AI in the Post-Exploit phase
        + shows how AI agents can potentially automate post-exploitation activities
      * Exploration vs. Exploitation
        + concept of balancing exploration and exploitation in AI decision-making.
      * Real-World Application
      * Foundation for Future
        + *While the paper primarily focuses on a conceptual framework, it discusses real-world implications.*
        + While novel - it explains the logical steps required to turn this concept into a practical application
        + emphasizes the need for additional study in this area
        + By presenting more novel ideas - can stimulate critical thinking to advance AI in PT

# Incorporation of Findings into the Course

## Discussion on how the findings from the literature review inform your course development.

* + - a

## Preliminary Ideas for Course Content based on the Literature Review

* + - Break down into PT steps
      * Pentesting Step 1 - Gathering Information (Module 1)
      * Pentesting Step 2 - Scanning (Module 2)
      * Pentesting Step 3 - Exploiting (Module 3)
      * Post-Breach/Exploit, Pentesting Steps 4,5, and 6 (Module 4)

# Table of Contents for the Course

## Module 1: Introduction to AI and ML in Penetration Testing

### Module Overview and Brief Description

* + - * Module 1, comprised of "Introduction to AI and ML in Penetration Testing" and "Gathering Information and Reconnaissance with AI," serves as the foundation for the course. It offers participants an in-depth introduction to the role of Artificial Intelligence (AI) and Machine Learning (ML) in the field of penetration testing. This module introduces the first step of the Penetration Testing process, Gathering Information. Through these submodules, participants will gain a comprehensive understanding of key concepts, ethical considerations, and the terminology essential for the course.

### Submodule 1.1: Introduction to AI and ML in Penetration Testing

* + - * Overview of the Course
      * Overview of Penetration Testing
        + Pentest steps/phases
      * Overview of AI
      * Overview of Machine Learning
        + Deep Learning
      * Examine Ethical Considerations
        + Discuss Ethical challenges
        + Responsible/ lawful use of AI for security assessments
      * Key concepts and terminology
        + threats, vulnerability, exploits, AI, penetration testing, Machine Learning …

### Submodule 1.2: Gathering Information and Reconnaissance with AI

* + - * Introduction to Phase 1: Gathering Information
      * Use of AI-Driven Tools for Information Gathering
        + Shodan and Censys (“Asset Discovery”)
      * Machine Learning in Data Collection and Analysis
      * Ethical Considerations in Information Gathering
      * Demonstration: Gathering Information with AI

## Module 2: Machine Learning for Vulnerability Assessment

### Module Overview and Brief Description

* + - * Module 2, consisting of "Scanning and Vulnerability Assessment with ML" and "Exploiting and AI-Enhanced Techniques," focuses on utilizing Machine Learning (ML) in the context of vulnerability assessment during penetration testing. This module offers participants an in-depth exploration of Penetration Testing Phases 2, Scanning and Phase, and 3, Exploitation. By highlighting the role of AI and ML in these critical phases, participants will learn about AI-powered vulnerability scanners, data collection and preprocessing techniques, exploit development with ML, and practical vulnerability assessment exercises.

### Submodule 2.1: Scanning and Vulnerability Assessment with ML

* + - * Introduction to Phase 2: Scanning
      * Using Machine Learning for Vulnerability Assessment
      * AI-Powered Vulnerability Scanners
      * Data Collection and Preprocessing for ML
        + NLP
      * Practical Exercise: Vulnerability Assessment with ML

### Submodule 2.2: Exploiting and AI-Enhanced Techniques

* + - * Introduction to Phase 3: Exploitation
      * AI-Driven Exploitation Tools and Frameworks
      * Machine Learning for Exploit Development
      * Real-World Examples of AI-Enhanced Exploits
      * Demonstration: AI-Powered Exploitation

## Module 3: Post-Exploitation AI and ML Techniques

### Module Overview and Brief Description

* + - * Module 3, "Post-Exploitation AI and ML Techniques," delves into the advanced phases of penetration testing, Maintaining Connection, Covering Tracks, and Reporting. By highlighting the role of AI and ML in these critical post-exploitation phases, participants will gain insights into enhancing evasion techniques, streamlining reporting and documentation through AI-generated reports, and applying advanced AI-enhanced post-exploitation and privilege escalation strategies. This module equips participants with advanced skills and insights into the transformative power of AI and ML in post-exploitation scenarios, enhancing their ability to navigate and assess cybersecurity landscapes effectively.

### Submodule 3.1: Maintaining Connection, Covering Tracks, and Reporting

* + - * Overview of Maintaining Connection and Covering Tracks
      * Role of AI and ML in Evasion and Stealth Techniques
      * Reporting and Documentation with AI-Generated Reports
      * Practical Exercise: Maintaining Connection and Covering Tracks with AI

### Submodule 3.2: AI-Enhanced Post-Exploitation and Privilege Escalation

* + - * Techniques for Post-Exploitation with AI
      * AI-Driven Privilege Escalation Strategies
      * Realistic Scenario Simulations with AI
      * Demonstration:

## Module 4: Deep Learning and Advanced Techniques

### Module Overview and Brief Description

* + - * Module 4, "Deep Learning and Advanced Techniques," concludes the course, focusing on cutting-edge topics in penetration testing. Participants will explore the potential of Deep Learning, AI, and ML in advanced penetration testing techniques. This module wraps up the course by revisiting key concepts, exploring future trends in AI and ML within penetration testing, and offering additional insights and suggested references for further exploration in this dynamic field.

### Submodule 4.1: Deep Learning and Advanced Techniques

* + - * Introduction to Deep Learning
      * Deep Learning Applications in Penetration Testing
      * Advanced AI-Enhanced Techniques
      * Machine Learning for Zero-Day Exploits
      * AI-Driven Red Team Operations
      * Demonstration:

### Submodule 4.2: Review/Conclusion

* + - * Review Key Concepts
      * Future Trends in AI and ML in Penetration Testing
      * Additional Insights
        + Discuss References and Further Reading

## References and Further Reading

* + - a

# Key Components of the Newly Developed Course

## Learning Objectives

* + - *outline what students should be able to achieve or demonstrate by the end of the course*
    - Understand fundamental concepts of pentesting, AI, and ML
    - Understand penetration testing phases
    - Familiarity with various pentesting tools

## Target Audience

* + - *group of individuals or professionals for whom the course is designed*

## Assessment Strategies

* + - *how students' knowledge and skills will be evaluated and measured throughout the course*
    - A quiz at the end of each module to test for understanding
    - Assorted practical exercises

## Supplementary Materials

* + - *“description of what you intend to include in your future final deliverable”*
    - reading materials
      * e.g., online technical blogs, articles, whitepapers
      * about any topics, tools, and related concepts mentioned
    - official tutorials of some
      * software and/or hardware tools mentioned/ used

# Conclusion

## Summary of Key Insights

* + - a

## Potential Impact of the Course

* + - a

# References

* + https://ieeexplore.ieee.org/document/5460718

# Additional Facts

* + “cybercriminals are increasingly doubling the effectiveness of their attack tools for half the cost every few months”
    - <https://ieeexplore.ieee.org/abstract/document/8963730>

# Additional Resources to Consider

* + Reinforcing Penetration Testing Using AI
    - <https://ieeexplore.ieee.org/abstract/document/9843459>
  + Reinforcement Learning for Intelligent Penetration Testing
    - <https://ieeexplore.ieee.org/document/8611595>
  + Survey On The Applications Of Artificial Intelligence In Cyber Security \*\*
    - https://www.researchgate.net/profile/Aaron-Achi/publication/355119649\_Survey\_On\_The\_Applications\_Of\_Artificial\_Intelligence\_In\_Cyber\_Security/links/623c6b5991e0810f44d62f22/Survey-On-The-Applications-Of-Artificial-Intelligence-In-Cyber-Security.pdf