JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

AI BASED ENUMERATION AND EXPLOIT SUGGESTER

R Abhishek Reddy[1] | Chiranjeevi G[2] | Eshwar Guptha[3]| Mohammed Zaid Z[4]

[1,2,3,4] STUDENT

[1,2,3,4] Department of Computer Application

[1,2,3,4] Jain(Deemed-to-be)University, Bangalore.

Abstract:
This research is intended to implement Artificial Intelligence with Cyber Security. Our main goal of this research is to make an AI that can gather information on a specified target and search for the best possible exploits and all the vulnerabilities on the target using tools that are used by a Cyber Sec professional during a Pentest. This tool uses all the methods and techniques used by a Black Hat hacker to enumerate the target completely and give back proper information this increases accuracy as the AI will learn new things every time it does an enumeration it uses various sources and techniques from the user which can help it increase its efficiency.

Keyword: Artificial Intelligence (AI), Cybersecurity (CS), Pentester, CVE, Json, TCP\UDP Network, Vulnerability, Exploit

I.Introduction

Hacking or pentesting is a process involving a lot of effort and Time, it is divided into parts such as - enumeration, recon, exploitation, post exploitation, and covering tracks. Aiming to reduce the effort by eliminating the first task which is enumeration and footprinting it is tedious work as we need to browse many data and this is the most important step to be avoided by any hacker as it gives the main information needed to know the weak spots of your target making the basis of your hacking process if not done properly this can lead to errors and bias in results. The program aims to make the process as accurate as possible and give an analysis with almost perfect results.[1]

We are enhancing the data collection done by previous scanners such as openVAS, Nmap, Metasploit, etc. The additional technology we are working on is password cracker using aircrack, hashcat, johntheripper, and rainbow crack. We are also working on a packet analysis tool that helps in filtering specific packets from a huge pcap file which can be used in forensics as we are implementing low powered FTP, ssh, and bruteforcer using Nmap scripting language, Hydra, and Medusa this makes the use case of the tool to another level as it focuses more on the offensive side now as it can help in testing packets crack passwords and if successful hijack accounts access with the right password. There are many other implementations as it is server-based we can install the docker instance or the entire application in a cloud machine with a dynamic DNS or IP and get remote testing working and make things simple yet versatile to use.[5] There is a chance that our scan may not work but we can do the most testing to make logs making the tool strong can give the highest level of accuracy.

II.RESEARCH METHODOLOGY:

The research to justify the use case of the combination of two different fields and studies to produce a even better output. The main aim of this research is to determine the most effective resource and details necessary to go through with this research and provide effective results. The method used is Qualitative Approach to determine the best methods and effective techniques to gain results.

III.Problem Statements.

This research is intended to implement Artificial Intelligence with Cyber Security. The main goal of this research is to make an AI that can gather information on a specified target and search for the best possible exploits and all the vulnerabilities on the target using tools that are used by a cybersecurity professional during a Pentest this tool uses all the methods and techniques used by a Black Hat hacker to enumerate the target completely and give back proper information this increases accuracy as the AI will learn new things every time it does an enumeration it uses various sources and techniques from the user which can help It increases its efficiency. In a professional environment, hackers use a lot of tools for scanning targets mainly use tools like Nmap, Masscan, Rustscan, CVE database (CSV gov) Searchsploit. For the information related to open ports and services the scan by default will be an aggressive as this is used during a pentest by the pentester but there will be an option for a stealth scan for a more accurate Black Hat hackers environment. The Code primarily automates all the possible scans based on the info we initially provided.[6] The AI reads the scans and starts to enumerate all the possible ways it can scan it will record all its findings in a temporary directory Then it analyses the raw data and enumerates the target based on the huge logic of various outcomes and inputs using information from APIs and huge amounts of Databases of CVEs or vulnerabilities. These APIs and Databases help in creating more accurate results. Problem Statements using the data it further finds specific payloads for that particular target and using Metasploit will generate a resource file for further exploitation. The AI uses a complex algorithm of various outcomes and works on finding the smallest and best route to exploit once it has found the best route it writes that in its report for later use by the Pentester. It was used in creating pattern matching tools that alert analysts to the security issues in the network.[7] The tools can outpace the ability of human analysts in responding. More automation is needed in all aspects of Cybersecurity using AI technology.[8] The AI can reduce the execution time of attacks/defense while increasing their effectiveness and strengths. The mechanism of attacks/defenses expands from intelligent agents acting humanly to thinking humanly.[2]

Using the data it further finds specific payloads for that particular target and using Metasploit will generate a resource file for further exploitation. The AI uses a complex algorithm of various outcomes and works on finding the smallest and best route to exploit once it has found the best route it writes that in its report for later use by the Pentester. It was used in creating pattern matching tools that alert analysts to the security issues in the network. The tools can outpace the ability of human analysts in responding. More automation is needed in all aspects of Cybersecurity using AI technology.[9] The AI can reduce the execution time of attacks/defenses while increasing their effectiveness and strengths. The mechanism of attacks/defenses expands from intelligent agents acting humanly to thinking humanly.

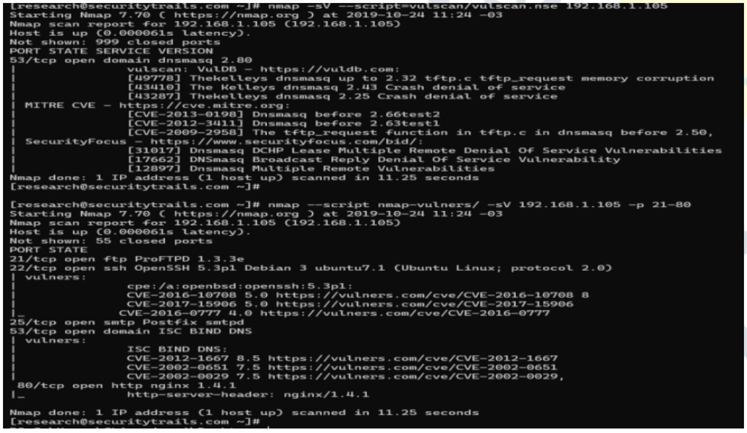
IV.Objective.

The main agenda of this tool is to ease the pentesting process a little by completing the enumeration and recon part of the process and this will also help in the final report writing. The tool aims to be better than other vulnerability scanners & enumeration tools. This is intended to automate processes with an increase in efficiency of the code via ai which can learn from its previous experience adding to it. This tool will be using a database that holds key information to increase its speed n efficiency. [3] The ai works on a system that each time it is presented with a difficult task it learns and stores the new info in its database which makes it a powerful tool over time and even adds to its efficiency. The Code primarily automates all the possible scans based on the info we initially provided and the AI reads the scans and starts to enumerate all the possible ways it can scan it will record all its findings in a temporary directory and then it analyzes the raw data to a more informative data and then enumerates the target based on the huge logic of various outcomes and inputs using information from API from websites containing huge amount Databases of CVE's or vulnerabilities[10]. Using the data it further finds specific payloads for that particular target and using Metasploit it will generate a resource file for further exploitation. The AI uses a complex algorithm of various outcomes and works on finding the smallest and best route to exploit once it has found the best route it writes that in its report for later use by the Pentester. Even though we have clear intel on how it works and we need to hard code it for it to be useful so there are 2 sages of coding this enormous tool.

V. Scope.

The tool uses a sequence of techniques to scan and gain more information on the target specified this can be better understood with the below example.

Scan Example:



```
[research@securitytrails.com ~]# nmap -sV --script=vulscan/vulscan.nse 192.168.1.105
Starting Nmap 7.70 ( https://nmap.org ) at 2019-10-24 11:24 -03
Nmap scan report for 192.168.1.105 (192.168.1.105)
Host is up (0.000061s latency).
Not shown: 999 closed ports
PORT STATE SERVICE VERSION
53/tcp open domain dnsmasq 2.80
|               vulscan: VulDB - https://vuldb.com:
|               [49778] Thekelleys dnsmasq up to 2.32 tftp.c tftp_request memory corruption
|               [43410] The Kelleys dnsmasq 2.43 Crash denial of service
|               [43287] Thekelleys dnsmasq 2.25 Crash denial of service
| MITRE CVE - https://cve.mitre.org:
|               [CVE-2013-0198] Dnsmasq before 2.66test2
|               [CVE-2012-3411] Dnsmasq before 2.63test1
|               [CVE-2009-2958] The tftp_request function in tftp.c in dnsmasq before 2.50,
| SecurityFocus - https://www.securityfocus.com/bid/:
|               [31017] Dnsmasq DCHP Lease Multiple Remote Denial Of Service Vulnerabilities
|               [17662] DNSmasq Broadcast Reply Denial Of Service Vulnerability
|               [12897] Dnsmasq Multiple Remote Vulnerabilities
Nmap done: 1 IP address (1 host up) scanned in 11.25 seconds
[research@securitytrails.com ~]#

[research@securitytrails.com ~]# nmap --script nmap-vulners/ -sV 192.168.1.105 -p 21-80
Starting Nmap 7.70 ( https://nmap.org ) at 2019-10-24 11:24 -03
Nmap scan report for 192.168.1.105 (192.168.1.105)
Host is up (0.000061s latency).
Not shown: 55 closed ports
PORT STATE
21/tcp open ftp ProFTPD 1.3.3e
22/tcp open ssh OpenSSH 5.3p1 Debian 3 ubuntu7.1 (Ubuntu Linux; protocol 2.0)
| vulners:
|               cpe:/a:openbsd:openssh:5.3p1:
|               CVE-2016-10708 5.0 https://vulners.com/cve/CVE-2016-10708 8
|               CVE-2017-15906 5.0 https://vulners.com/cve/CVE-2017-15906
|_             CVE-2016-0777 4.0 https://vulners.com/cve/CVE-2016-0777
25/tcp open smtp Postfix smtpd
53/tcp open domain ISC BIND DNS
| vulners:
|               ISC BIND DNS:
|               CVE-2012-1667 8.5 https://vulners.com/cve/CVE-2012-1667
|               CVE-2002-0651 7.5 https://vulners.com/cve/CVE-2002-0651
|               CVE-2002-0029 7.5 https://vulners.com/cve/CVE-2002-0029,
 80/tcp open http nginx 1.4.1
|_              http-server-header: nginx/1.4.1

Nmap done: 1 IP address (1 host up) scanned in 11.25 seconds
[research@securitytrails.com ~]#
```

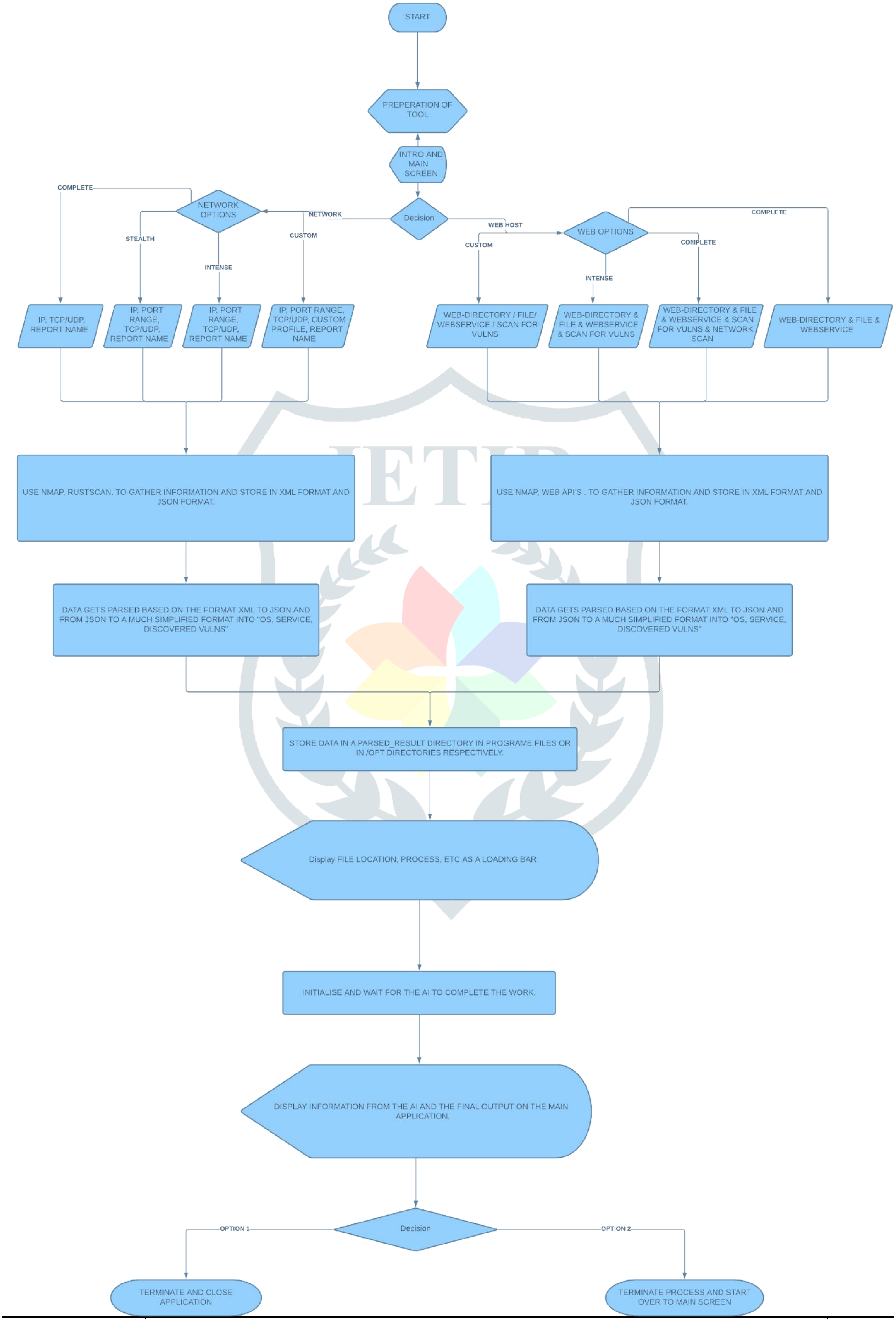
In the above example, the pentester is scanning the target with a tool known as nmap here he can check for the open ports with the help of the TCP request mechanism.[12] The TCP packets if successful then the ports are open and using the same method the tool can extract the version and service that's been running on the port. In the above example, the pentester is using a script that is implemented to checks for the service and version of the service with a few online database APIs and is returning the corresponding CVE or vulnerabilities related to it with its respective intensities.[4] The tool proposed is using these functionalities to perform such scans and attacks based on these TCP and UDP network scans to determine various information related to the target and also by reducing the complexity of the output into a more readable format.

VI. Limitations of existing System:

The main disadvantage is that the tools are
Nonautomated, Needs manual data reading and parsing. The complexity of use.
which makes this even more difficult to complete. The scanners and enumerators are really good but separated by various processes making them even slower and leading to a lot of issues.[11] We tend to eliminate this issue using our research.

CODE MAP

https://www.mindmeister.com/map/2106059128?t=MR3PSC8ATu


START
PREPERATION OF TOOL
INTRO AND MAIN SCREEN
Decision
NETWORK
WEB HOST
NETWORK OPTIONS
COMPLETE
STEALTH
INTENSE
CUSTOM
IP, TCP/UDP, REPORT NAME
IP, PORT RANGE, TCP/UDP, REPORT NAME
IP, PORT RANGE, TCP/UDP, REPORT NAME
IP, PORT RANGE, TCP/UDP, CUSTOM PROFILE, REPORT NAME
WEB OPTIONS
CUSTOM
INTENSE
COMPLETE
COMPLETE
WEB-DIRECTORY / FILE/ WEBSERVICE / SCAN FOR VULNS
WEB-DIRECTORY & FILE & WEBSERVICE & SCAN FOR VULNS
WEB-DIRECTORY & FILE & WEBSERVICE & SCAN FOR VULNS & NETWORK SCAN
WEB-DIRECTORY & FILE & WEBSERVICE
USE NMAP, RUSTSCAN. TO GATHER INFORMATION AND STORE IN XML FORMAT AND JSON FORMAT.
USE NMAP, WEB API'S . TO GATHER INFORMATION AND STORE IN XML FORMAT AND JSON FORMAT.
DATA GETS PARSED BASED ON THE FORMAT XML TO JSON AND FROM JSON TO A MUCH SIMPLIFIED FORMAT INTO "OS, SERVICE, DISCOVERED VULNS"
DATA GETS PARSED BASED ON THE FORMAT XML TO JSON AND FROM JSON TO A MUCH SIMPLIFIED FORMAT INTO "OS, SERVICE, DISCOVERED VULNS"
STORE DATA IN A PARSED_RESULT DIRECTORY IN PROGRAME FILES OR IN /OPT DIRECTORIES RESPECTIVELY.
Display FILE LOCATION, PROCESS, ETC AS A LOADING BAR
INITIALISE AND WAIT FOR THE AI TO COMPLETE THE WORK.
DISPLAY INFORMATION FROM THE AI AND THE FINAL OUTPUT ON THE MAIN APPLICATION.
Decision
OPTION 1
OPTION 2
TERMINATE AND CLOSE APPLICATION
TERMINATE PROCESS AND START OVER TO MAIN SCREEN

AI METHODOLOGY

The Ai algorithm and neural network we are thinking to be implemented must be capable of comparing, deciding, and ranking the input values based on the critical level by comparing the JSON value found positive by the scans it should work based on the scenario:

If the value is opened, vulnerable, or found a CVE go ahead or else stop. Compare the Data with the data inside the vulnerable database and assign a rank based on criticality. If the comparison goes well assign a rank-wise system and then create a pie chart based on the number of critical points and present a graph. After comparison assigns all the data according to the format and generates a pdf report.

The AI must be fast and must also incorporate the shortest and best route methodology by implementing these key factors. We can reduce the amount of wasted data in the process we have to manage to create a specific kind of database that holds highly repeating vulnerabilities. In the long run and implement the database as a primary source much like a primary memory making this even faster and less power draining.

The AI is a really important aspect of the entire project as it plays the main role in deciding the accuracy of the scan and if the accuracy is of then the scans results is close to being of no use and can be a total waste of time. By hard coding the various necessary aspects that make a penetration test scan accurate we can make the AI be more precise and also by maintaining a separate rule that makes the AI know how to improve we can ensure that the AI learns a new shortcut or trick to make the scan faster and accurate.

The AI in mind that we decide for these works are:

1) Decision tree algorithm: Uses a binary tree structure with nodes linking to each other the more the nodes or more the relations the more accurate the result is.
2) Random Forests algorithm: Uses a series of decision trees with recurring searches and analysis the more the trees the more accurate the result gets and is based on the “Majority Wins Model”.
3) Naïve Bayes algorithm: Based on probability and the bayes theory it finds the most likely output and also gives out the closest possible output.
4) Neural Network: Based on the human brain it uses several layers and also gives out a accurate output after going through its multiple layers and searching for the best output.

Based on the accuracy needed we can use the above algorithms and theories to find the most accurate result as all of the above have significance in the matters of complete analysis, regression and also output.

VII.CONCLUSION

In conclusion we can tell that using the data analyzed it is possible to create a well-maintained program that can automate cybersecurity enumeration we can implement the above information to achieve more than this in the future it just needs a bit more modifications there are possibilities that we can completely automate a pen testing process but it may take extensive testing and modifications but it is possible.

As of now there is a clear possibility that an advanced enumeration system can be introduced in the field with the above information and some more detailed analysis and further proofing this can help improve the service provided to a more greater extent.

REFERENCES:

1. Yampolskiy, R. V., & Spellchecker, M. S. (2016). Artificial intelligence safety and cybersecurity: A timeline of AI failures. arXiv preprint arXiv:1610.07997.

2. Morel, B. (2011, October). Artificial intelligence and the future of cybersecurity. In Proceedings of the 4th ACM workshop on Security and artificial intelligence (pp. 93-98).

3. Sarker, I. H., Furhad, M. H., & Nowrozy, R. (2021). Ai-driven cybersecurity: an overview, security intelligence modeling and research directions. SN Computer Science, 2(3), 1-18.

4. Chan, L., Morgan, I., Simon, H., Alshabanat, F., Ober, D., Gentry, J., ... & Cao, R. (2019, June). Survey of AI in cybersecurity for information technology management. In 2019 IEEE technology & engineering management conference (TEMSCON) (pp. 1-8). IEEE.

5. Wirkuttis, N., & Klein, H. (2017). Artificial intelligence in cybersecurity. Cyber, Intelligence, and Security, 1(1), 103-119.

6. Darraj, E., Sample, C., & Justice, C. (2019, July). Artificial intelligence cybersecurity framework: Preparing for the here and now with ai. In ECCWS 2019 18th European Conference on Cyber Warfare and Security (p. 132). Academic Conferences and publishing limited.

7. Zeadally, S., Adi, E., Baig, Z., & Khan, I. A. (2020). Harnessing artificial intelligence capabilities to improve cybersecurity. Ieee Access, 8, 23817-23837.

8. Singer, P. W., & Friedman, A. (2014). *Cybersecurity: What everyone needs to know*. oup usa.

9. Kemmerer, R. A. (2003, May). Cybersecurity. In *25th International Conference on Software Engineering, 2003. Proceedings.* (pp. 705-715). IEEE.

10. Sarker, I. H., Kayes, A. S. M., Badsha, S., Alqahtani, H., Watters, P., & Ng, A. (2020). Cybersecurity data science: an overview from machine learning perspective. *Journal of Big data*, *7*(1), 1-29.

11. Martínez Torres, J., Iglesias Comesaña, C., & García-Nieto, P. J. (2019). Machine learning techniques applied to cybersecurity. *International Journal of Machine Learning and Cybernetics*, *10*(10), 2823-2836.

12. Ten, C. W., Manimaran, G., & Liu, C. C. (2010). Cybersecurity for critical infrastructures: Attack and defense modeling. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, *40*(4), 853-865.