



#### **Exercise Structure**

- OTCEP25 folder on Windows VM desktop with the necessary exercise files
- PROMPT.txt
  - Contains the command to be copy-pasted into the Chat Panel
- COMMANDS.txt
  - Contains the command to be copy-pasted into the terminal after the generated code have been accepted
- ANSWER.txt
  - Contains an example of the code generated during preparation using the same prompt within PROMPT.txt





#### **Disclaimer**

- GenAl produces non-deterministic output
  - Same prompt (with same context) may not produce the same output
- If the code produced by prompt does not work:
  - Revert changes and try again with the same prompt,
  - Copy-paste the error message(s) into LLM to debug, or
  - Use the code provided in ANSWER.txt



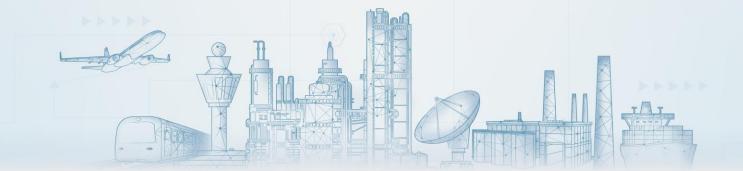


#### Scenario

- This workshop is a snapshot of a Purple Team Exercise with 2 exercises each for Attacks and Defenders
- Attacker
  - Leveraging GenAl to bypass AV signature detection
  - Leveraging GenAl to execute TTPs that blend into the background
- Defender
  - Leverage GenAl to supplement detection rules from open-sources repos
  - Leverage GenAl to improve and refine detection rules

# **Scenario 1: Payload Obfuscation**

Attacker





#### **AttackExercise 1: Motivation**

Recently, advancements in the code understanding capabilities of LLMs have raised concerns about criminals using LLMs to generate novel malware. Although LLMs struggle to create malware from scratch, criminals can easily use them to rewrite or obfuscate existing malware, making it harder to detect.

Adversaries have long used common obfuscation techniques and tools to avoid detection. We can easily fingerprint or detect off-the-shelf obfuscation tools because they are well known to defenders and produce changes in a predefined way However, criminals can prompt LLMs to perform transformations that are much more natural-looking, which makes detecting this malware more challenging

Source: https://unit42.paloaltonetworks.com/using-llms-obfuscate-malicious-javascript/
Date: Dec 2024

# Findings: Government-Backed Threat Actors Misusing Gemini

At a Glance: Government-Backed Attackers

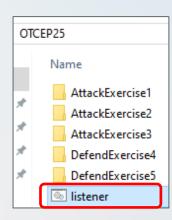
Government-backed attackers attempted to use Gemini for coding and scripting tasks, gathering information about potential targets, researching publicly known vulnerabilities, and enabling post-compromise activities, such as defense evasion in a target environment.

Source: https://cloud.google.com/blog/topics/threat-intelligence/adversarial-misuse-generative-ai Date: Jan 2025



### AttackExercise 1: Set Up

- Listener that simulates attacker Command & Control
   (C2) infrastructure (listener.bat)
  - A batch script that simulates the attacker's command & control server
  - Uses `ncat.exe` to listen on port 1234 for incoming connections

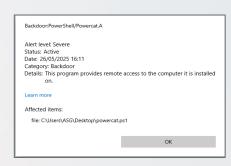




### **AttackExercise 1: Thought Process**

- PowerCat is a PowerShell script allows an attacker to connect a victim machine to his system
  - Able to execute code on victim machine <u>remotely</u>
- The script has been signatured and is being detected by Win Defender
- Attacker needs to obfuscate the code in order to execute it







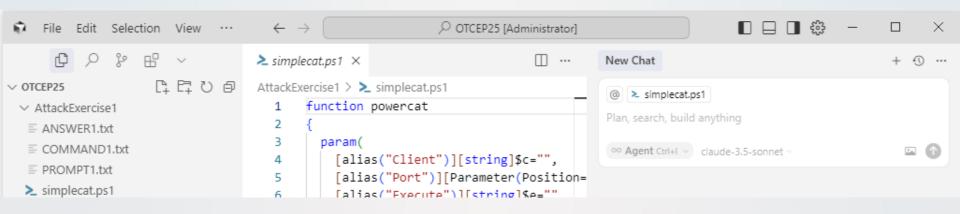
#### **AttackExercise 1: Obfuscation Introduction**

- What is obfuscation?
  - A technique hackers use to disguise malicious code (the "Payload") to sneak it past Antivirus software.
  - Imagine writing a harmful message in a secret code (Obfuscation) so security guards can't read it.
- Examples of obfuscation:
  - Renaming function and variable names
  - String obfuscation



#### **AttackExercise 1: Execution**

 Task: Use Cursor to modify and obfuscate the existing code, target file is simplecat.ps1 in the AttackExercise1 folder





#### **AttackExercise 1: Execution**

- Copy the prompt in PROMPT1.txt to the agent chat window and hit enter
- The obfuscated code will be "added" to a new file named Exercise1.ps1
- Click 'Accept file' to save changes made in Exercise1.ps1

```
Rename all parameters and function names to short random alphanumeric chars and save it to file Exercise1.ps1 in the same folder. Use xyz as the function name, a for client mode, b for port and c for execute.
```



### **AttackExercise 1: Prompt Engineering**

Rename all parameters and function names to short random alphanumeric chars and save it to file Exercise1.ps1 in the same folder.

Use  $\underline{xyz}$  as the function name, a for client mode, b for port and c for execute.

Specific instruction to rename parameters and function names

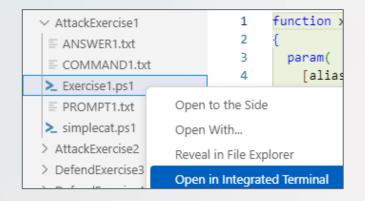
More consistent output than generic 'obfuscate' instruction

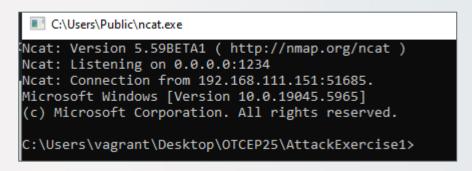
▼ For consistency in this exercise
 Not necessary for actual engagements



#### **AttackExercise 1: Execution**

- Run listener.bat, if it is not running yet
- Right click Exercise1.ps1 and click "Open in Integrated Terminal"
  - Copy the text in COMMAND1.TXT into the Terminal and hit enter
- Observe a shell connection has been made to the attacker listener.bat





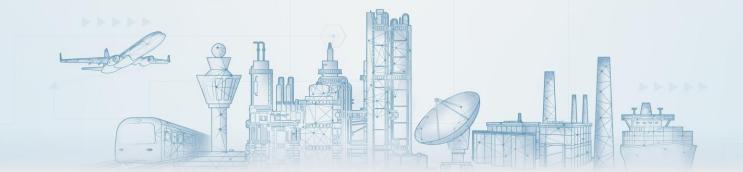


#### **AttackExercise 1: Conclusion**

- New obfuscated code was successful in establishing a reverse shell
- Code is no longer detected by Windows Defender (copy out to Desktop to test)
  - Test out the evasion by copying the file to Desktop and see if it gets flagged
  - Renaming function and variable names worked, WinDef was detecting malicious code based on this
- Defender team is now able to hunt for the payload behaviour and create alert rules

# **Scenario 2: Payload Conversion**

Attacker





#### **AttackExercise 2: Motivation**



Code metamorphism refers to a computer programming exercise wherein the program modifies its own code (partial or entire) consistently and automatically while retaining its core functionality. This technique is often

This paper explores the potential of several prominent LLMs for software code mutation that may be used to reconstruct (with mutation) existing malware code bases or create new forms of embedded mutation engines for next-gen metamorphic malwares. In this work, we introduce a framework for creating self-testing program



# **AttackExercise 2: Payload Objective**

#### Evading AV

- Most AV engines are optimized to detect threats written in widely abused languages like PowerShell, C#, and Python.
- Using less commonly analyzed languages or runtimes, such as JavaScript (Node.js), can help evade detection by avoiding high-sensitivity code paths.
- Blending into the environment
  - Payloads executed through legitimate interpreters (e.g. node.exe) often look like normal activity.
  - This makes it challenging for defenders to differentiate malicious behavior from normal activity.



### **AttackExercise 2: Thought Process**

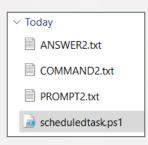
- Attacker has gotten a foothold onto the system in Exercise 1, now needs a way to maintain access even through system reboots – Persistence Technique
- A common method is by using Windows Scheduled Tasks. A scheduled task has 3 main parts:
  - Trigger When should it run? (e.g., at 11 PM, on startup)
  - Action What should it do? (e.g., run a program or command)
  - Conditions/Settings Optional rules (e.g., only if plugged in)

Name Status Triggers		Triggers	s	Next Run Time	
MicrosoftEd Ready Multip		Multipl	e triggers defined	19/7/2025 10:17:14 am	
	Action		Details		
	Start a program		$\label{lem:capprox} \hbox{C:\Program Files (x86)\Microsoft\EdgeUpdate\MicrosoftEdgeUpdate.exe/c} \\$		



### **AttackExercise 2: Thought Process**

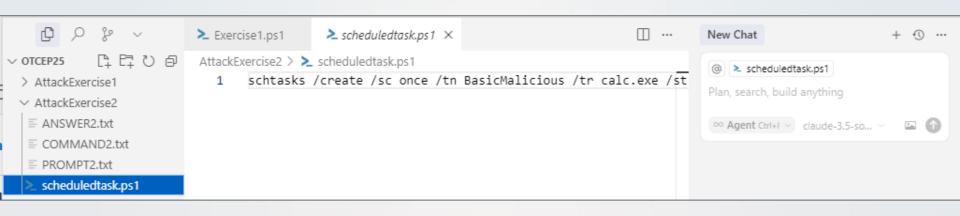
- Attacker has obtained a PowerShell script to create the scheduled task
- Knows that PowerShell logging was enabled on the target system and wants to tests the defences by choosing a stealthier method
  - Notices that Node is installed on the system, which is less common attack vector
  - Needs to find way to run through NodeJS





#### **AttackExercise 2: Execution**

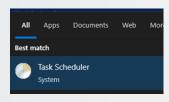
 Task: Use Cursor to convert the existing PowerShell script to be executable in NodeJS, target file is scheduledtask.ps1 in the AttackExercise2 folder

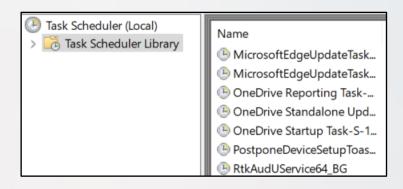




#### **AttackExercise 2: Execution**

- Search for Task Scheduler from Start button and open it
- Observe that there isn't any Malicious tasks currently



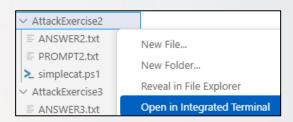




#### **AttackExercise 2: Execution**

- Copy the prompt in PROMPT2.txt to the agent chat window and hit enter
- New code written in NodeJS will be saved to Exercise2.js
- Right click Exercise2.js and click "Open in Integrated Terminal"
  - Copy the text in COMMAND2.TXT into the Terminal and hit enter





PS C:\Users\vagrant\Desktop\OTCEP25\AttackExercise2> node .\Exercise2.js
Task created successfully:
SUCCESS: The scheduled task "BasicMalicious" has successfully been created.



# **AttackExercise 2: Prompt Engineering**

Rewrite the logic of a PowerShell using Node.js instead of PowerShell.

Convert this PowerShell script to a NodeJS script and save as Exercise2.js in this folder.

The script should not require any further installation and

should only consist of a single file.

Ensure the script can run on any machine with Node.js, without additional setup

No splitting code into modules or using external files. This makes the script easy to run, review, and deploy

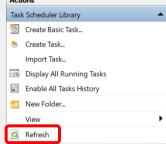


#### **AttackExercise 2: Conclusion**

- Refresh the task scheduler and look for BasicMalicious task
- Attacker was able to recreate the Persistence technique in a different attack vector
  - PowerShell logging will not be useful for detection
- Blue teams will now be able to determine if there is sufficient telemetry to

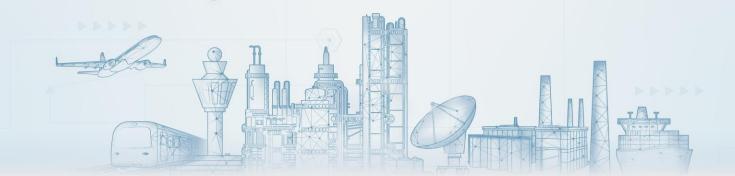
detect this type of attack

Name	Status	Triggers	Next Run Time	Last Run Time
BasicMalicious	Ready	At 11:59 PM on 6/30/2025	6/30/2025 11:59:00 PM	11/30/1999 12:00:00 AM



# **Scenario 3: Rule Conversion**

Defender





# **DefendExercise 3: Sigma Rules**

- Sigma is an open standard for writing detection rules in a generic, readable format (YAML).
- There is a repository of Sigma rules available opensource
  - https://github.com/SigmaHQ/sigma
  - Hundreds of pre-built rules for Windows, Linux, cloud, and application logs.
- Browse and select from the available rules to match your needs

```
title: Suspicious Scheduled Task Creation via schtasks.exe
id: a1b2c3d4-e5f6-7890-abcd-ef1234567890
status: experimental
description: Detects creation of scheduled tasks using schtask
    - https://attack.mitre.org/techniques/T1053/005/
author: Security Analyst
date: 2024/01/01
modified: 2024/01/01

    attack.persistence

    - attack.privilege_escalation
    - attack.t1053.005
logsource:
    product: windows
    service: security
detection:
    selection:
        CommandLine|contains|all:
            - 'schtasks'
            - '/create'
    filter privs:
        CommandLine|contains:
            - 'NT AUT'
            - ' SYSTEM'
            - 'HIGHEST'
    condition; all of selection and not filter privs
falsepositives:
    - Legitimate administrative tasks
    - System maintenance scripts
    - Software installations
 evel: medium
```



# **DefendExercise 3: Sigma Rules**

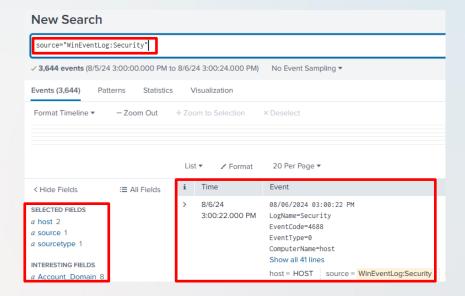
- Sigma repo contains a rule related to suspicious scheduled task creation
- Defender system is using Splunk as the SIEM, search query needs to be in SPL
- Conversion from Sigma to SPL needs to be done for it to be useful

```
title: Suspicious Scheduled Task Creation via schtasks.exe
id: a1b2c3d4-e5f6-7890-abcd-ef1234567890
status: experimental
description: Detects creation of scheduled tasks using schtask
   - https://attack.mitre.org/techniques/T1053/005/
author: Security Analyst
date: 2024/01/01
modified: 2024/01/01
   - attack.persistence
   - attack.privilege_escalation
   - attack.t1053.005
logsource:
   product: windows
   service: security
detection:
        CommandLine|contains|all:
            - 'schtasks'
           - '/create'
   filter privs:
        CommandLine|contains:
            - 'NT AUT'
            - ' SYSTEM'
           - 'HIGHEST'
   condition; all of selection and not filter privs
falsepositives:
    - Legitimate administrative tasks
   - System maintenance scripts
    - Software installations
evel: medium
```



# **DefendExercise 3: Splunk**

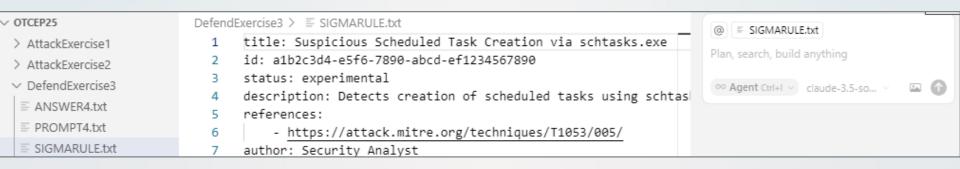
- Search box to enter your query
  - Searches through various log sources
- Results consist of the events returned based on the search query
  - Fields in the left column can be added to the search query to finetune it
- Look through the events to confirm if the event is a legitimate attack





#### **DefendExercise 3: Execution**

Task: Use Cursor to convert the selected Sigma rule to SPL, target file is
 SIGMARULE.TXT in the DefendExercise3 folder





#### **DefendExercise 3: Execution**

- Copy the prompt in PROMPT3.txt to the agent chat window and hit enter
- New SPL query will be saved to Exercise3.txt
- Access Splunk in the Host VM Edge favorites bar, click Search & Reporting
  - Username: admin, Pass: password123

#### SIGMARULE.txt

Convert the Sigma rule into an efficient Splunk Processing Language (SPL) query using main as the index.

Convert `CommandLine` -> `Message`.

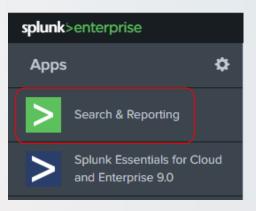
Translate the `detection` logic from the Sigma rule into an optimized SPL query using the inferred context and the required field mapping.c".

Do not add comments.

Save the output to Exercise3.txt in this folder folder.

Do not refer to any other files.







# **DefendExercise 3: Prompt Engineering**

Convert the Sigma rule into an efficient / Splunk Processing Language (SPL) query using main as the index.

Convert `CommandLine` -> `Message`.

Translate the `detection` logic from the Sigma rule into an optimized SPL query using the inferred context and the required field mapping".

Task cursor to do the conversion, specifying main as our index (based on our known setup)

We know that the CommandLine in Sigma maps to Message in SPL

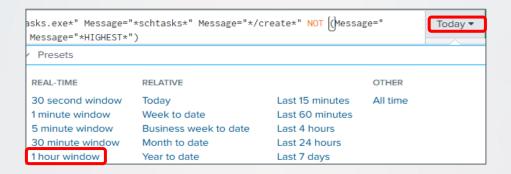
Get Cursor to read through the Sigma rule and try to map to SPL correctly



# **DefendExercise 3: Threat Hunting**

- Copy in the SPL query from Exercise3.txt to the search bar
- Set limit to past hour

# New Search index=\* EventCode=4688 Message="\*\\schtasks.exe\*" Message="\*schtasks\*" Message="\*/create\*" NOT (Message="\*NT AUT\*" OR Message="\* SYSTEM\*" OR Message="\*HIGHEST\*")





# **DefendExercise 3: Threat Hunting**

Look in the results for BasicMalicious Task

>	6/30/25 10:27:30.000 AM	06/30/2025 10:27:30 AM LogName=Security EventCode=4688 EventType=0 ComputerName=host			
			Process	Information:	
				New Process ID:	0xcdc
				New Process Name:	C:\Windows\System32\schtasks.exe
				Token Elevation Type:	%%1936
				Mandatory Label:	S-1-16-12288
				Creator Process ID:	0x1584
				Creator Process Name:	C:\Windows\System32\WindowsPowerShell\v1.0\powershell.exe
				Process Command Line:	"C:\Windows\system32\schtasks.exe" /create /sc once /tn BasicMalicious /tr calc.exe /st 23:59
		Process Information: New Process ID: New Process Name: Token Elevation Type: Mandatory Label: Creator Process Name: Process Command Line:	S-1-16-12288 0x1584 C:\Windows\System32\Windo	sks.exe wsPowerShell\v1.0\powershell.exe asks.exe" /create /sc once /tn BasicMalicious /t	r calc eve (st 23-59

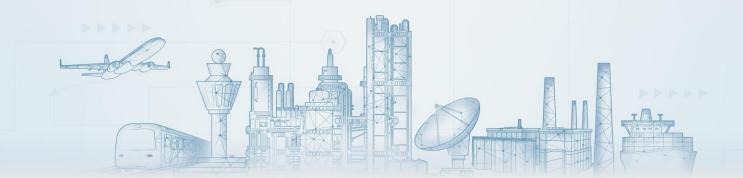


#### **DefendExercise 3: Conclusion**

- Defender was able to detect the creation of new scheduled task
- An open-source Sigma rule was used, utilizing crowdsourcing to speed up detection engineering
- Could be translated to others query language than SPL

# Scenario 4: Improve Rule

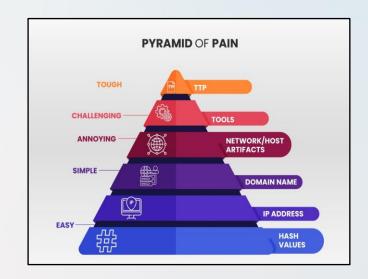
Defender





#### **DefendExercise 4: TTP**

- The Pyramid of Pain shows how hard it is for attackers to adapt when you detect different things
  - It is 'trivial' for attackers to get around detection mechanisms the lower it is on the pyramid
  - The higher you go on the pyramid, the more pain you cause attackers — and the more durable your detection becomes
- We will look at how we can leverage GenAl to 'level up' our detection strategy





## **DefendExercise 4: Preparation**

- In DefendExercise4 folder, right-click advancedtask.ps1 and click "Run with PowerShell"
- Refresh the Task Scheduler and observe that there are 2 tasks now





## **DefendExercise 4: Preparation**

- Re-run the detection rule in Splunk to look for any additional results
  - Observe that there isn't any result for AdvancedMalicious
- Defender observes that while there are 2 new tasks created in the system, the SPL rule is only able to detect 1
- Defender needs to produce a better rule that can detect both types of attack



### **DefendExercise 4: Execution**

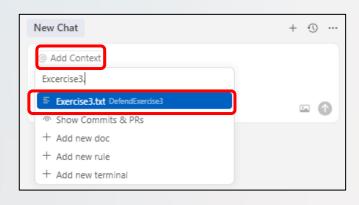
 Task: Use Cursor to generate a better SPL query, target file is Exercise3.txt in the DefendExercise3 folder

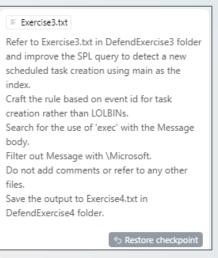




#### **DefendExercise 4: Execution**

- Copy the prompt in PROMPT4.txt to the agent chat window
- Ensure context is set to Exercise3.txt and hit enter
- New SPL query will be saved to Exercise4.txt







## **DefendExercise 4: Prompt Engineering**

Refer to Exercise3.txt in DefendExercise3 folder and improve the SPL query to detect a new scheduled task creation using main as the index.

Craft the rule based on event id for task creation rather than LOLBINs.

Search for the use of 'exec' with the Message body.

Filter out Message with \Microsoft.

Task cursor to find improvements to existing code rather than reinvent the wheel

Create the detection rule using specific Windows Event IDs rather than LOLBINs which are more generic

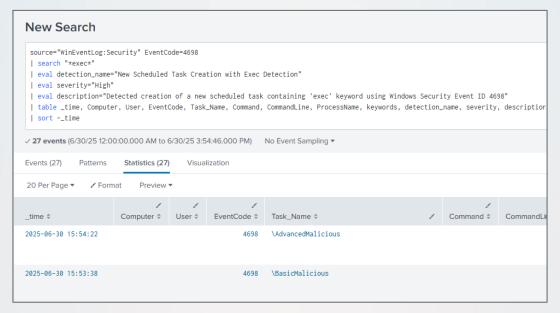
Identify command execution activity related to the task creation, giving more context or evidence of malicious intent

Remove false positives based on our environment baseline



#### **DefendExercise 4: Conclusion**

Copy the SPL query from Exercise4.txt to the search bar, set limit to past hour



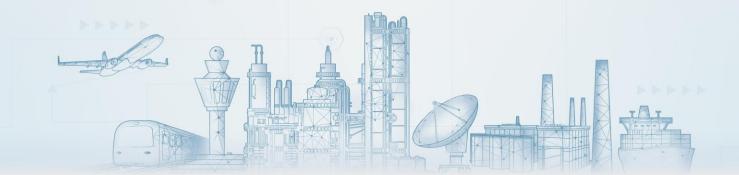


#### **DefendExercise 4: Conclusion**

- Previous detection from SIGMARULE.txt detects scheduled task by looking for execution of schtasks.exe
- AdvancedMalicious task was created by PowerShell's Register-ScheduledTask function instead and was therefore not detected
- New rule generated was based on the Task Creation EventID that captures any new task creation event, regardless of the method

# Bonus Scenario 5: Payload Enhancement

**Attacker** 





#### **ExtraExercise 5: Motivation**

- This scenarios is to show how 'supply chain attacks' can be carried out on open-sourced codebase.
- Attacker has access to an open-source GitHub project and wants to 'poison' it to run malicious code
  - Code will run cmd.exe to connect back to the Attacker controlled IP
  - This will give Attacker command-line access to the Victim machine



#### **ExtraExercise 5: Intended Behaviour**

- The project is a Node based TCP proxy that acts as a middleman between a client and a server, forwarding data between them
  - node tcp-proxy-cli.js -n 192.168.111.151 -s 1111 -p 2222



-p 2222: The proxy will listen on port 2222

-n 192.168.111.151: The proxy will forward traffic to localhost

-s 1111 : The proxy will forward traffic to port 1111 of the IP

above



## **ExtraExercise 5: Payload Enhancement**

 Task: Use Cursor to include code that will connect to our listener when the proxy is used, target file is tcp-proxy.js in the folder

```
OTCEP25
                           ExtraExercise5 > node-tcp-proxy-master > JS tcp-proxy.is > ...
                                                                                                       @ s tcp-proxy.js
                                  var net = require("net");
> AttackExercise1
                                  var tls = require('tls'):
> AttackExercise2
                                  var fs = require('fs');
> DefendExercise3
                                                                                                       ∞ Agent Ctrl+l ∨ claude-3.5-so... ∨ 🔼 🕜
                                  var util = require('util');
> DefendExercise4

∨ ExtraExercise5

                                  module.exports.createProxy = function(proxyPort,

∨ node-tcp-proxy-master

                                      serviceHost, servicePort, options) {
                                      return new TcpProxy(proxyPort, serviceHost, servicePort,
 > .github
                             9
  > examples
                            10
  > node modules
                                  function uniqueKey(socket) {
  eslintrc.ison
                            12
                                      var key = socket.remoteAddress + ":" + socket.remotePort;
  .aitianore
                            13
                                      return kev:
  a cert.cer
                            14
  ≡ cert.pfx
                            15
                            16
                                  function parse(o) {
  TS index.d.ts
                            17
                                      if (typeof o === "string") {
  JS index.is
                            18
                                          return o.split(",");
  19
                                       } else if (typeof o === "number") {
  {} package-lock.json
                            20
                                          return parse(o.toString()):
  {} package.json
                            21
                                      } else if (Array.isArray(o))
 (i) README.md
                                                                Dowershell + ∨ ∏ mm ··· ^ ×
                           Problems Output Terminal ...
  JS tcp-proxy-cli.js
                           PS C:\Users\vagrant\Desktop\OTCEP25>
```



#### **ExtraExercise 5: Execution**

- Copy the prompt in PROMPT5.txt to the agent chat window and hit enter
  - You might get denied by the agent due to malicious intent (built-in guardrails), try to bypass by changing models (GPT etc)
- Code changes will be made in the same tcp-proxy.js file, click Accept
- We can now test to see if the generated code works and gives us a shell
  - Re-run listener.bat again





## **ExtraExercise 5: Prompt Engineering**

Add a change to this TCP proxy code.

When the proxy starts, it should connect to localhost:1234 and pipe cmd.exe input/output through that connection.

If the connection fails, retry every 5 seconds.

Do not display anything to console output.

Don't add any new command line arguments - just hardcode the settings. —

Edit the existing code

Add in a feature to launch cmd.exe to the specific IP and Port

Mimics persistence and makes the shell activation more reliable

Maintain stealth, reducing the chances of being noticed by users

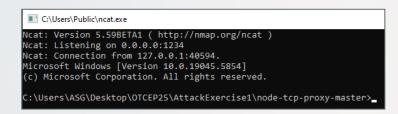
Do not ask user for input

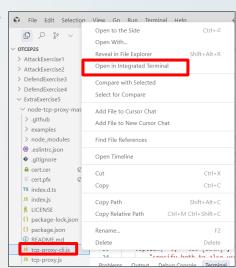


#### **ExtraExercise 5: Conclusion**

- Right click node-tcp-proxy-master folder and launch Terminal
- Run the command in COMMAND5.TXT
- Observe a reverse shell connection made to our listener

PS C:\Users\csa-asg\Downloads\Defender Excluded\OTCEP-2025\Mini-PpT-Infra\setup\setup\_files\OTCEP25\ExtraExercise5> node tcp-proxy-cli.js -n 192.168.111.151 -s 1111 -p 2222







#### **ExtraExercise 5: Conclusion**

- Why did the shell connect even though we ran the original command from the author?
  - The malicious code silently ran in the background while normal operation was carried out
- Defender can now hunt for the shell being spawned

