Cyber Threat Intelligence

Cyber Threat Modeling

How Large Language Models can help to classify Cyber Threats and give information about them ?

Baptiste Viera

FALL 2023

Table of contents

01

What is the problem in Cyber Threat Modeling ?

02

How to use LLM for Cyber Threat Modeling ?

03

What datasets are used ?

04

What are the implementation and the results ?

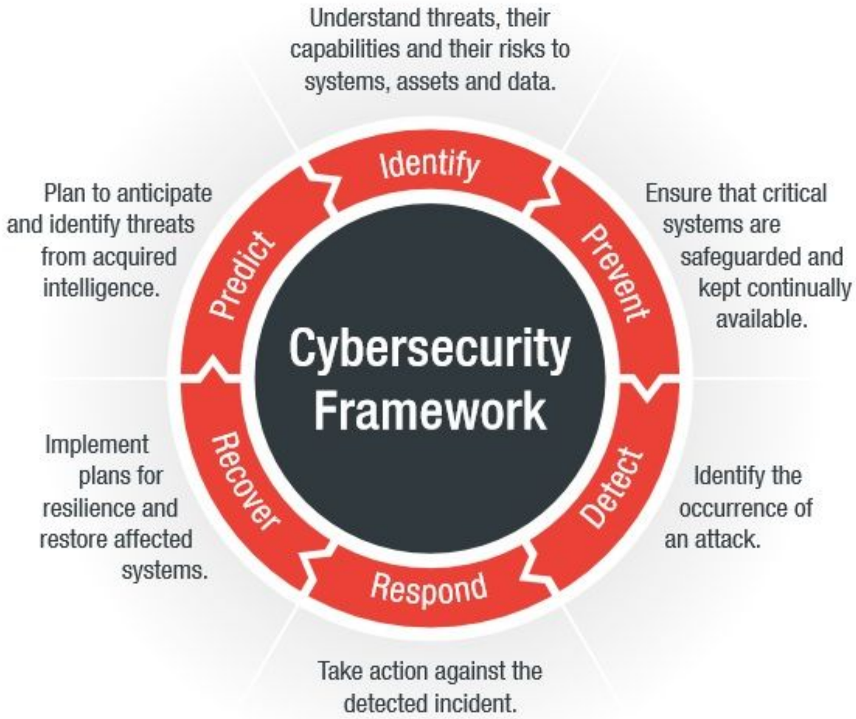
05

What are the Implications & Main Contributions ?

01

What is the problem in Cyber Threat Modeling ?

What is Cyber Threat Modeling ?



Understand threats, their capabilities and their risks to systems, assets and data.

Ensure that critical systems are safeguarded and kept continually available.

Identify the occurrence of an attack.

Take action against the detected incident.

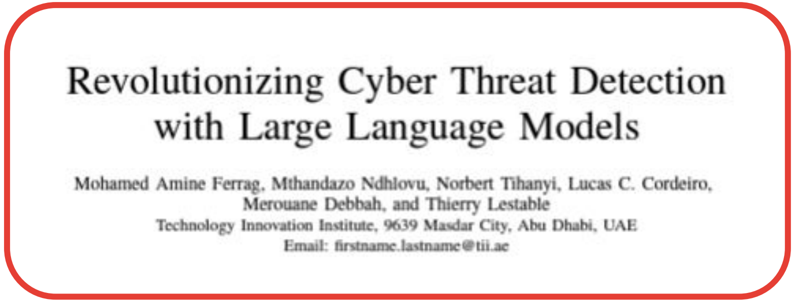
Implement plans for resilience and restore affected systems.

Plan to anticipate and identify threats from acquired intelligence.

What is the problem in Cyber Threat Modeling ?

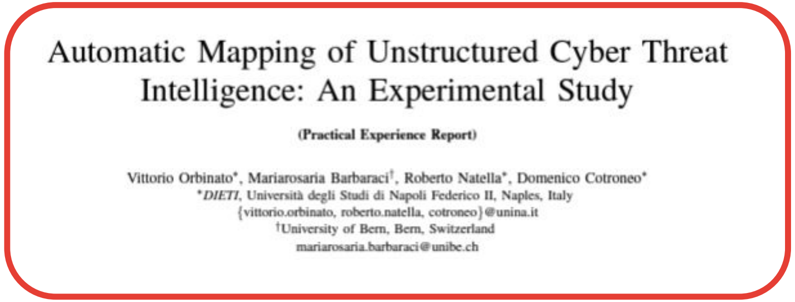
- It's difficult to identify and map cyber threats with adversary techniques from Mitre Attack
- It's difficult to get reliable information about detection or mitigation for instance
- It could be dangerous to rely on Large Language Model (LLM) Application connected to internet (prompt injection)
- Not having a LLM connected to an up to date knowledge base could be dangerous too (hallucination)

Related papers



Revolutionizing Cyber Threat Detection with Large Language Models

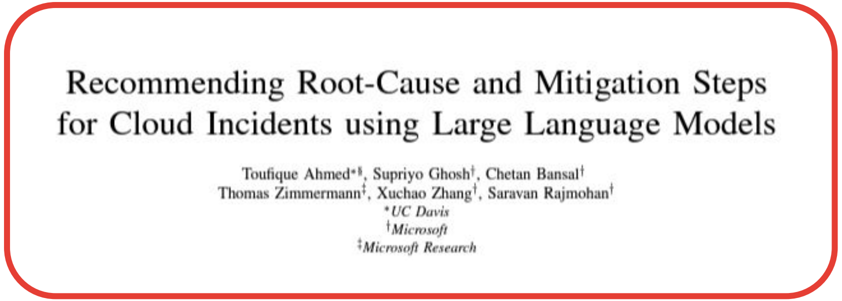
Mohamed Amine Ferrag, Mthandazo Ndhlovu, Norbert Tihanyi, Lucas C. Cordeiro,
Merouane Debbah, and Thierry Lestable
Technology Innovation Institute, 9639 Masdar City, Abu Dhabi, UAE
Email: firstname.lastname@tii.ae



Automatic Mapping of Unstructured Cyber Threat Intelligence: An Experimental Study

(Practical Experience Report)

Vittorio Orbinato[*], Mariarosaria Barbaraci[†], Roberto Natella[*], Domenico Cotroneo[*]
[*]*DIETI*, Università degli Studi di Napoli Federico II, Naples, Italy
{vittorio.orbinato, roberto.natella, cotroneo}@unina.it
[†]University of Bern, Bern, Switzerland
mariarosaria.barbaraci@unibe.ch



Recommending Root-Cause and Mitigation Steps for Cloud Incidents using Large Language Models

Toufique Ahmed[*§], Supriyo Ghosh[†], Chetan Bansal[†]
Thomas Zimmermann[‡], Xuchao Zhang[†], Saravan Rajmohan[†]
[*]*UC Davis*
[†]*Microsoft*
[‡]*Microsoft Research*

02

How to use LLM for Cyber Threat Modeling ?

How to use LLM for Cyber Threat Modeling ?

Use general LLM (GPT3.5, GPT4, Command...)

Use fine-tuned LLM

Use LLM with a specific knowledge base

How to use LLM for Cyber Threat Modeling ?

Research Questions

RQ1: To what extent could fine-tuned LLM improve the modeling of cyber threats?

RQ2: How reliable is Retrieval Augmented Generation (RAG) with a vector and/or graphical knowledge database for modeling cyberthreats?

RQ3: Does prompt engineering help to model cyber threat?

03

What datasets are used ?

What datasets are used ?

MITRE ATT&CK™

ATOMIC READ TEAM

Sigma

SIEM Detection Format

MITRE ENGENUITY™

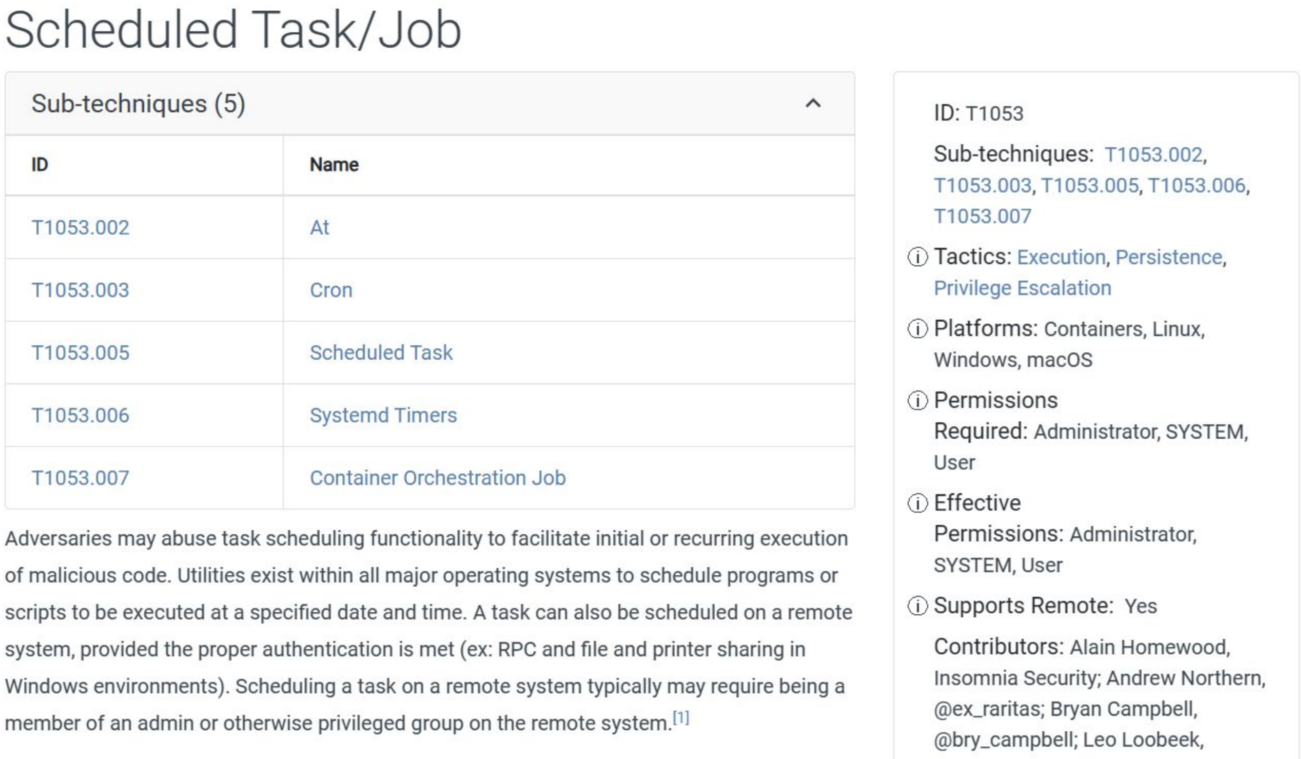
A Foundation for Public Good

Dataset: MITRE ATT&CK

- Tactics : **14**
- Techniques : **201**
- Subtechniques : **424**
- Datasources: **41**

- Software: **760**
- Campaigns: **24**
- Groups : **143**

Dataset: MITRE ATT&CK



Scheduled Task/Job

Sub-techniques (5)

ID	Name
T1053.002	At
T1053.003	Cron
T1053.005	Scheduled Task
T1053.006	Systemd Timers
T1053.007	Container Orchestration Job

Adversaries may abuse task scheduling functionality to facilitate initial or recurring execution of malicious code. Utilities exist within all major operating systems to schedule programs or scripts to be executed at a specified date and time. A task can also be scheduled on a remote system, provided the proper authentication is met (ex: RPC and file and printer sharing in Windows environments). Scheduling a task on a remote system typically may require being a member of an admin or otherwise privileged group on the remote system.[1]

ID: T1053

Sub-techniques: T1053.002, T1053.003, T1053.005, T1053.006, T1053.007

ⓘ Tactics: Execution, Persistence, Privilege Escalation

ⓘ Platforms: Containers, Linux, Windows, macOS

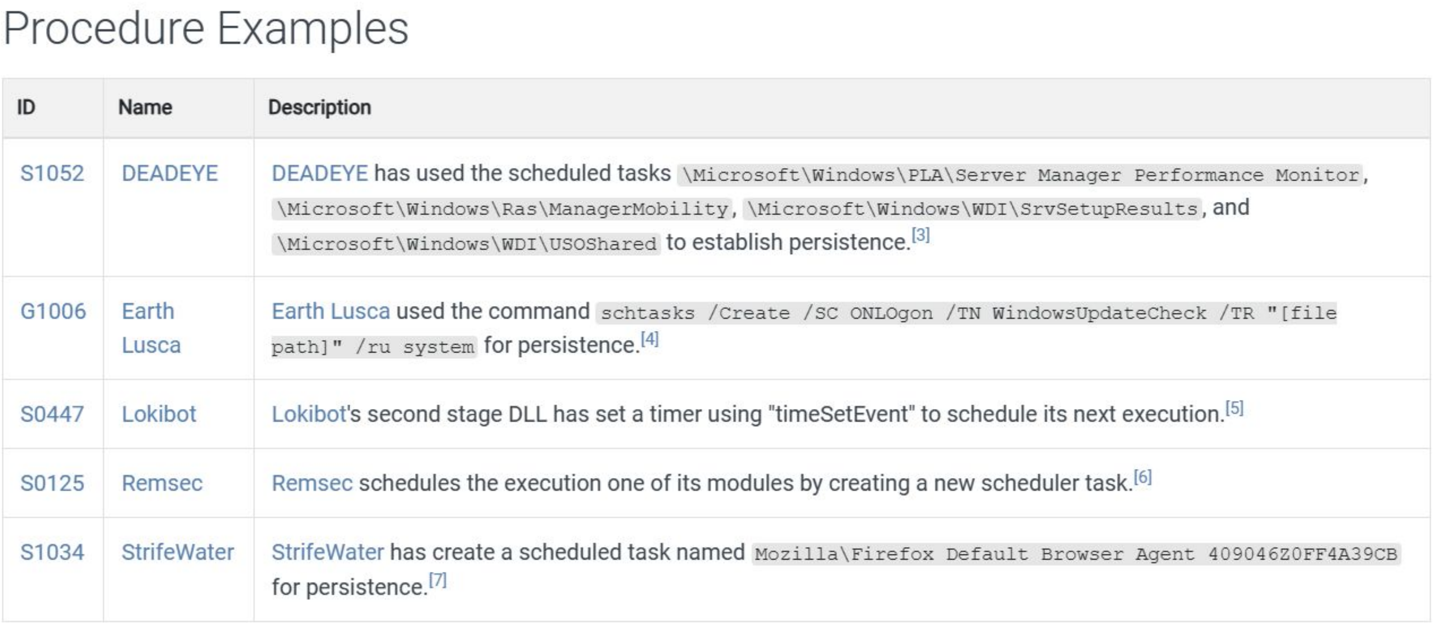
ⓘ Permissions Required: Administrator, SYSTEM, User

ⓘ Effective Permissions: Administrator, SYSTEM, User

ⓘ Supports Remote: Yes

Contributors: Alain Homewood, Insomnia Security; Andrew Northern, @ex_raritas; Bryan Campbell, @bry_campbell; Leo Loobeek,

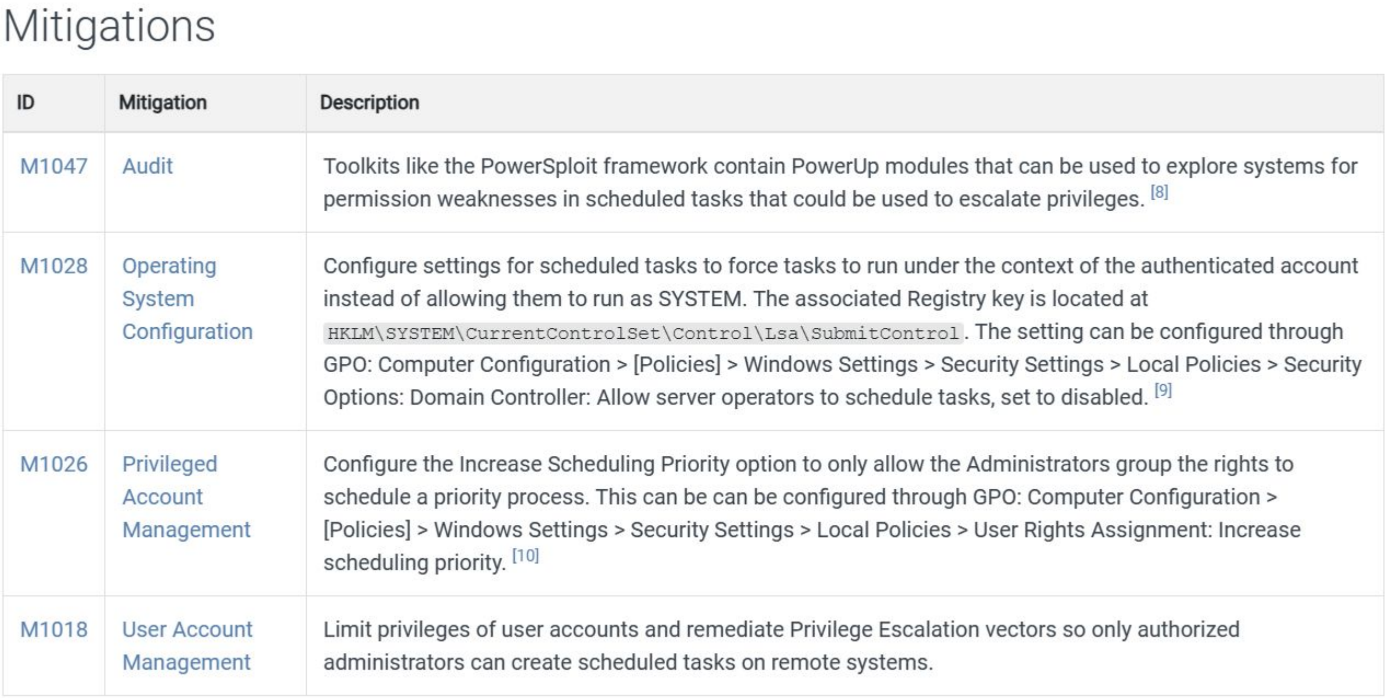
Dataset: MITRE ATT&CK



Procedure Examples

ID	Name	Description
S1052	DEADEYE	DEADEYE has used the scheduled tasks `\Microsoft\Windows\PLA\Server Manager Performance Monitor`, `\Microsoft\Windows\Ras\ManagerMobility`, `\Microsoft\Windows\WDI\SrvSetupResults`, and `\Microsoft\Windows\WDI\USOShared` to establish persistence.[3]
G1006	Earth Lusca	Earth Lusca used the command `schtasks /Create /SC ONLOgon /TN WindowsUpdateCheck /TR "[file path]" /ru system` for persistence.[4]
S0447	Lokibot	Lokibot's second stage DLL has set a timer using "timeSetEvent" to schedule its next execution.[5]
S0125	Remsec	Remsec schedules the execution one of its modules by creating a new scheduler task.[6]
S1034	StrifeWater	StrifeWater has create a scheduled task named `Mozilla\Firefox Default Browser Agent 409046Z0FF4A39CB` for persistence.[7]

Dataset: MITRE ATT&CK



Mitigations

ID	Mitigation	Description
M1047	Audit	Toolkits like the PowerSploit framework contain PowerUp modules that can be used to explore systems for permission weaknesses in scheduled tasks that could be used to escalate privileges. [8]
M1028	Operating System Configuration	Configure settings for scheduled tasks to force tasks to run under the context of the authenticated account instead of allowing them to run as SYSTEM. The associated Registry key is located at `HKLM\SYSTEM\CurrentControlSet\Control\Lsa\SubmitControl`. The setting can be configured through GPO: Computer Configuration > [Policies] > Windows Settings > Security Settings > Local Policies > Security Options: Domain Controller: Allow server operators to schedule tasks, set to disabled. [9]
M1026	Privileged Account Management	Configure the Increase Scheduling Priority option to only allow the Administrators group the rights to schedule a priority process. This can be can be configured through GPO: Computer Configuration > [Policies] > Windows Settings > Security Settings > Local Policies > User Rights Assignment: Increase scheduling priority. [10]
M1018	User Account Management	Limit privileges of user accounts and remediate Privilege Escalation vectors so only authorized administrators can create scheduled tasks on remote systems.

Dataset: ATOMIC RED TEAM

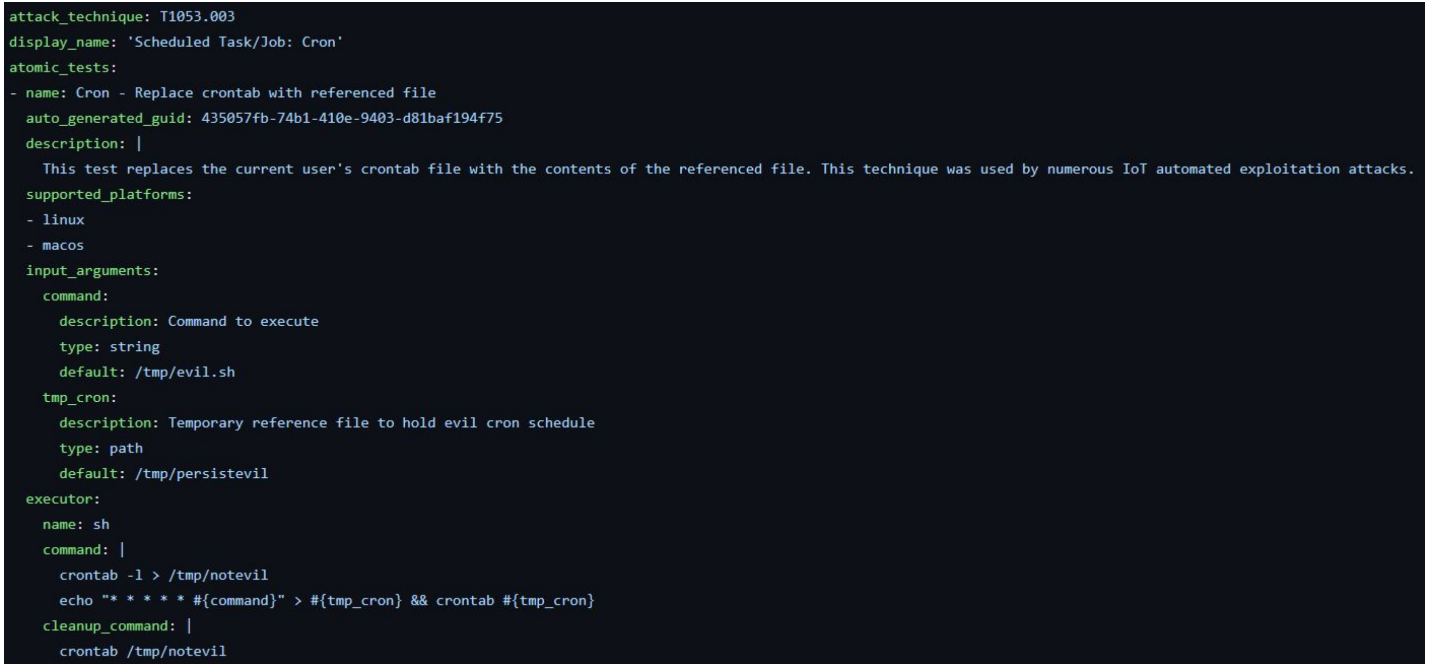
ATOMIC RED
TEAM

- Atomic Red Team has tests for 296 MITRE ATT&CK Techniques for all of the platforms

- The community has created 1592 Atomic Tests for all of the platforms.

Dataset: ATOMIC RED TEAM

ATOMIC RED TEAM



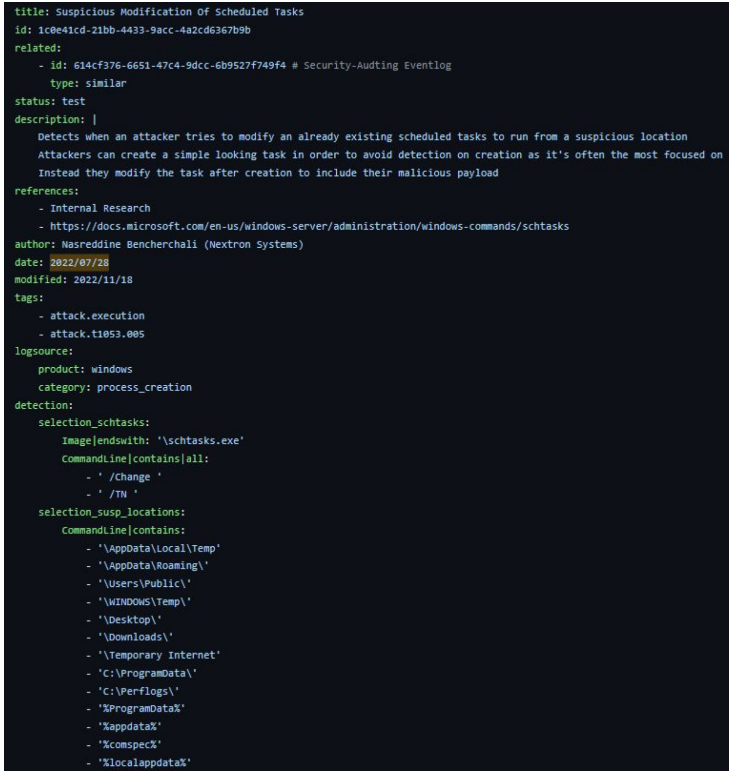
```
attack_technique: T1053.003
display_name: 'Scheduled Task/Job: Cron'
atomic_tests:
- name: Cron - Replace crontab with referenced file
  auto_generated_guid: 435057fb-74b1-410e-9403-d81baf194f75
  description: |
    This test replaces the current user's crontab file with the contents of the referenced file. This technique was used by numerous IoT automated exploitation attacks.
  supported_platforms:
  - linux
  - macos
  input_arguments:
    command:
      description: Command to execute
      type: string
      default: /tmp/evil.sh
    tmp_cron:
      description: Temporary reference file to hold evil cron schedule
      type: path
      default: /tmp/persistevil
  executor:
    name: sh
    command: |
      crontab -l > /tmp/notevil
      echo "* * * * * #{command}" > #{tmp_cron} && crontab #{tmp_cron}
    cleanup_command: |
      crontab /tmp/notevil
```

Dataset: SIGMA

Sigma

SIEM Detection Format

- Sigma is composed by more than 3000 detection rules

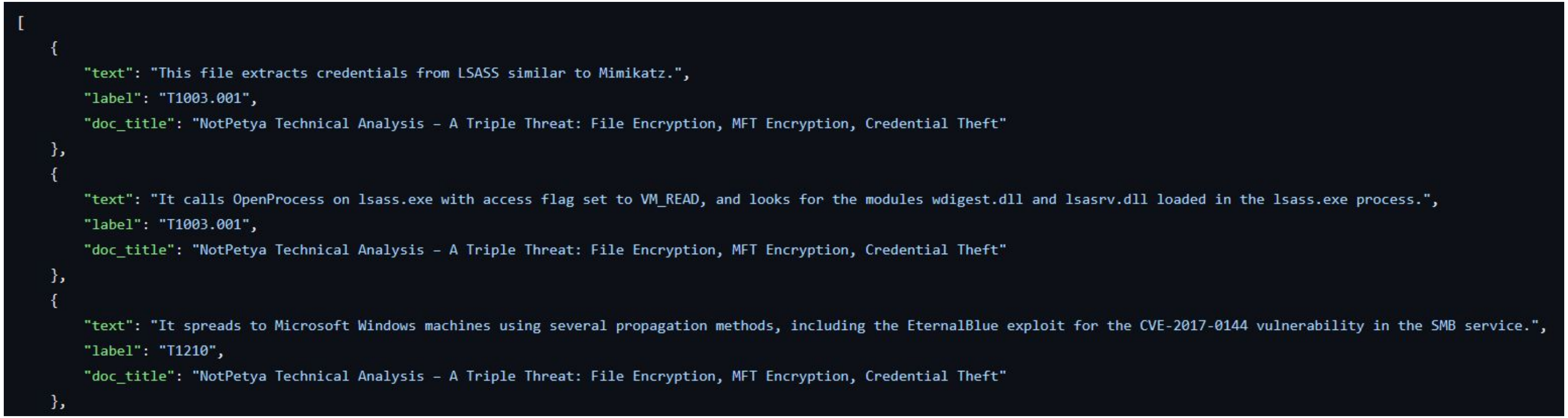


```
title: Suspicious Modification Of Scheduled Tasks
id: 1c0e41cd-21bb-4433-9acc-4a2cd6367b9b
related:
    - id: 614cf376-6651-47c4-9dcc-6b9527f749f4 # Security-Audting Eventlog
      type: similar
status: test
description: |
    Detects when an attacker tries to modify an already existing scheduled tasks to run from a suspicious location
    Attackers can create a simple looking task in order to avoid detection on creation as it's often the most focused on
    Instead they modify the task after creation to include their malicious payload
references:
    - Internal Research
    - https://docs.microsoft.com/en-us/windows-server/administration/windows-commands/schtasks
author: Nasreddine Bencherchali (Nextron Systems)
date: 2022/07/28
modified: 2022/11/18
tags:
    - attack.execution
    - attack.t1053.005
logsource:
    product: windows
    category: process_creation
detection:
    selection_schtasks:
        Image|endswith: '\schtasks.exe'
        CommandLine|contains|all:
            - ' /Change '
            - ' /TN '
    selection_susp_locations:
        CommandLine|contains:
            - '\AppData\Local\Temp'
            - '\AppData\Roaming\'
            - '\Users\Public\'
            - '\WINDOWS\Temp\'
            - '\Desktop\'
            - '\Downloads\'
            - '\Temporary Internet'
            - 'C:\ProgramData\'
            - 'C:\Perflogs\'
            - '%ProgramData%'
            - '%appdata%'
            - '%comspec%'
            - '%localappdata%'
```

Dataset: CYBER THREAT REPORTS

MITRE ENGENUITY
A Foundation for Public Good

- 50 common ATT&CK techniques
- 5089 statements from Cyber Threat Reports

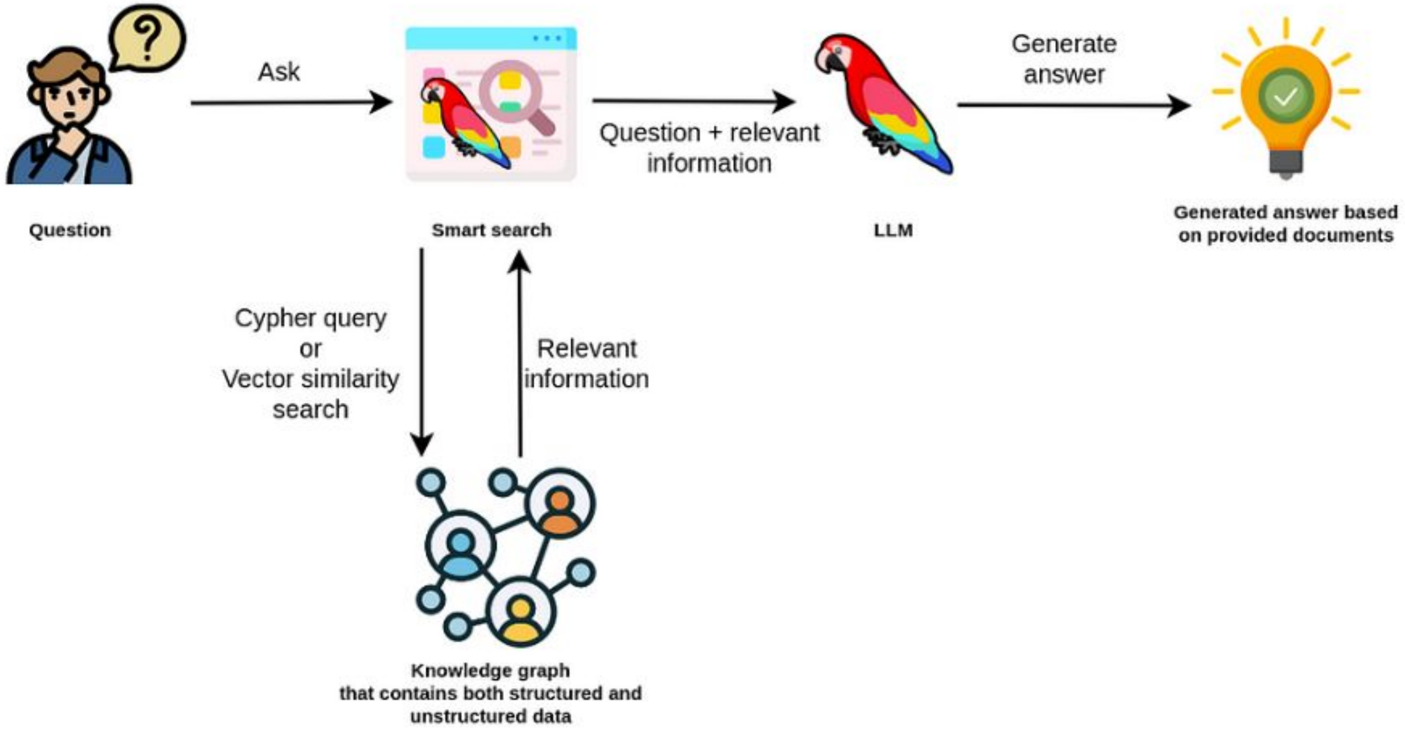


```
[
    {
        "text": "This file extracts credentials from LSASS similar to Mimikatz.",
        "label": "T1003.001",
        "doc_title": "NotPetya Technical Analysis – A Triple Threat: File Encryption, MFT Encryption, Credential Theft"
    },
    {
        "text": "It calls OpenProcess on lsass.exe with access flag set to VM_READ, and looks for the modules wdigest.dll and lsasrv.dll loaded in the lsass.exe process.",
        "label": "T1003.001",
        "doc_title": "NotPetya Technical Analysis – A Triple Threat: File Encryption, MFT Encryption, Credential Theft"
    },
    {
        "text": "It spreads to Microsoft Windows machines using several propagation methods, including the EternalBlue exploit for the CVE-2017-0144 vulnerability in the SMB service.",
        "label": "T1210",
        "doc_title": "NotPetya Technical Analysis – A Triple Threat: File Encryption, MFT Encryption, Credential Theft"
    },
```

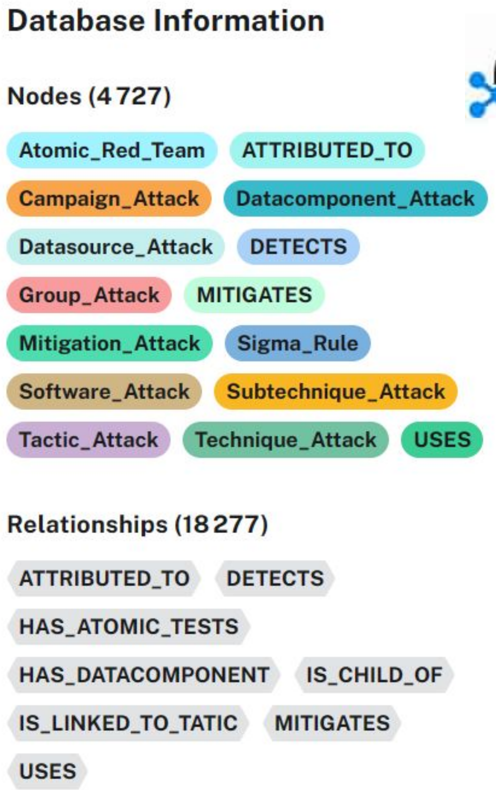
04

What are the implementation and the results ?

The Framework


Ask
Question + relevant information
Generate answer
Question
Smart search
LLM
Generated answer based on provided documents
Cypher query or Vector similarity search
Relevant information
Knowledge graph that contains both structured and unstructured data

Creation of a Cyber Threat Knowledge Graph



Database Information

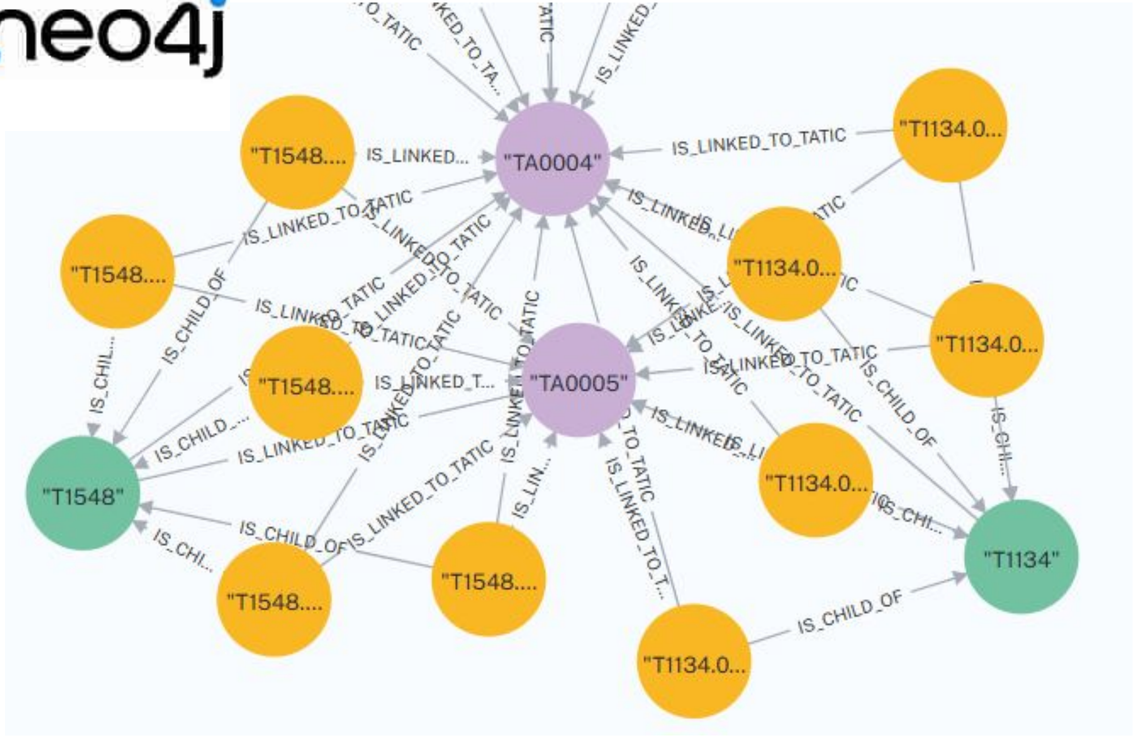
Nodes (4 727)

Atomic_Red_Team ATTRIBUTED_TO Campaign_Attack Datacomponent_Attack Datasource_Attack DETECTS Group_Attack MITIGATES Mitigation_Attack Sigma_Rule Software_Attack Subtechnique_Attack Tactic_Attack Technique_Attack USES

Relationships (18 277)

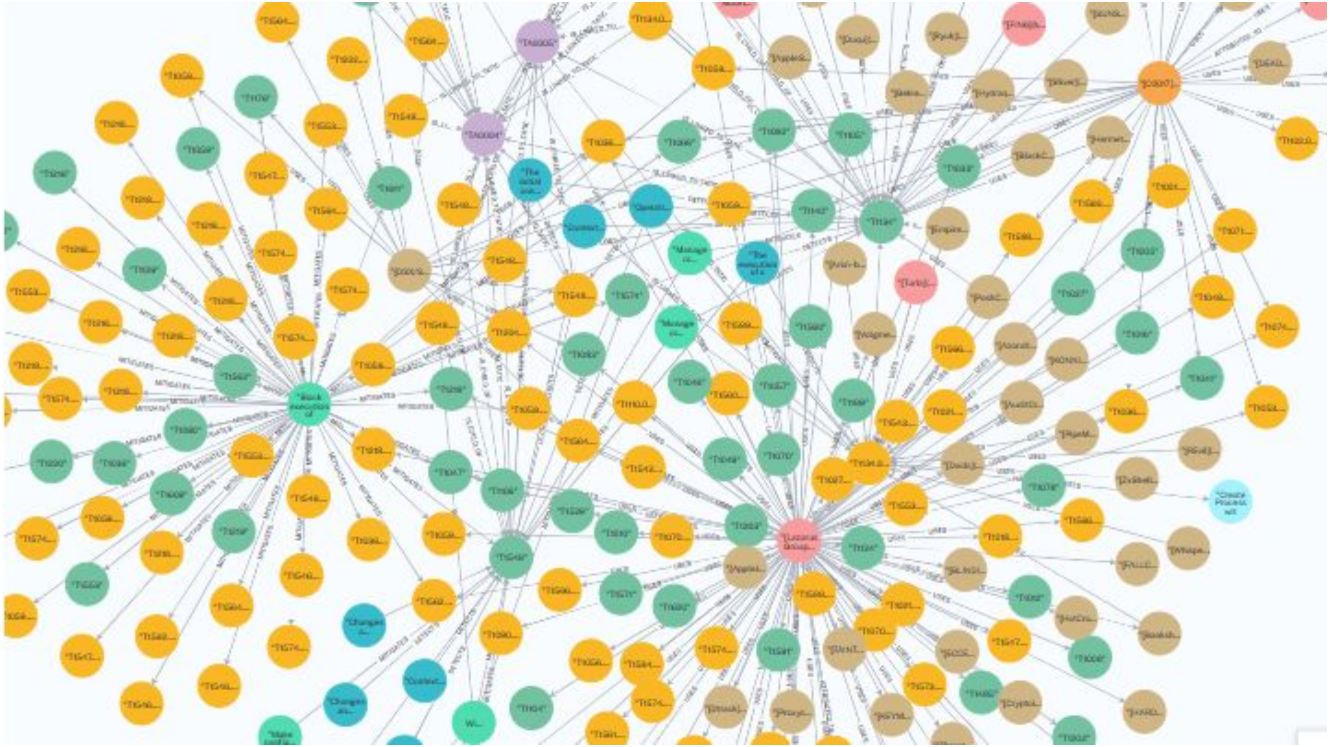
ATTRIBUTED_TO DETECTS HAS_ATOMIC_TESTS HAS_DATACOMPONENT IS_CHILD_OF IS_LINKED_TO_TATIC MITIGATES USES

neo4j

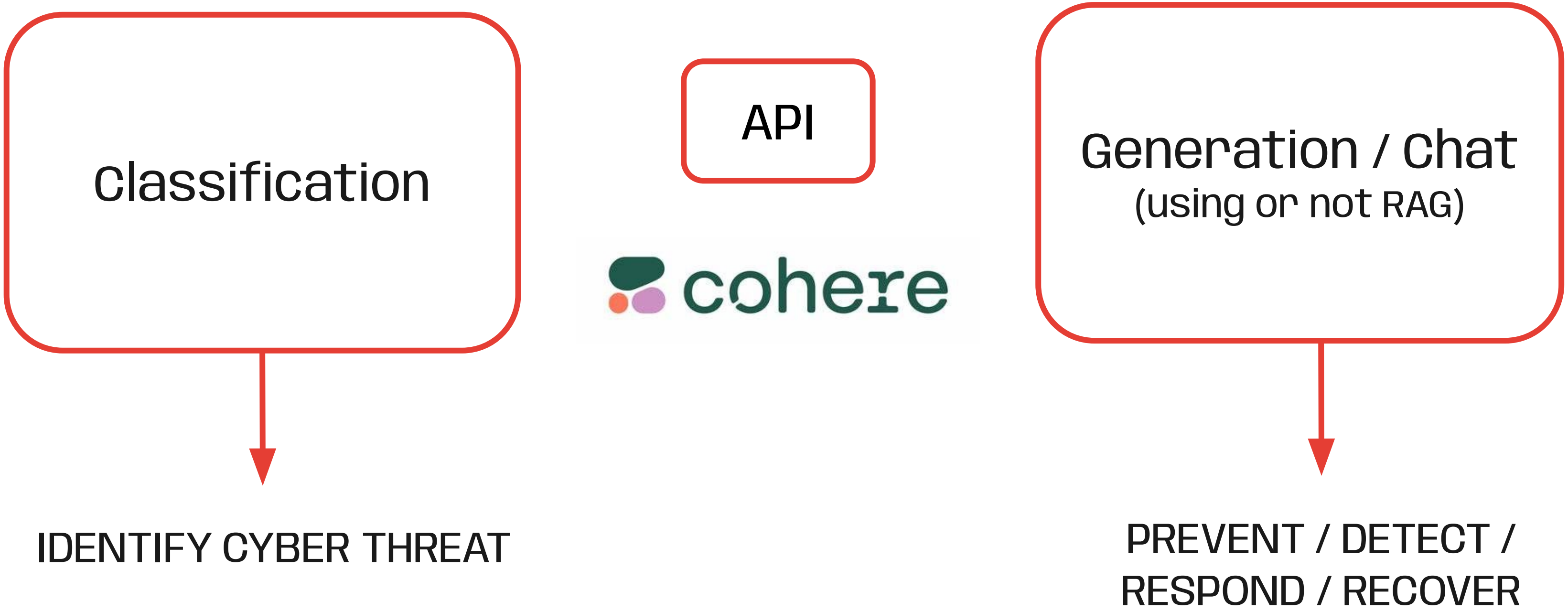


Creation of a Cyber Threat Knowledge Graph

neo4j



Different types of LLM for different usages



Classification

IDENTIFY CYBER THREAT

API

cohere

Generation / Chat
(using or not RAG)

PREVENT / DETECT /
RESPOND / RECOVER

Fined-Tuning improves classification performance

Research Question 1

Example 1:

Input :

```
C:\programdata\procdump64.exe -accepteula
-ma lsass.exe C:\ProgramData\lsass.dmp
```

Output: T1003.001: LSASS Memory

Fined-Tuning improves classification performance

Research Question 1

Example 1:

Input :

```
C:\programdata\procdump64.exe -accepteula
-ma lsass.exe C:\ProgramData\lsass.dmp
```

Output: T1003.001: LSASS Memory

Example 2:

Input :
creating SQL queries to produce the database tables to store the stolen data

Output: T1074.001: Local Data Staging

Fined-Tuning improves classification performance

Research Question 1

Example 1:

Input :
C:\programdata\procdump64.exe -accepteula -ma lsass.exe C:\ProgramData\lsass.dmp

Output: T1003.001: LSASS Memory

Example 2:

Input :
creating SQL queries to produce the database tables to store the stolen data

Output: T1074.001: Local Data Staging

	Classical	Fine-Tuned
Accuracy	0.5625	0.9166
Precision (macro)	0.5051	0.8619
Recall (macro)	0.5599	0.8773
F1 Score (macro)	0.5050	0.8628

Model: *embed-english-v3.0*

Test Set Size: 96 (this is a hard limit fixed by Cohere API)

Train Set Size: 4,838

RAG reliability depends on the method & use case

Research Question 2

Vector Search : Retriever

Example:

Input :
How could I detect T1055: Process Injection?

Output: <List of ranked documents>

RAG reliability depends on the method & use case

Research Question 2

Vector Search : Retriever

Example:

Input :
How could I detect T1055: Process Injection?

Output: <List of ranked documents>

Notes :

The documents are the “Technique_Attack” nodes in the Neo4j graph.

Only the first document of the list is selected.

RAG reliability depends on the method & use case

Research Question 2

Vector Search : Retriever

Example:

Input :
How could I detect T1055: Process Injection?

Output: <List of ranked documents>

	Classic	Rerank	Rerank Fine-tuned
Accuracy	0.7308	0,7692	0.8461

Test Set Size : 26 (the most commun Mitre Techniques)

Notes :

The documents are the "Technique_Attack" nodes in the Neo4j graph.

Only the first document of the list is selected.

RAG reliability depends on the method & use case

Research Question 2

Vector Search : Retriever

Example:

Input :
How could I detect T1055: Process Injection?

Output: <List of ranked documents>

Notes :

The documents are the "Technique_Attack" nodes in the Neo4j graph.

Only the first document of the list is selected.

	Classic	Rerank	Rerank Fine-tuned
Accuracy	0.7308	0,7692	0.8461

Test Set Size : 26 (the most commun Mitre Techniques)

Rerank Fine-Tuned:

First, for each 201 "Technique" nodes, 3 questions have been generated based on their content.

Second, the rerank model has been trained on the dataset created at the first step.

RAG reliability depends on the method & use case

Research Question 2

Graph Search : GraphCypherQAChain

Example:

Input :
How could I detect T1055: Process Injection?

Intermediate Output:

```
MATCH (m:Technique_Attack {id_attack:\"T1055\"})
RETURN m.id_attack as ID, m.name_attack as
NAME, m.detection as DETECTION
```

Context: <Result of Cypher Command>

Output: <Response>

RAG reliability depends on the method & use case

Research Question 2

Graph Search : GraphCypherQAChain

Example:

Input :
How could I detect T1055: Process Injection?

Intermediate Output:

```
MATCH (m:Technique_Attack {id_attack:\"T1055\"})
RETURN m.id_attack as ID, m.name_attack as
NAME, m.detection as DETECTION
```

Context: <Result of Cypher Command>

Output: <Response>

	2 Few-Shot
Accuracy	0,8461

CYPHER GENERATION PROMPT TEMPLATE:

```
Task: Generate Cypher statement to query a graph database.
Instructions:
Use only the provided relationship types and properties in the schema.
Do not use any other relationship types or properties that are not provided.
Schema:
{schema}
Note: Do not include any explanations or apologies in your responses.
Do not respond to any questions that might ask anything else than for you to construct a Cypher
statement.
Do not include any text except the generated Cypher statement.
Examples: Here are a few examples of generated Cypher statements for particular questions:
# How could I detect T1548: Abuse Elevation Control Mechanism ?
MATCH (m:Technique_Attack {{id_attack:"T1548"}})
RETURN m.id_attack as ID, m.name_attack as NAME, m.detection as DETECTION
# How could I detect T1595: Active Scanning ?
MATCH (m:Technique_Attack {{id_attack:"T1595"}})
RETURN m.id_attack as ID, m.name_attack as NAME, m.detection as DETECTION

The question is:
{question}"""
```

05

What are the Implications & Main Contributions ?

What are the Implications & Main Contributions ?

- Fine-tuned large language model helps for classification task
- Rerank Fine-Tuned helps to retrieve documents in RAG system for Vector Search
- Graph Cypher Chain helps to retrieve efficiently documents in RAG system without fine-tuning, only with few-shot
- The knowledge graph database on Cyber Threat Intelligence that I made, allows to make hundreds of experiments

THANKS FOR YOUR “ATTENTION” !

CREDITS: This presentation template was created by **Slidesgo**, and includes icons by **Flaticon**, and infographics & images by **Freepik**