When Knowledge Graph Meets TTPs: Highly Automated and Adaptive Executable TTP Intelligence for Security Evaluation

Jack Tang, Lorin Wu, Porot Mo

About US

- Jack Tang @360 Digital Security Group

Jack, the team leader, has over 15 years of expertise in the security industry and is presently focusing on the use of MITRE ATT&CK® in security operations and threat intelligence. He is knowledgeable on kernel and virtualization vulnerability research for Android, Mac, and Windows. He ranked Top 16 on the MSRC in 2016 and Top 34 in 2015. In 2016, he was awarded the Microsoft Mitigation Bypass Bounty. Jack has lectured at security conferences such as CanSecWest, Black Hat, HITCon, and PacSec.

- Lorin Wu @360 Digital Security Group

Building an offensive and defensive knowledge graph for cyber security is what Lorin is currently working on. He spent many years working at Trend Micro, where he concentrated on the creation of heuristic patterns and mobile sandbox technologies. During this period, he identified various international cyber security operations that were reported to INTERPOL and Google Security Team.

- Porot Mo @360 Digital Security Group

Porot received a master's degree from the University of Chinese Academy of Sciences after graduating from the University of Science and Technology of China. He is currently devoted to the study of offensive and defensive technologies and has three years of expertise in sandbox development.

Agenda

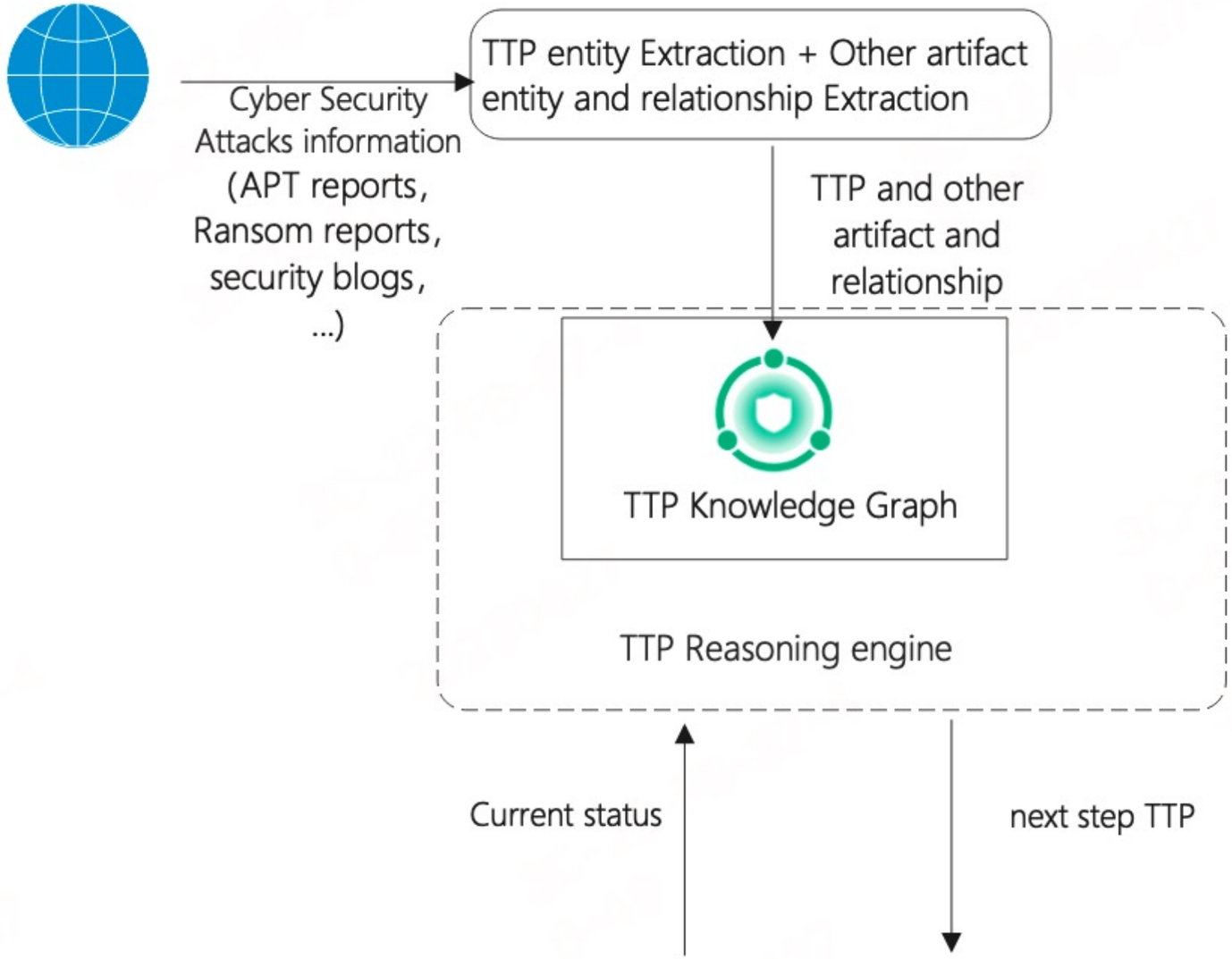
- Background
- Solution Overview
- TTP(Tactics, Techniques, Procedures) Knowledge Graph Construction
  - TTPs Extraction Automatically
- Adaptive attack path reasoning for BAS (Breach and Attack Simulation)

Background

BAS (Breach and Attack Simulation) increasingly needs:

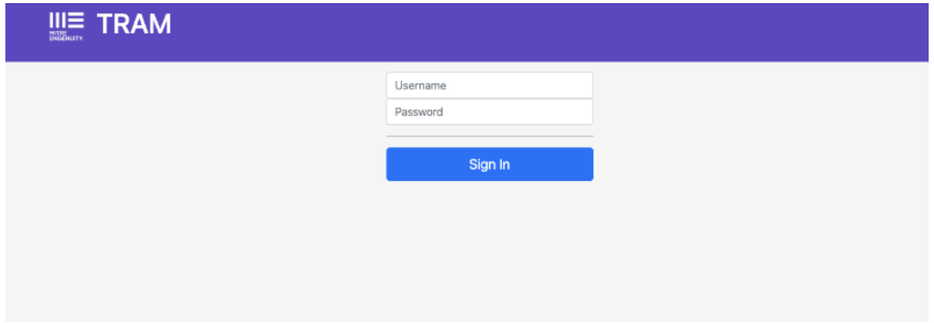
- Keeping up with the TTPs of attackers.
- Selecting the appropriate TTP simulation according to the actual situation of the target organization.
- Using the attack path (sequential TTP) to assess the entire defense-in-depth of the target organization.

Solution Overview


TTP entity Extraction + Other artifact entity and relationship Extraction
Cyber Security Attacks information (APT reports, Ransom reports, security blogs, ...)
TTP and other artifact and relationship
TTP Knowledge Graph
TTP Reasoning engine
Current status
next step TTP

TTP Knowledge Graph Construction - TTPs Extraction Automatically

The Prior Art



Token feature engineering, ML based, MITER Tram project
https://github.com/center-for-threat-informed-defense/tram/

Take NER extraction as the task
Token feature engineering, ML based, MITER Tram project

OpenAI

GPT3.5, GPT 4.0 based
https://chat.openai.com/

Comparisons in actually encountered TTPs data

- TTPExtractor vs. ChatGPT

	OpenAI ChatGPT	Luwak TTPExtractor
Precision	0.2015	0.7241
Recall	0.4927	0.5769
F1 Score	0.2861	0.6422

1. The ChatGPT version is 3.5, the test date is 5/5/2023;
2. The test set consists of 5 Chinese reports and 5 English reports, and the final score is obtained after reviewing the predicted results by security experts;
3. Please refer to another attachment for details [Comparative data of TTPExtractor and ChatGPT], includes predicted results for each report, expert review results for each report, and prompts for using GPT.

Key problems to improve the accuracy of TTP extraction

- Distinguish primary and secondary tactics and techniques, and extract them based on actual attack scenarios

November 12, 2022

Disabled antivirus (AV) programs such as Windows Defender

add "HKLM\Software\Policies\Microsoft\Windows Defender" /v DisableAntiVirus /t REG_DWORD /d 1 /f

add "HKLM\Software\Policies\Microsoft\Windows Defender" /v DisableAntiSpyware /t REG_DWORD /d 1 /f

add "HKLM\Software\Policies\Microsoft\Windows Defender\MpEngine" /v MpEnablePus /t REG_DWORD /d 0 /f

Primary	Tactic	Technique
True	Defense Evasion (TA0005)	Disable or Modify Tools (T1562.001)
False	Defense Evasion (TA0005)	Modify Registry (T1112)

Key problems to improve the accuracy of TTP extraction

- Extract tactics and techniques involved in attacks from multiple perspectives such as command lines, tools, and code snippets

The fetched payload is supposed to be saved in %Profile%\update.dll. Eventually, the fetched file is spawned with the following commands:

- Command #1: rundll32.exe %Profile%\update.dll,#1 5pOygllrsNaAYqx8JNZSTouZNjo+j5XEFHzxqllqpQ==
- Command #2: rundll32.exe %Profile%\update.dll,#1 5oGygYVhos+laqBlNdFaVJSfMiwhh4LCDn4=

Primary	Tactic	Technique
True	Defense Evasion (TA0005)	Rundll32 (T1218.011)

Key problems to improve the accuracy of TTP extraction

- Extract based on the context of the attack description in the report

The first stage macro checks for the presence of a Kaspersky security solution on the victim's machine by trying the following file paths:

- C:\Windows\avp.exe # Kaspersky AV
- C:\Windows\Kavsvc.exe # Kaspersky AV
- C:\Windows\clisve.exe # Unknown

If a Kaspersky security solution is indeed installed on the system, it enables trust access for Visual Basic Application (VBA) by setting the following registry key to '1':

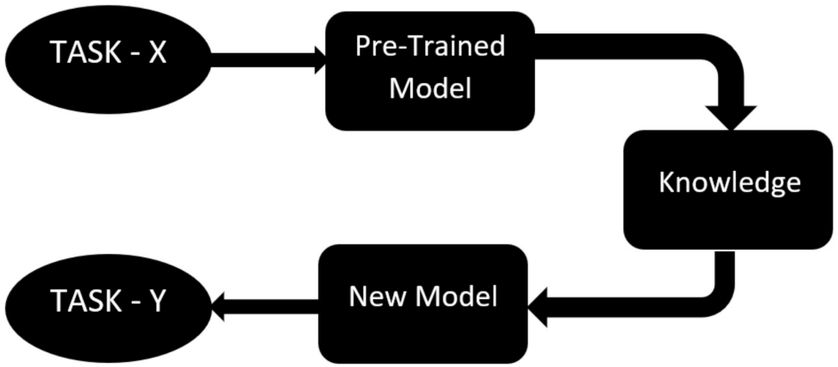
```
HKEY_CURRENT_USER\Software\Microsoft\Office\[Application.Version]\Word\Security\AccessVBOM
```

By doing so, Microsoft Office will trust all macros and run any code without showing a security warning or requiring the user's permission. Next, the macro creates a mutex named 'sensiblemtv16n' and opens the malicious file once more. Thanks to the "trust all macros" setting, the macro will be executed automatically.

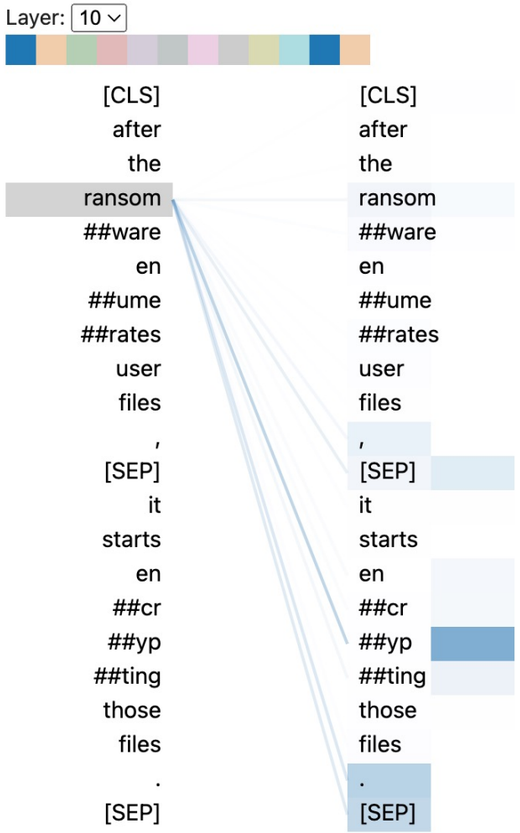
Primary	Tactic	Technique
True	Defense Evasion (TA0005)	Modify Registry (T1112)
False	Discovery (TA0007)	Security Software Discovery (T1518.001)

Extract TTPs using pretrained language models and transfer learning

- Pretrained language models
  - BERT
  - Whole Word Masking(WWM) technology
- Transfer learning
  - Finetune



https://www.javatpoint.com/transfer-learning-in-machine-learning



After the ransomware enumerates user files, the ransomware starts encrypting those files.

bert-base-uncased, English

通过修改注册表启动项，实现持久化。

bert-wwm-ext, Chinese

Extract TTPs using pretrained language models and transfer learning

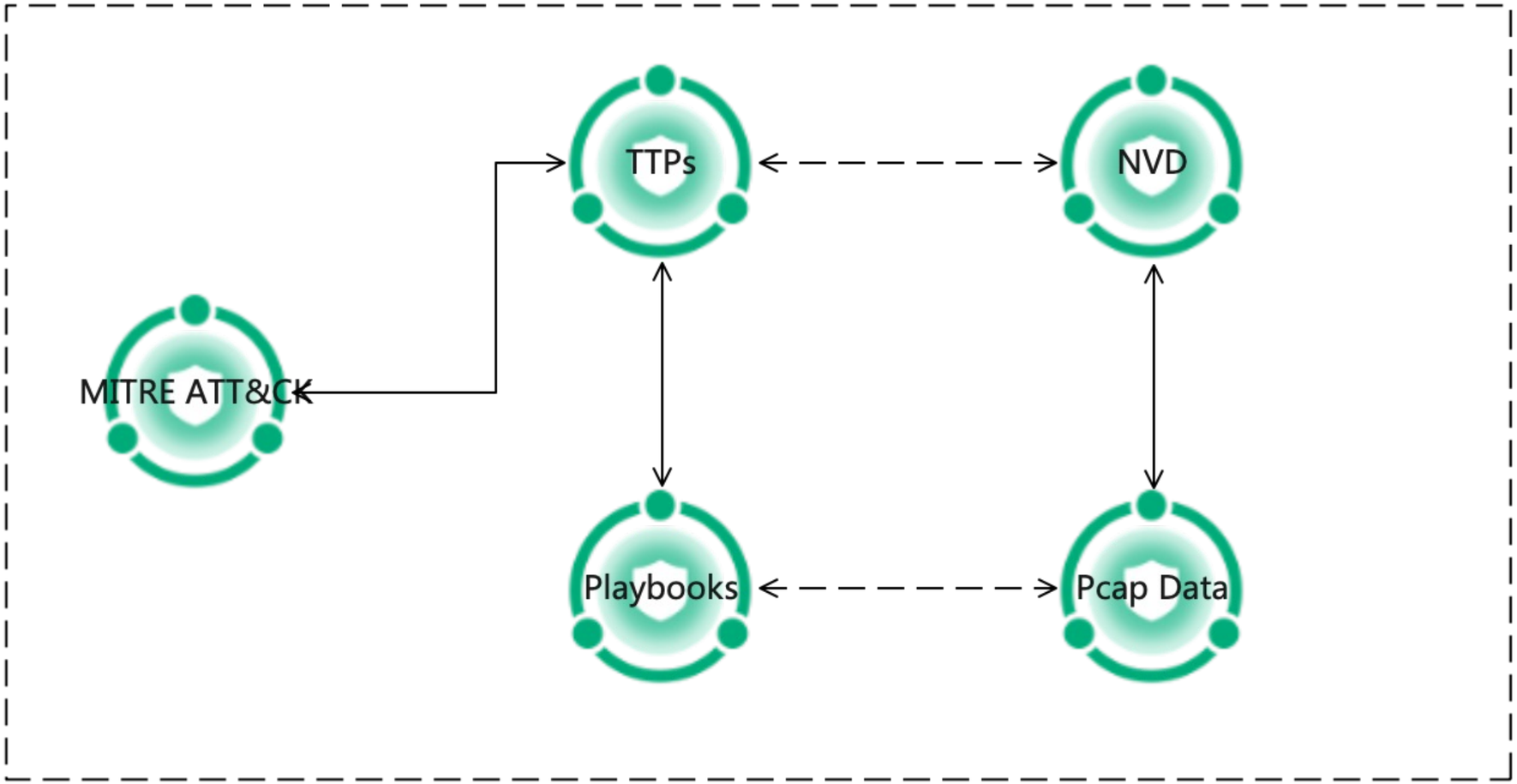
- The pipeline of extract TTPs from unstructured text

Shellcode first checks whether there is a Kaspersky main process avp.exe or Avast main process AvastSvc.exe in the current system, and if it exists, execute the shell command "/c schtasks /create /sc minute /mo 1 /tn WindowsUpdate /tr C: \\ProgramData\\OneDrive.exe".

Primary	Tactic	Technique
True	Persistence (TA0003)	Scheduled Task (T1053.005)
False	Execution (TA0002)	Windows Command Shell (T1059.003)
False	Defense Evasion (TA0005)	Masquerade Task or Service (T1036.004)
False	Discovery (TA0007)	Security Software Discovery (T1518.001)

TTP Knowledge Graph Construction - Semantic Web Building

TTP Knowledge Graph Construction



ATT&CK semantic web building

- Appropriate offensive entities and relationships
  - Technique – [platformRequires] –> Platform
  - Technique – [serviceRequires] –> ServiceState

ID: T1021.001

Sub-technique of: T1021

ⓘ Tactic: Lateral Movement

ⓘ Platforms: Windows

ⓘ System Requirements: RDP service enabled, account in the Remote Desktop Users group

ⓘ Permissions Required: Remote Desktop Users, User

Contributors: Matthew Demaske, Adaptforward

Version: 1.1

Created: 11 February 2020

Last Modified: 30 March 2023

T1021.001

T1021.001
URI: https://attack.mitre.org/enterprise-attack#T1021.001
Data property assertions:
T1021.001 description "Adversaries may use [Valid Accounts](https://attack.mitre.org/techniques/T1078) to log into a computer using the Remote Desktop Protocol (RDP). The adversary may then perform act...
T1021.001 serviceRequires "rdp:running"
T1021.001 id "attack-pattern--eb062747-2193-45de-8fa2-e62549c37ddf"
T1021.001 name "Remote Desktop Protocol"
T1021.001 tid "T1021.001"
T1021.001 type "attack-pattern"
Annotations:
platformRequires Windows
achieves TA0008
permissionObtains User
permissionObtains Remote-Desktop-Users
permissionRequires User
permissionRequires Remote-Desktop-Users

ATT&CK semantic web building

- Appropriate offensive entities and relationships
  - Technique – [achieves] –> Tactic
  - Technique – [permissionRequires] –> Permission
  - Technique – [permissionObtains] –> Permission

ID: T1548.002
Sub-technique of: T1548
ⓘ Tactics: Privilege Escalation, Defense Evasion
ⓘ Platforms: Windows
ⓘ Permissions Required: Administrator, User
ⓘ Effective Permissions: Administrator
ⓘ Defense Bypassed: Windows User Account Control
Contributors: Casey Smith; Stefan Kanthak
Version: 2.0
Created: 30 January 2020
Last Modified: 19 April 2022

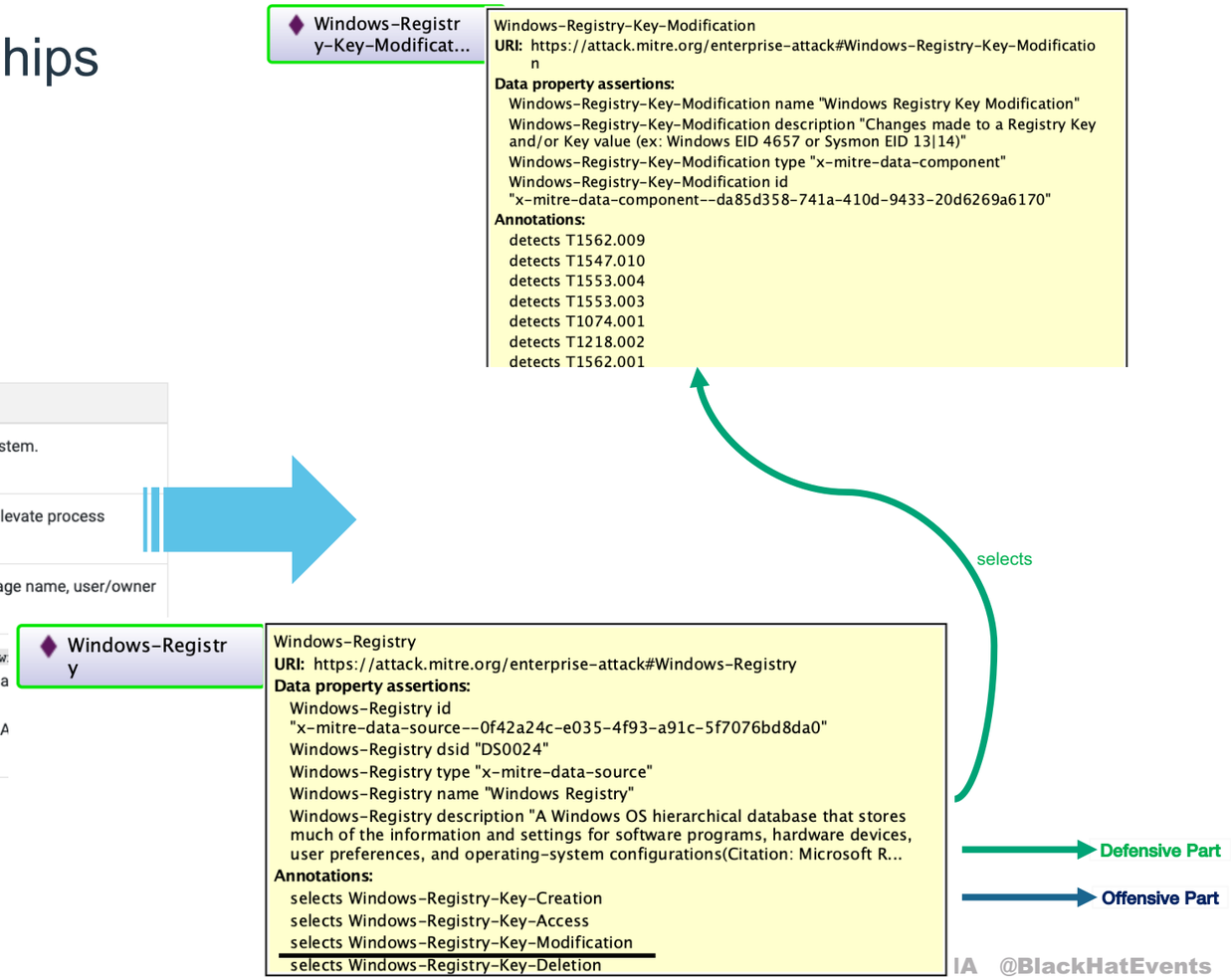
T1548.002

T1548.002
URI: https://attack.mitre.org/enterprise-attack#T1548.002
Data property assertions:
T1548.002 id "attack-pattern--120d5519-3098-4e1c-9191-2aa61232f073"
T1548.002 name "Bypass User Account Control"
T1548.002 type "attack-pattern"
T1548.002 description "Adversaries may bypass UAC mechanisms to elevate process privileges on system. Windows User Account Control (UAC) allows a program to elevate its privileges (tracked as integrity levels ranging from...
T1548.002 tid "T1548.002"
Annotations:
achieves TA0004
permissionObtains Administrator
achieves TA0005
permissionRequires Administrator
permissionRequires User
platformRuns Windows

ATT&CK semantic web building

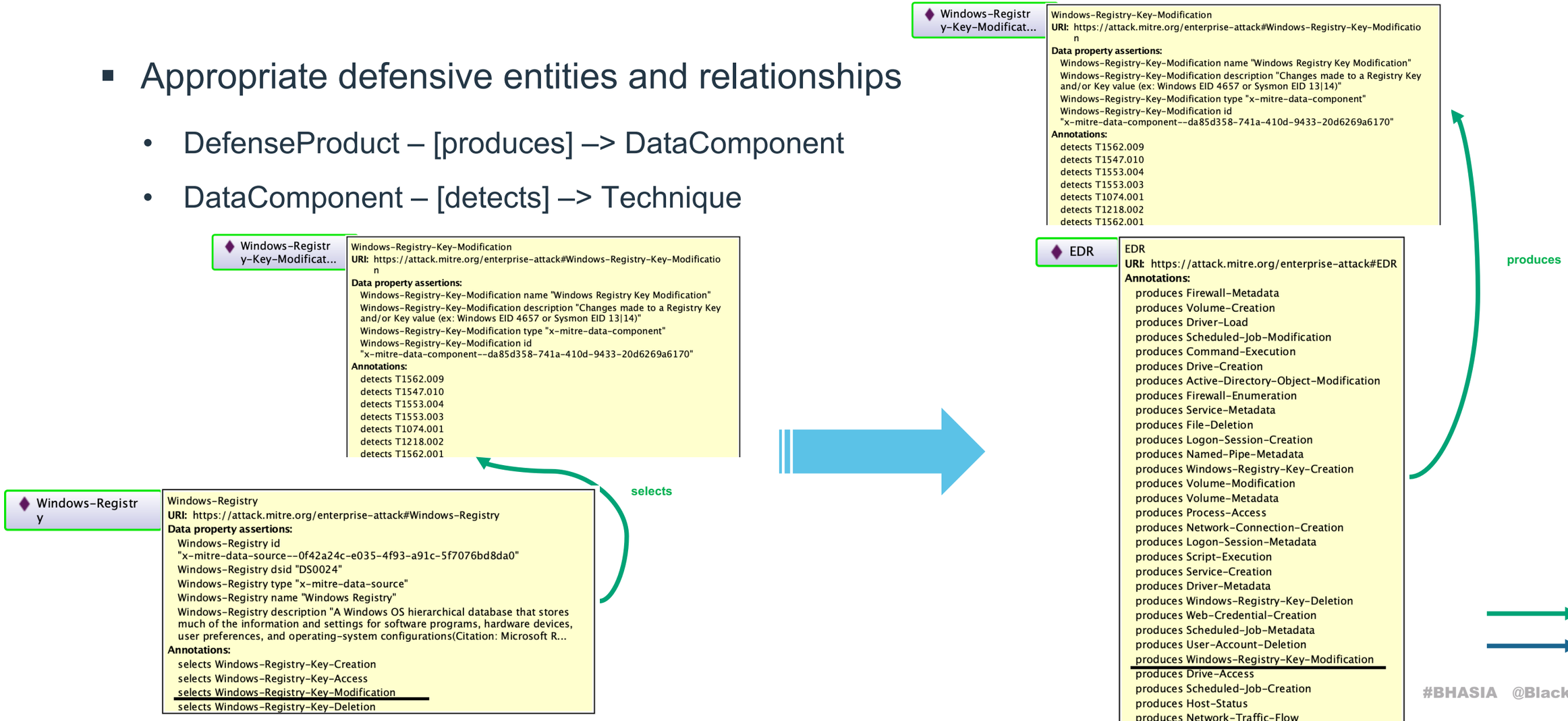
- Appropriate defensive entities and relationships
  - DataSource – [selects] –> DataComponent
  - DataComponent – [detects] –> Technique

ID	Data Source	Data Component	Detects
DS0017	Command	Command Execution	Monitor executed commands and arguments that may bypass UAC mechanisms to elevate process privileges on system.
DS0009	Process	Process Creation	Monitor newly executed processes, such as `eventvwr.exe` and `sdclt.exe`, that may bypass UAC mechanisms to elevate process privileges on system.
		Process Metadata	Monitor contextual data about a running process, which may include information such as environment variables, image name, user/owner that may bypass UAC mechanisms to elevate process privileges on system.
DS0024	Windows Registry	Windows Registry Key Modification	Some UAC bypass methods rely on modifying specific, user-accessible Registry settings. For example:* The `eventvw` the `[HKEY_CURRENT_USER]\Software\Classes\mscfile\shell\open\command` Registry key.[6]* The `sdclt.exe` bypa `[HKEY_CURRENT_USER]\Software\Microsoft\Windows\CurrentVersion\App Paths\control.exe` and `[HKEY_CURRENT_USER]\Software\Classes\exefile\shell\runas\command\isolatedCommand` Registry keys.[65][66]A monitor these Registry settings for unauthorized changes.

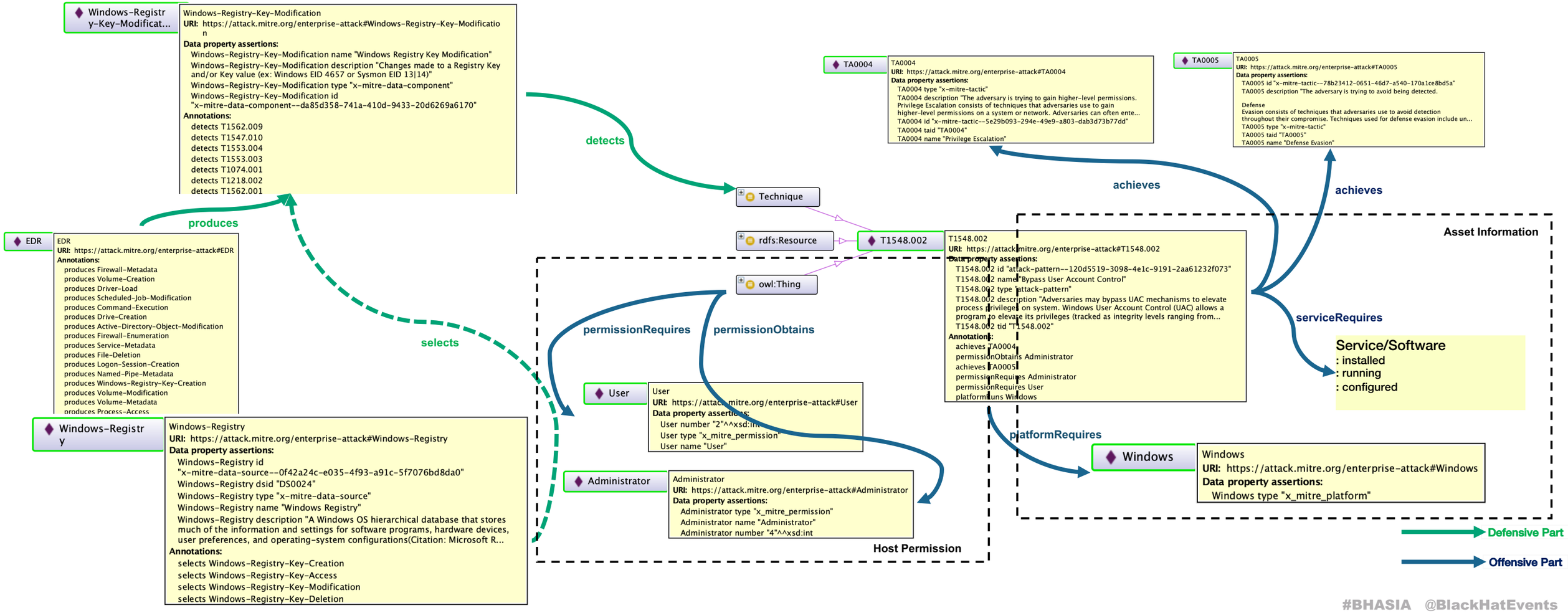


ATT&CK semantic web building

- Appropriate defensive entities and relationships
  - DefenseProduct – [produces] –> DataComponent
  - DataComponent – [detects] –> Technique

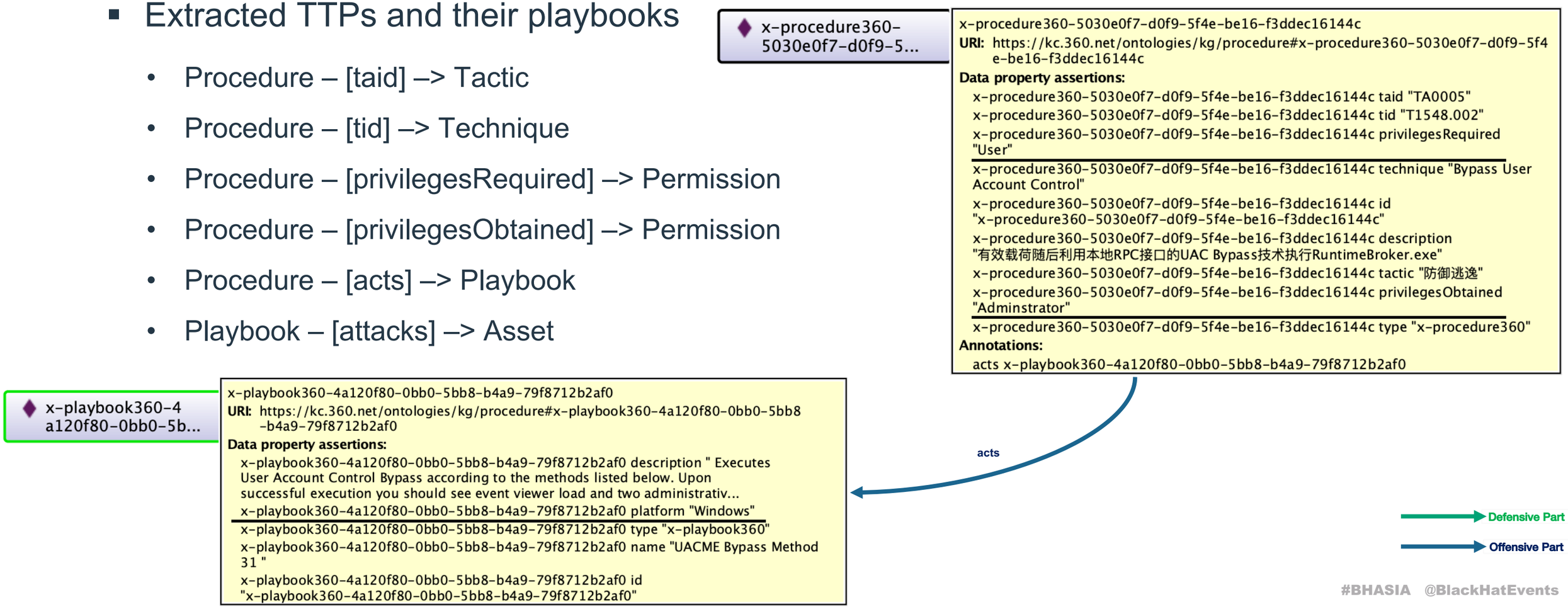


ATT&CK semantic web building



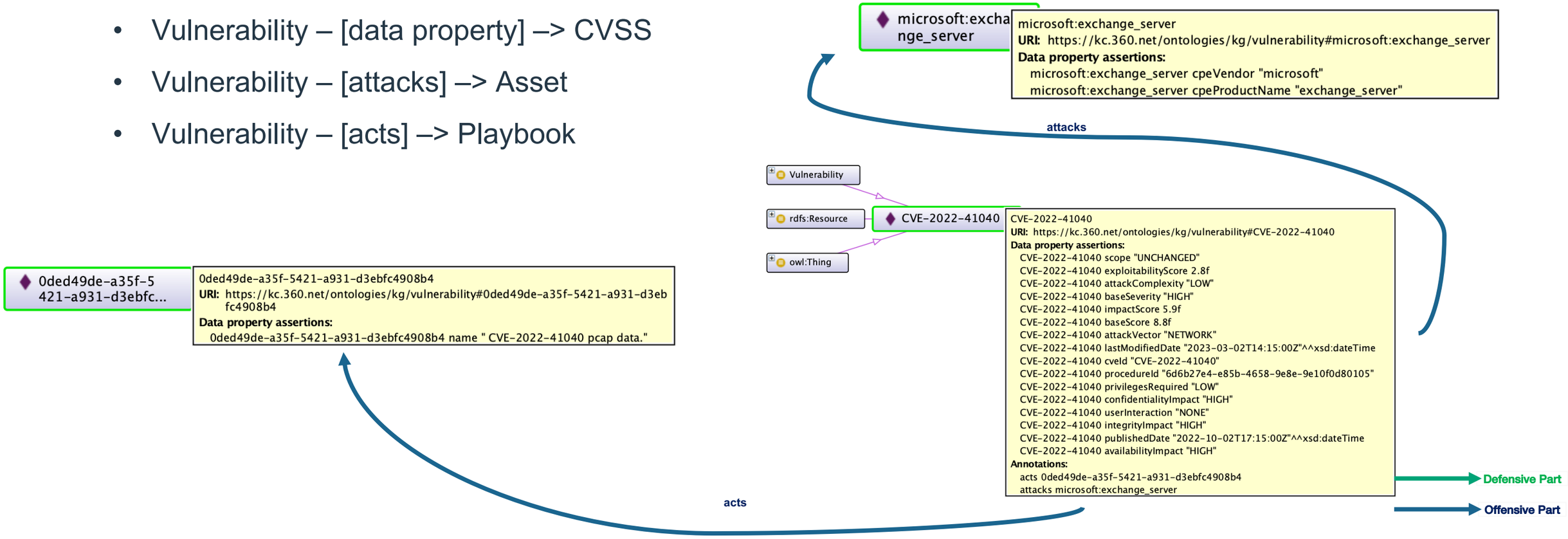
TTPs and playbooks semantic web building

- Extracted TTPs and their playbooks
  - Procedure – [taid] –> Tactic
  - Procedure – [tid] –> Technique
  - Procedure – [privilegesRequired] –> Permission
  - Procedure – [privilegesObtained] –> Permission
  - Procedure – [acts] –> Playbook
  - Playbook – [attacks] –> Asset



TTPs and playbooks semantic web building

- NVD vulnerabilities and their PCAP data
  - Vulnerability – [data property] –> CVSS
  - Vulnerability – [attacks] –> Asset
  - Vulnerability – [acts] –> Playbook

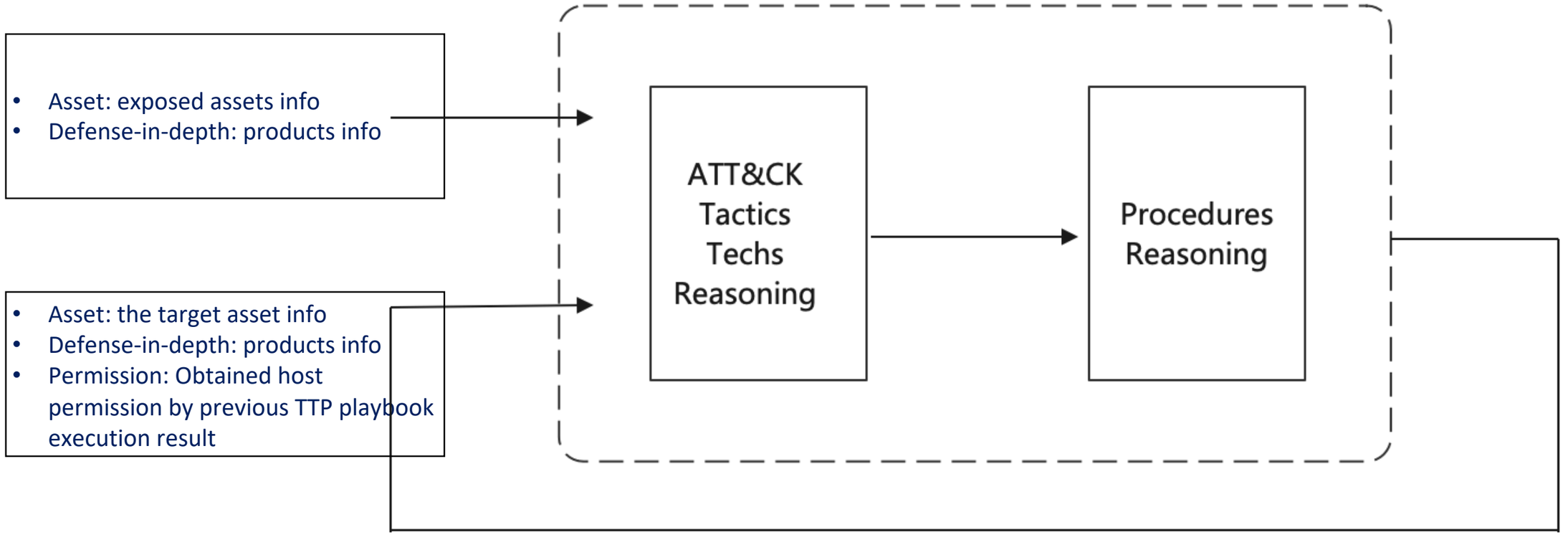


Adaptive attack path reasoning for BAS

Input data for assessment

- Asset information for the target organization
  - Results of asset mapping tools
- Network typology configuration for the target organization
  - Ensure the authenticity of the network topology where the assets are located as much as possible, e.g. determine the location of the assets, DMZ, Office and network connectivity
- Defense-in-depth typology configuration for the target organization
  - Keep asset-based security topologies as real as possible, e.g. determine which assets are protected by which security products

TTP Reasoning Engine

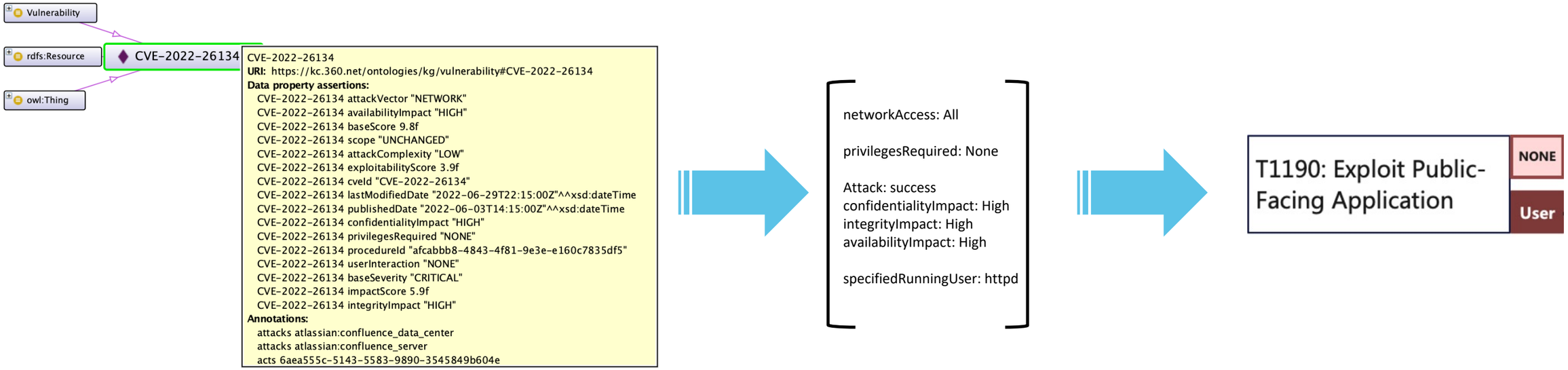


TTP Reason Engine: Tactics, techniques reasoning

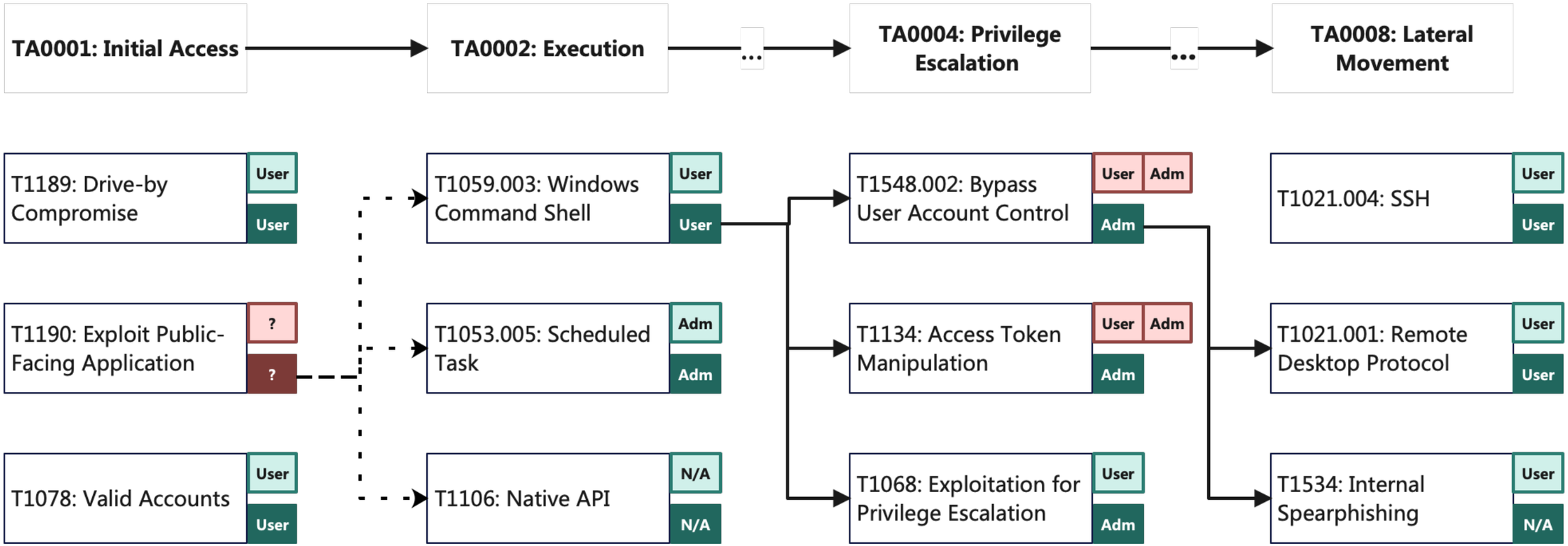
- Reasoning based on MITRE ATT&CK
  - The first dimension: The MITRE ATT&CK kill-chain phase determines the tactic route
    - ✓ Start from Initial Access (TA0001)
    - ✓ Put Credential Access (TA0006) and Lateral Movement (TA0008) last
  - The second dimension: Using the results of the previous step simulation attack, reason the techniques that can be used in the next phase
    - ✓ Host permission: obtained from the previous step's simulated attack, meet the techniques
    - ✓ Asset: The asset condition and platform that meet the techniques
    - ✓ Defense-in-depth: The techniques that enable defense products to produce detection data

TTP Reason Engine: Tactics, techniques reasoning

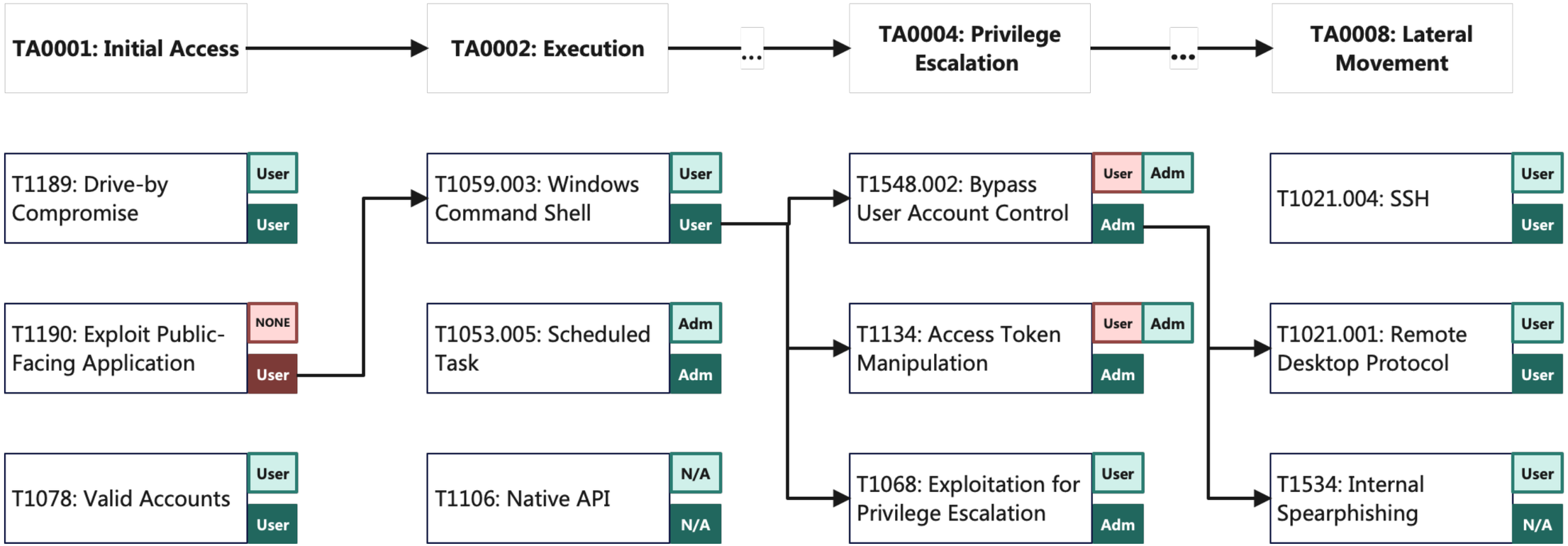
- Reasoning based on MITRE ATT&CK
  - Reasoning based on permission levels
    - ✓ E.g. for Windows, system permission can act as one of [system, Administrator, User and None]



TTP Reason Engine: Tactics, techniques reasoning



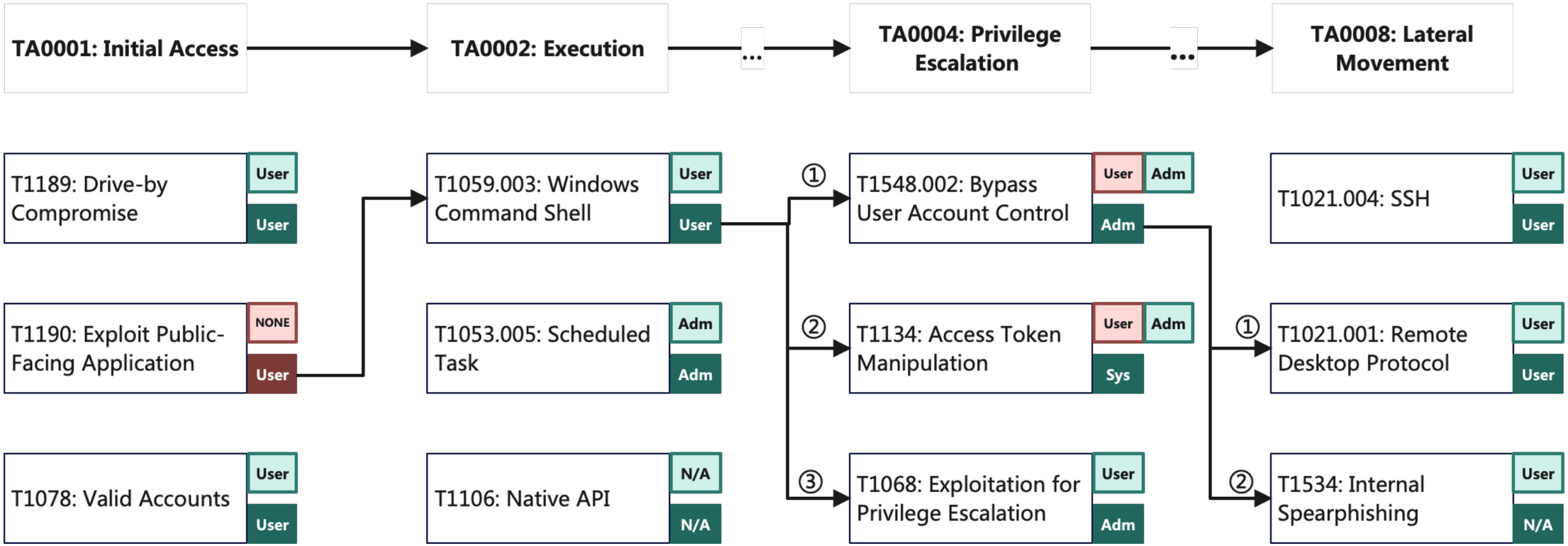
TTP Reason Engine: Tactics, techniques reasoning



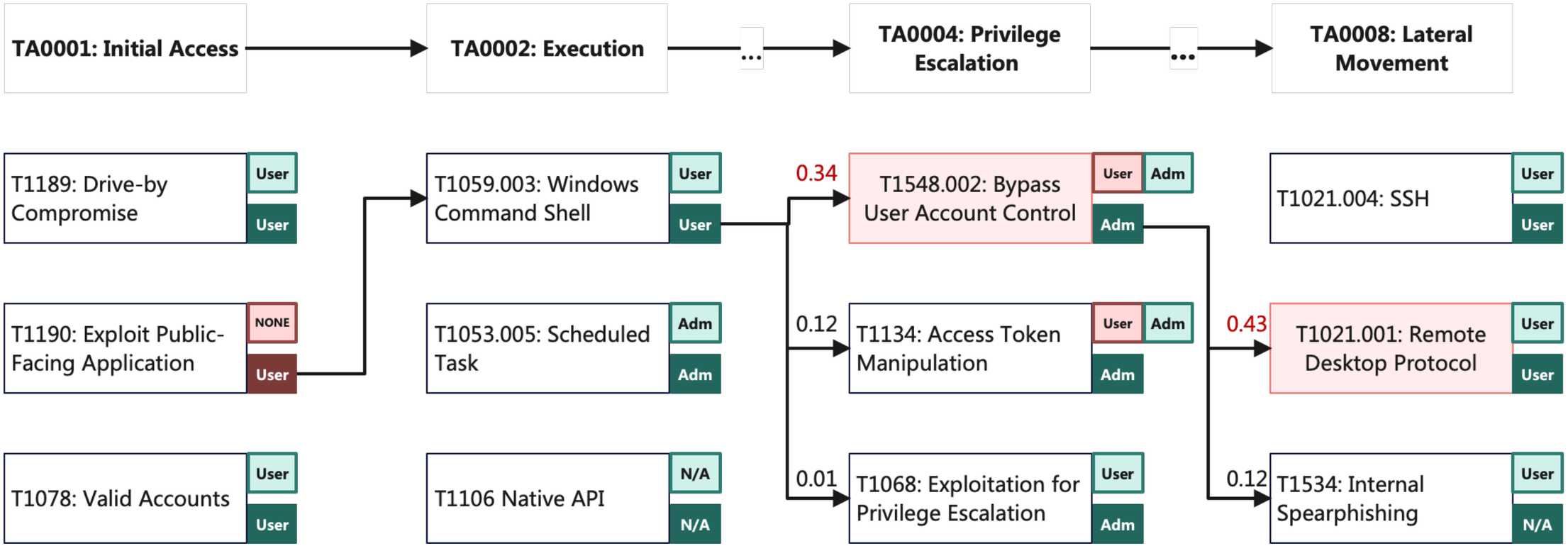
TTP Reason Engine: Procedures reasoning

- Based on real world procedures distribution
  - Continuously collect procedures by TTP Extraction approach
- Determine possibility which TTP to use in the next step
  - In the TTP chains we collected in real cybersecurity attacks
    - ✓ in current state: permission owned or obtained, asset
    - ✓ The most possible procedures used in attacks: the quantity, the popularity

TTP Reason Engine: Procedures reasoning



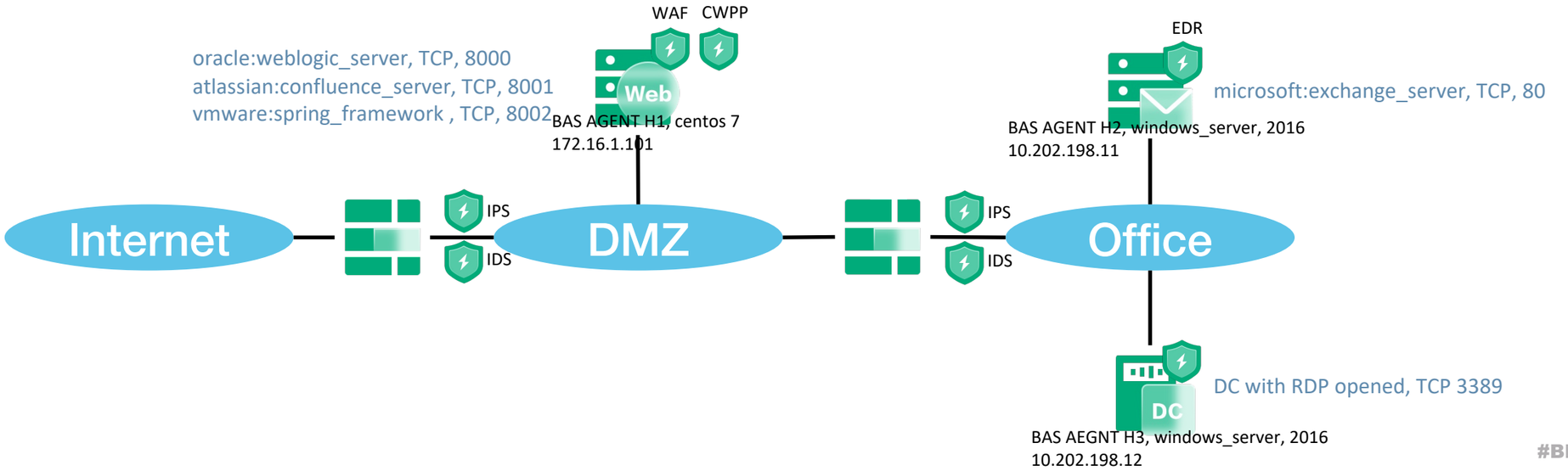
TTP Reason Engine: Procedures reasoning



Adaptive attack path reasoning for BAS

- A Small Real-World Example - Key attack path
  - Target services are weblogic server, confluence server, spring framework, exchange server, RDP
  - Target operation systems are centos and windows server

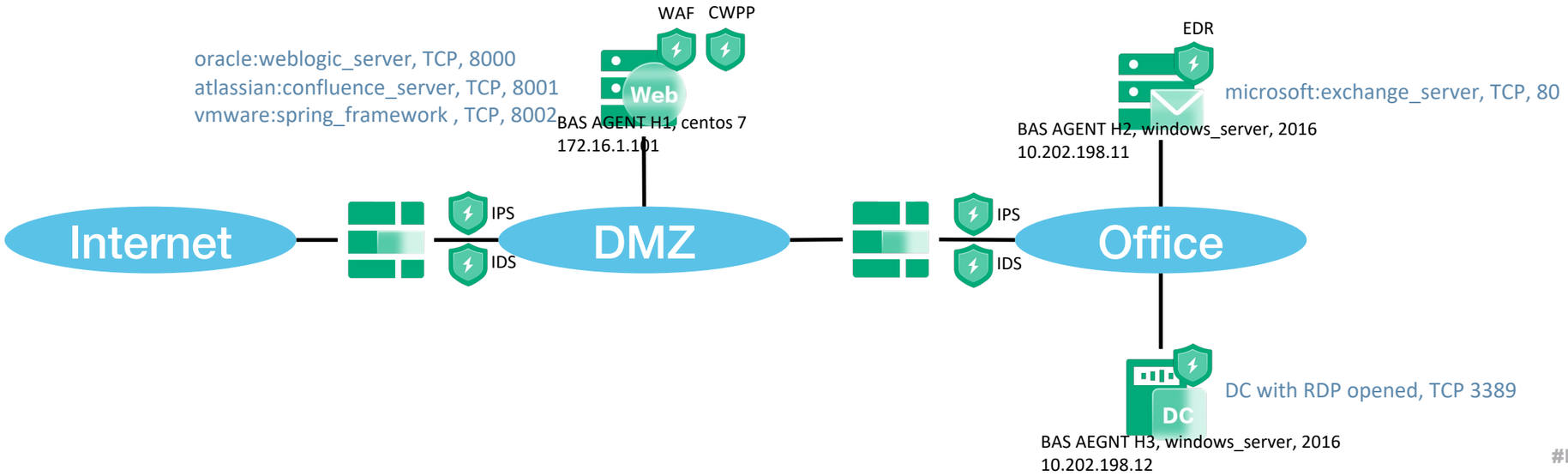
Therefore, the TTP Reason Engine will only reason the attack path around above assets.



Adaptive attack path reasoning for BAS

- A Small Real-World Example - Key attack path
  - Defense-in-depth topology consists of IPS, IDS, WAF, CWPP, and EDR

Therefore, the TTP Reason Engine will only reason the attack path that these security products will generate detection data.



Adaptive attack path reasoning for BAS

- A Small Real-World Example - Key attack path
  - BAS AGENT H1, centos 7

    1. According to exposed *[oracle:weblogic_server]*, *[atlassian:confluence_server]* and *[vmware:spring_server]*, selects corresponding vulnerability pcap playbooks.

Step	Result	Security Products	Area	Agent IP	Asset	Privilege Obtained	Tactic	Technique	Possibility	Privilege Required	Playbook
18	PREVENTED	IPS, CWPP, IDS, WAF	DMZ	172.16.1.101	oracle:weblogic_server	None	TA0001 Initial Access	T1190 Exploit Public-Facing Application	0.15650742	None	x-playbook360-db193392-e328-510a-9298-f0982b97c9ce CVE-2022-21371 pcap data. 入侵者模拟团队
19	PREVENTED	IPS, CWPP, IDS, WAF	DMZ	172.16.1.101	oracle:weblogic_server	None	TA0001 Initial Access	T1190 Exploit Public-Facing Application	0.15650742	None	x-playbook360-e4f96f9b-5d3d-5f64-b8f5-e8714a383983 CVE-2022-21441 pcap data. 入侵者模拟团队
20	SUCCESS	IPS, CWPP, IDS, WAF	DMZ	172.16.1.101	oracle:weblogic_server	User	TA0001 Initial Access	T1190 Exploit Public-Facing Application	0.15650742	None	x-playbook360-4f177683-f6a0-5b38-a36a-453f8f2df71b CVE-2022-24839 pcap data. 入侵者模拟团队

Adaptive attack path reasoning for BAS

- A Small Real-World Example - Key attack path
  - BAS AGENT H1, centos 7

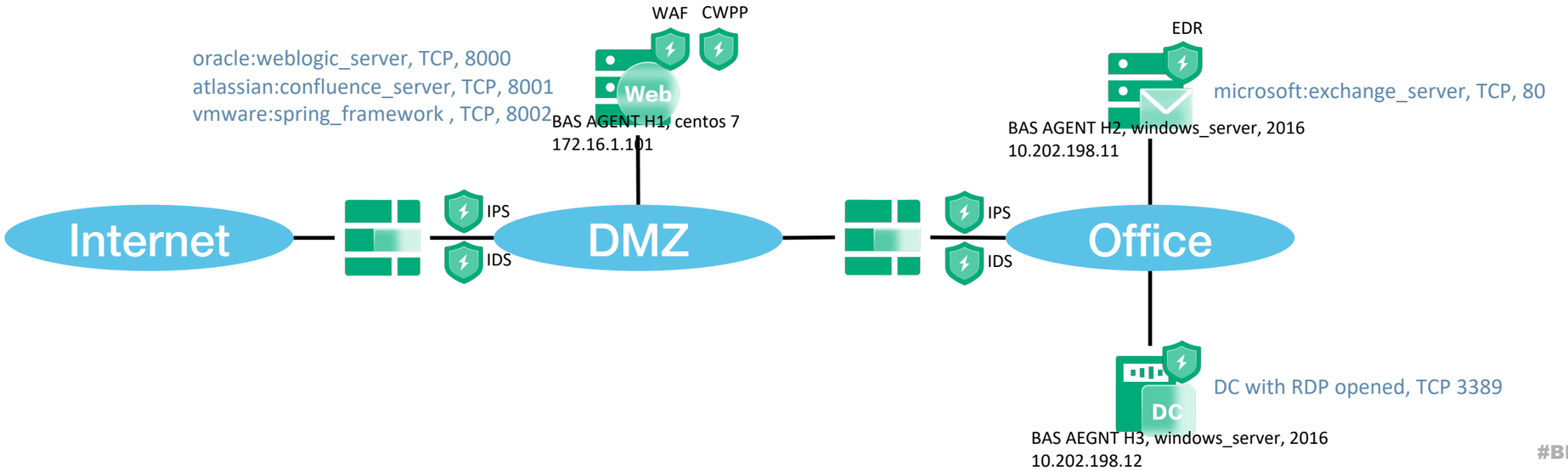
    2. According to previous step simulation attack result: obtained *[User]* permission, reason next TTP playbooks to attack *[Linux]*, loop until Credential Access phase.

Step	Result	Security Products	Area	Agent IP	Asset	Privilege Obtained	Tactic	Technique	Possibility	Privilege Required	Playbook
20	SUCCESS	IPS, CWPP, IDS, WAF	DMZ	172.16.1.101	oracle:weblogic_server	User	TA0001 Initial Access	T1190 Exploit Public-Facing Application	0.15650742	None	x-playbook360-4f177683-f6a0-5b38-a36a-453f8f2df71b CVE-2022-24839 pcap data. 入侵者模拟团队
21	PREVENTED	IPS, CWPP, IDS, WAF	DMZ	172.16.1.101	Linux	User	TA0003 Persistence	T1078.003 Local Accounts	0.026086956	User, Administrator	x-playbook360-1a1e2ac6-33c6-5ec9-adcd-b519528c9471 Login as nobody (Linux) 入侵者模拟团队
22	PREVENTED	IPS, CWPP, IDS, WAF	DMZ	172.16.1.101	Linux	User	TA0004 Privilege Escalation	T1078.003 Local Accounts	0.026086956	User, Administrator	x-playbook360-1a1e2ac6-33c6-5ec9-adcd-b519528c9471 Login as nobody (Linux) 入侵者模拟团队

Adaptive attack path reasoning for BAS

- A Small Real-World Example - Key attack path
  - BAS AGENT H1, centos 7

    3. Lateral Movement phase: based on previous step simulation attack result: owned *[User]* permission of H1, and running service *[microsoft:exchange_server]* on BAS AGENT H2.



Adaptive attack path reasoning for BAS

- A Small Real-World Example - Key attack path
  - BAS AGENT H2, windows server 2016
    1. According to exposed *[microsoft:exchange_server]*, selects corresponding vulnerability pcap playbooks

67	PREVENTED	IPS, EDR, EPP, IDS	OFFICE	10.202.198.11	microsoft:exchange_server	None	TA0001 Initial Access	T1190 Exploit Public-Facing Application	0.15650742	None	x-playbook360-ce518c10-d4ec-5c17-934d-dd71f27c8886 CVE-2022-41082 pcap data. 入侵者模拟团队
68	PREVENTED	IPS, EDR, EPP, IDS	OFFICE	10.202.198.11	microsoft:exchange_server	None	TA0001 Initial Access	T1190 Exploit Public-Facing Application	0.15650742	None	x-playbook360-5f8d9520-4f8a-534f-b972-885e8eef5d1a CVE-2023-21529 pcap data. 入侵者模拟团队
69	SUCCESS	IPS, EDR, EPP, IDS	OFFICE	10.202.198.11	microsoft:exchange_server	SYSTEM	TA0001 Initial Access	T1190 Exploit Public-Facing Application	0.15650742	None	x-playbook360-47ac2637-4bf8-521d-aa77-2cae41275a88 CVE-2023-21706 pcap data. 入侵者模拟团队

Adaptive attack path reasoning for BAS

- A Small Real-World Example - Key attack path
  - BAS AGENT H2, windows server 2016

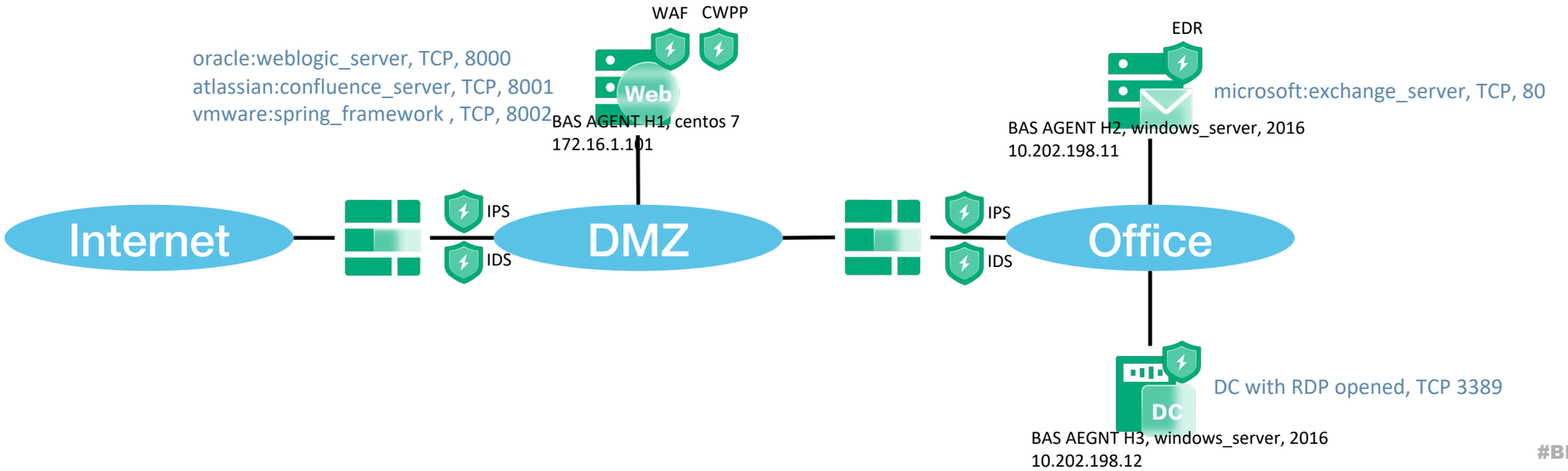
    2. According to previous step simulation attack result: obtained *[System]* permission, reason next TTP playbooks to attack *[windows server 2016]*, loop until Credential Access phase.

145	PREVENTED	IPS, EDR, EPP, IDS	OFFICE	10.202.198.11	Windows	SYSTEM	TA0006 Credential Access	T1003.001 LSASS Memory	0.15946844	None	x-playbook360-faddfa59-4811-5366-aff3-303ed6fff1e3 Dump LSASS.exe Memory using NanoDump 入侵者模拟团队
146	PREVENTED	IPS, EDR, EPP, IDS	OFFICE	10.202.198.11	Windows	SYSTEM	TA0006 Credential Access	T1003.001 LSASS Memory	0.15946844	None	x-playbook360-7074f1bc-9363-5d79-b264-83b4b4e75182 Dump LSASS.exe Memory using Out-Minidump.ps1 入侵者模拟团队
147	SUCCESS	IPS, EDR, EPP, IDS	OFFICE	10.202.198.11	Windows	SYSTEM	TA0006 Credential Access	T1003.001 LSASS Memory	0.15946844	None	x-playbook360-76699fac-380d-59a8-bea4-57d2133df8f5 Dump LSASS.exe Memory using ProcDump 入侵者模拟团队

Adaptive attack path reasoning for BAS

- A Small Real-World Example - Key attack path
  - BAS AGENT H2, windows server 2016

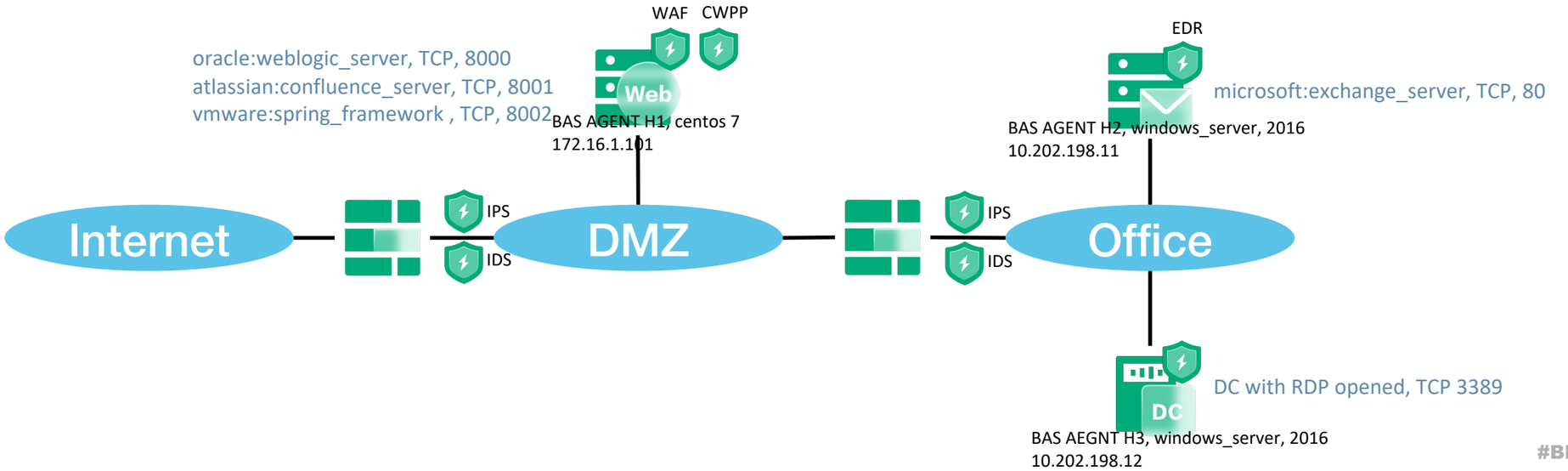
    3. Lateral Movement phase: based on previous step attack result: *[credential of H3 Administrator dump successfully]*, running service on BAS AGENT H3 *[RDP].*



Adaptive attack path reasoning for BAS

- A Small Real-World Example - Key attack path
  - BAS AGENT H3, windows server 2016

    1. If successfully traversed to this server via RDP, reason next TTP playbooks to attack *[windows server 2016]* in a loop until the end.



Adaptive attack path reasoning for BAS

- Technology Stack
  - Protégé, RDF/OWL, SPAQL
  - Jena with hybrid rule engine
    - ✓ based on the standard RETE algorithm, incrementally compute support
    - ✓ Logic Programming Engine with Tabling
- Performance
  - JVM
    - ✓ Xms1024m, Xmx10240m
  - Average reason speed
    - ✓ 30s/step

Demo

The Tool

Live soon: https://github.com/Qihoo360/Luwak

BLACK HAT SOUND BYTES

- Three key problems to improve the accuracy of TTP extraction helps defender keep up with the TTPs of attackers.

- A practical approach for building TTP-oriented knowledge graph can help BAS reason more adaptive attack paths to assess the entire defense-in-depth of the target organization.