Forecasting ATT&CK Flow based on APT by Recommendation System

Agenda

01 Introduction

02 Preliminary

03 About our tool

Agenda

01 Introduction

02 Preliminary

03 About our tool

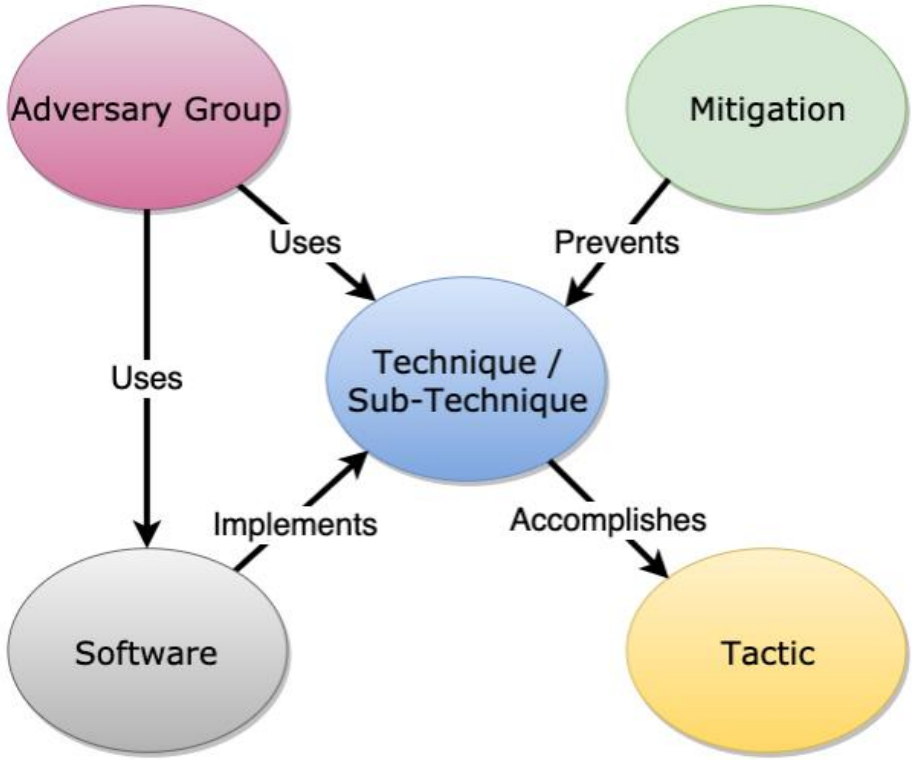
Motivation

- Cyber attacks are causing tremendous damage around the world
- To protect against attacks, many organizations have established or outsourced Security Operation Centers(SOCs)
- Large volumes of logs need to be analyzed to detect signs of an attack quickly in SOC.
- Therefore, there is a need for a method of efficiently analyzing logs

We propose a novel tool that uses collaborative filtering to forecast and visualize attacker behavior from MITRE ATT&CK data

ATT&CK

- Knowledge base provided by MITRE, a non-profit organization in the U.S
- Based on actual observed attackers(groups) and their tactics・techniques
  - Adversary Group(Group)
    - Attacker
  - Technique
    - Technology used in the attack.
  - Tactic
    - Objective to be achieved by technique
  - Software
    - Tools used by the attacker
  - Mitigation
    - Measure to mitigate against attacks.



The five elements in ATT&CK

https://attack.mitre.org/docs/ATTACK_Design_and_Philosophy_March_2020.pdf

ATT&CK Group

- There is information on which groups have used which techniques in the past.
- Each group is assigned a 5-digit ID like Gxxxx.
- There are 133 groups

  (v11, April 25, 2022)

https://attack.mitre.org/groups/G0045/

menuPass

G0045

menuPass is a threat group that has been active since at least 2006. Individual members of menuPass are known to have acted in association with the Chinese Ministry of State Security's (MSS) Tianjin State Security Bureau and worked for the Huaying Haitai Science and Technology Development Company.[1][2]

menuPass has targeted healthcare, defense, aerospace, finance, maritime, biotechnology, energy, and government sectors globally, with an emphasis on Japanese organizations. In 2016 and 2017, the group is known to have targeted managed IT service providers (MSPs), manufacturing and mining companies, and a university.[3][4][5][6][7][1][2]

Techniques Used

ATT&CK® Navigator Layers

Domain	ID		Name	Use
Enterprise	T1087	.002	Account Discovery: Domain Account	menuPass has used the Microsoft administration tool csvde.exe to export Active Directory data.[11]
Enterprise	T1583	.001	Acquire Infrastructure: Domains	menuPass has registered malicious domains for use in intrusion campaigns.[1][2]
Enterprise	T1560		Archive Collected Data	menuPass has encrypted files and information before exfiltration.[1][2]
		.001	Archive via Utility	menuPass has compressed files before exfiltration using TAR and RAR.[6][11][8]
Enterprise	T1119		Automated Collection	menuPass has used the Csvde tool to collect Active Directory files and data.[8]
Enterprise	T1059	.001	Command and Scripting Interpreter: PowerShell	menuPass uses PowerSploit to inject shellcode into PowerShell.[11][8]
		.003	Command and Scripting Interpreter: Windows Command Shell	menuPass executes commands using a command-line interface and reverse shell. The group has used a modified version of pentesting script wmiexec.vbs to execute commands.[6][11][12][10] menuPass has used malicious macros embedded inside Office documents to execute files.[9][10]
Enterprise	T1005		Data from Local System	menuPass has collected various files from the compromised computers.[1][8]
Enterprise	T1039		Data from Network Shared Drive	menuPass has collected data from remote systems by mounting network shares with `net use` and using Robocopy to transfer data.[6]
Enterprise	T1074	.001	Data Staged: Local Data Staging	menuPass stages data prior to exfiltration in multi-part archives, often saved in the Recycle Bin.[6]
		.002	Data Staged: Remote Data Staging	menuPass has staged data on remote MSP systems or other victim networks prior to exfiltration.[6][8]

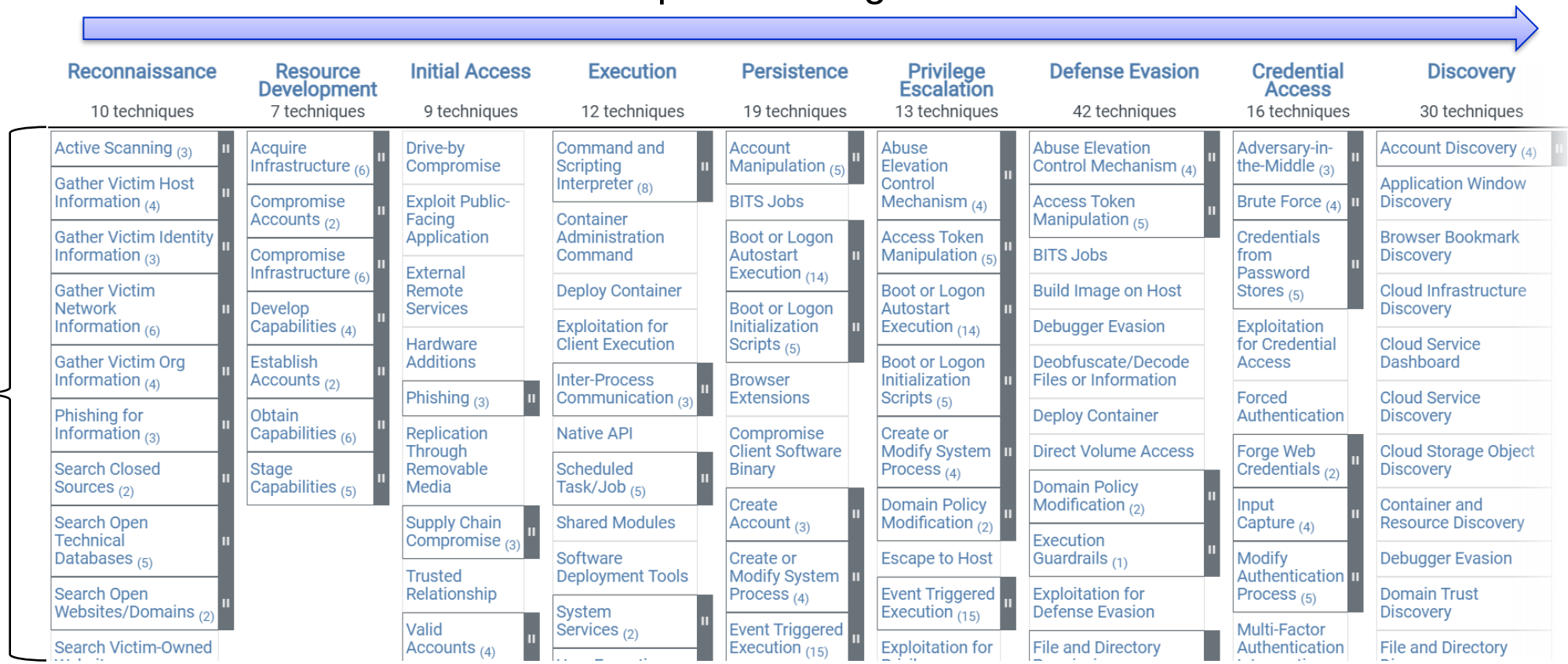
ATT&CK Tactic and Technique

191techniques (5-digit ID like Txxxx)
14 tactics
(v11, April 25, 2022)

- How to View Enterprise Matrix

Tactics：Represent stages of the attack.

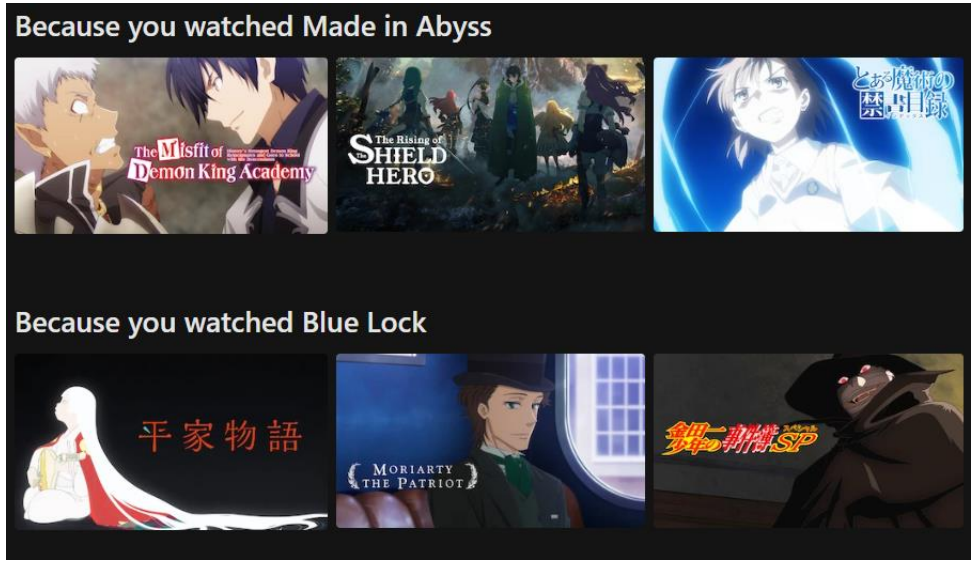
Techniques



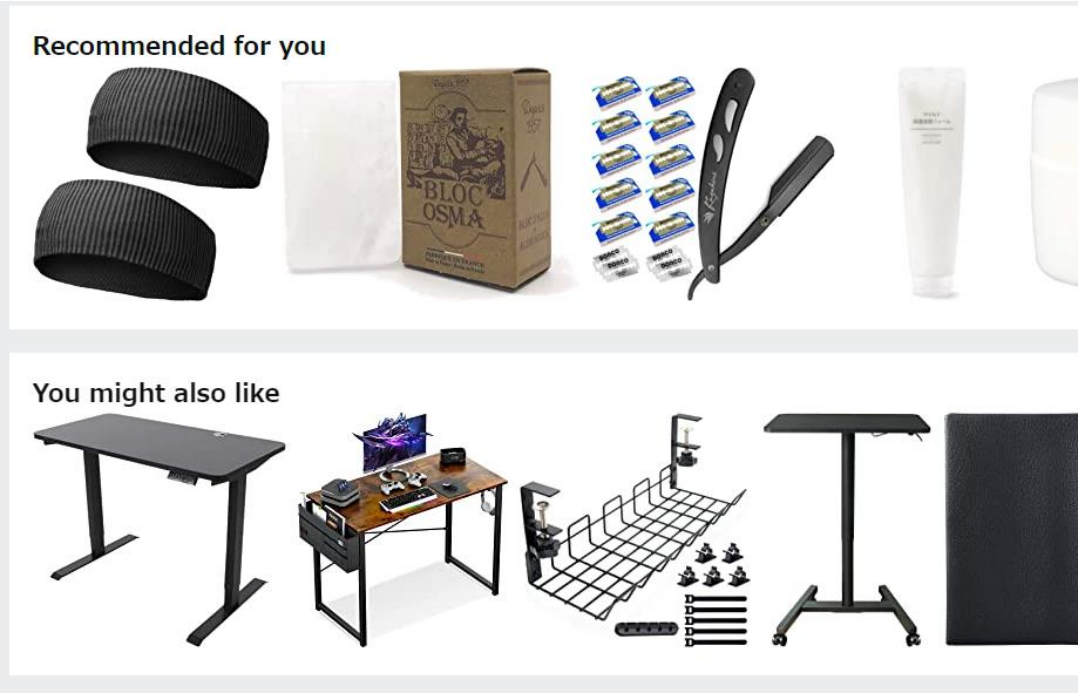
https://attack.mitre.org/matrices/enterprise/

Recommendation System

- Recommendation systems suggest products based on user's tastes.
- You will be surprised at how well the system guesses your preferences!
- It can predict your future behavior.



Netflix



Amazon

Our Core Idea

- Replacing with ATT&CK, each group can be considered as a user, and techniques used by that group can be considered as a user's purchase history
- It is possible to predict which techniques an attacker may use in the future, based on the techniques already detected

Agenda

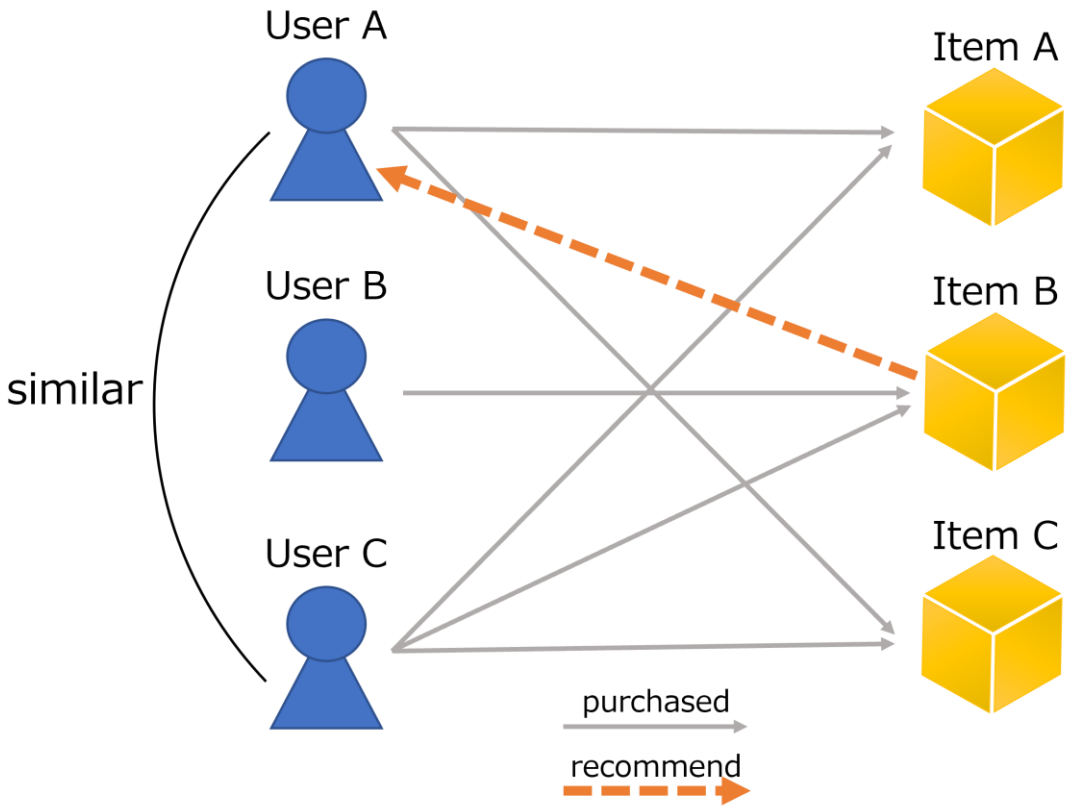
01 Introduction

02 Preliminary

03 About our tool

User based Collaborative Filtering

- User-based recommends products based on the similarity of purchase history between users.



Since Item A and C are common, User A and User C are considered similar.

The system recommends Item B to User A, which has been purchased by User C with a high similarity and has not yet been purchased by User A.

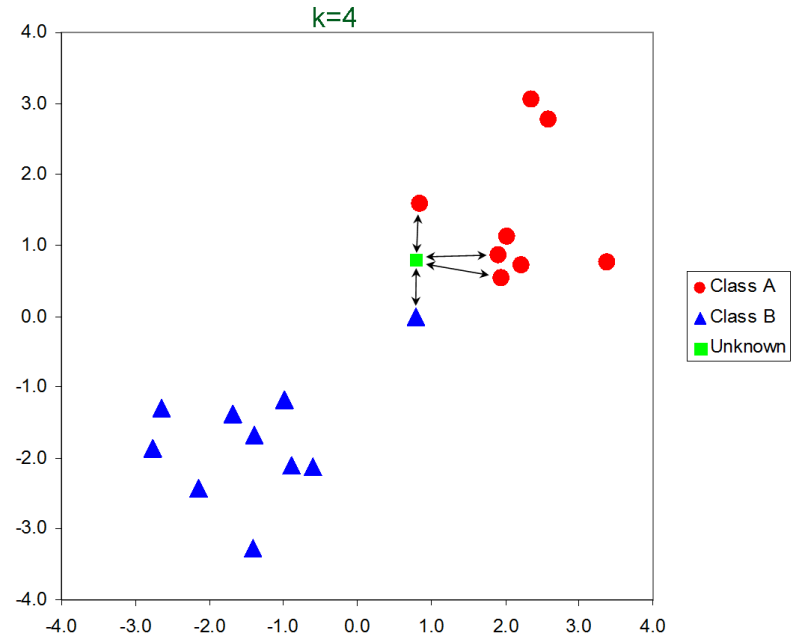
k-Nearest-Neighbor（*k*NN）

- One of Classification Method (Also for Regression)
  - Algorithm is as follows
    1. Calculate distance to between each data with known class belongings and unknown
    2. Select the *k* known data closest to the unknown data
    3. Take a majority vote for class with the *k*
       (The value of *k* is selected)
    4. Classify unknown data into the most voted class
  - Example for *k*=4 (Figure2)

    The 4 closest data from Unknown data

    ●(class A) × 3　▲(class B) × 1

    → Unknown data is classified as class A

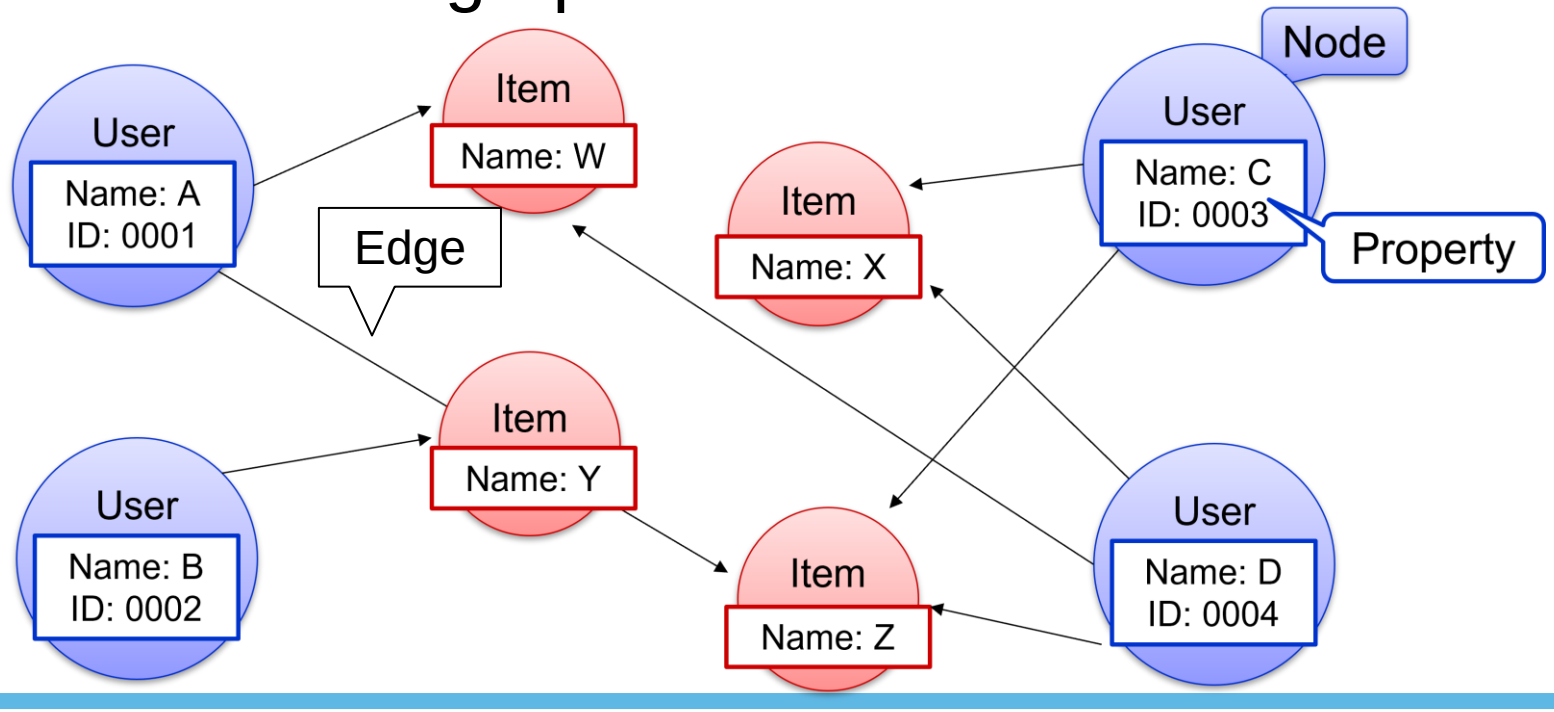


- We used W*k*NN for collaborative filtering algorithm
- WkNN considers distance as a weight in the majority vote process, giving more weight to those that are closer in distance.

http://www.scholarpedia.org/article/K-nearest_neighbor

Graph Database

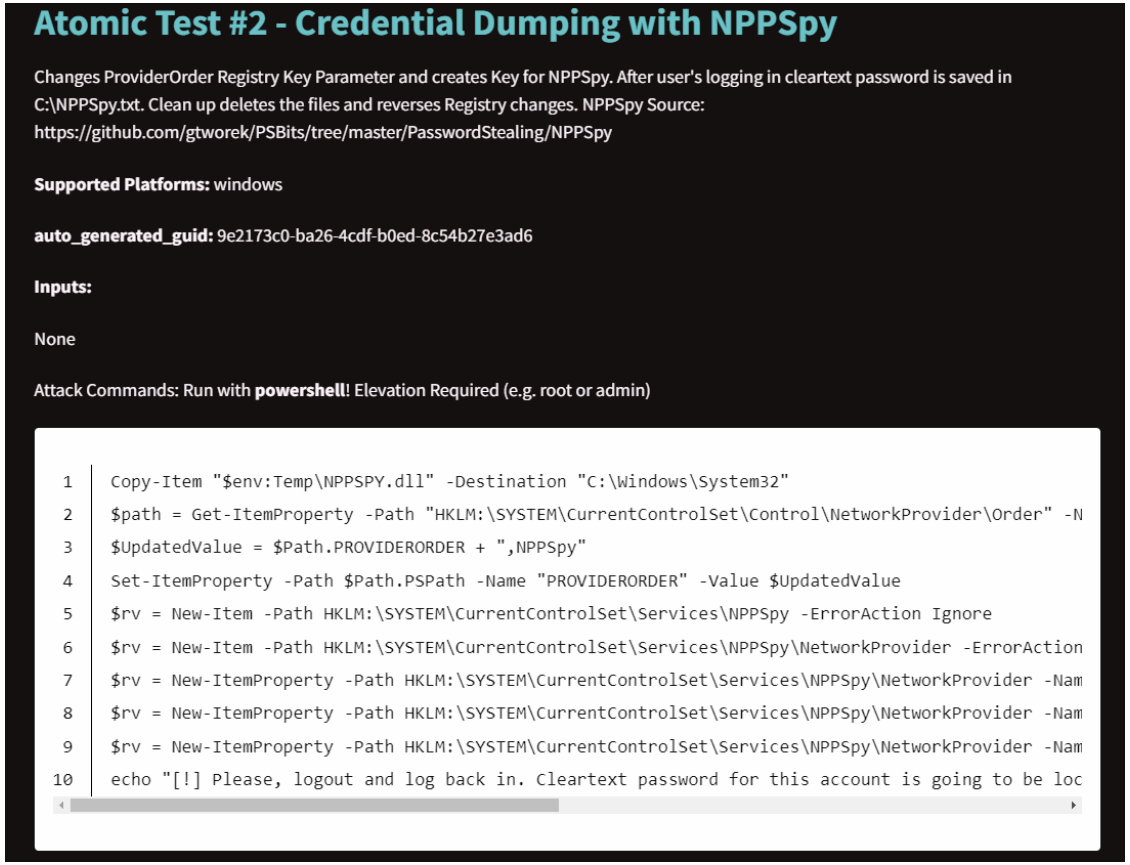
- A database based on a graph structure consisting of three elements: nodes, edges, and properties.
- It has the advantage of searching faster.
- Neo4j is used to create the graph database



Atomic Red Team

- Atomic Red Team is a test library based on ATT&CK framework.
- Command lines, etc. can be mapped to ATT&CK technique

https://atomicredteam.io/credential-access/T1003/



Agenda

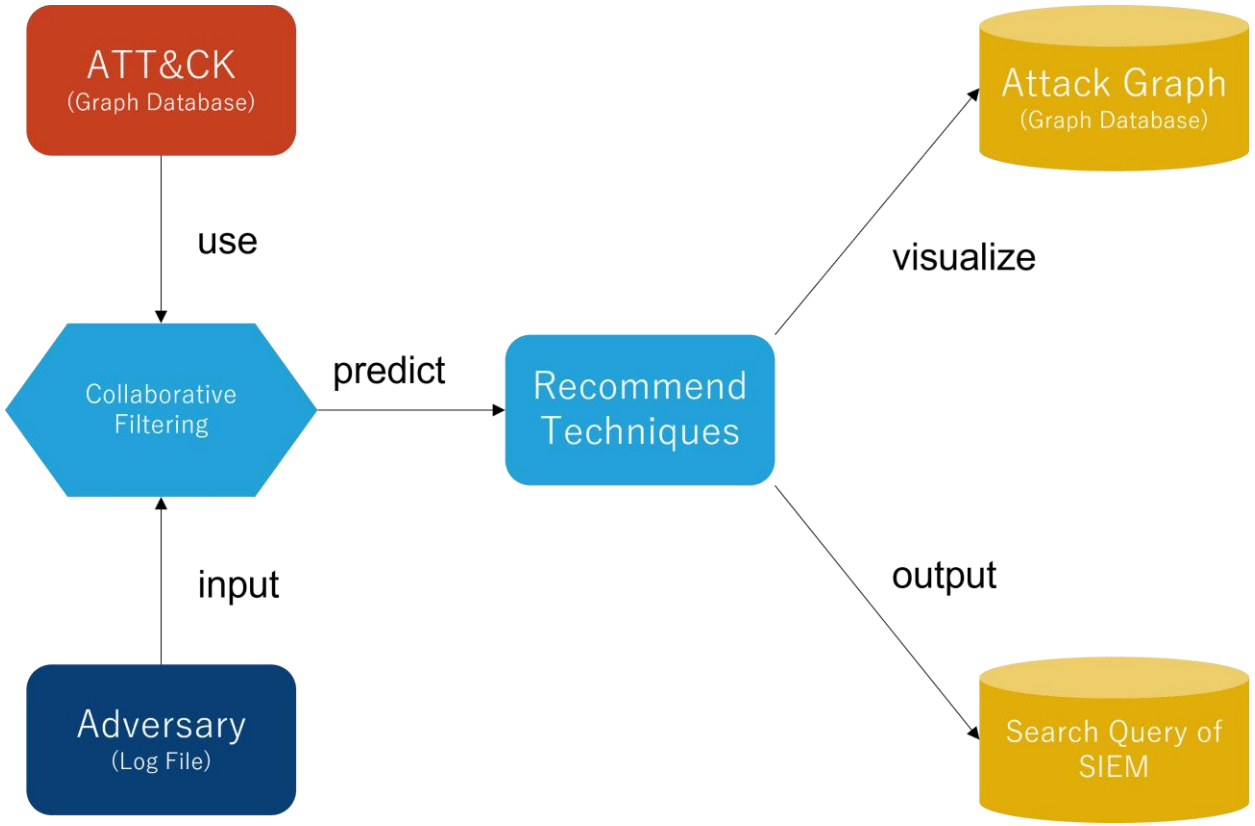
01 Introduction

02 Preliminary

03 About our tool

Our Tool (Prediction Flow)

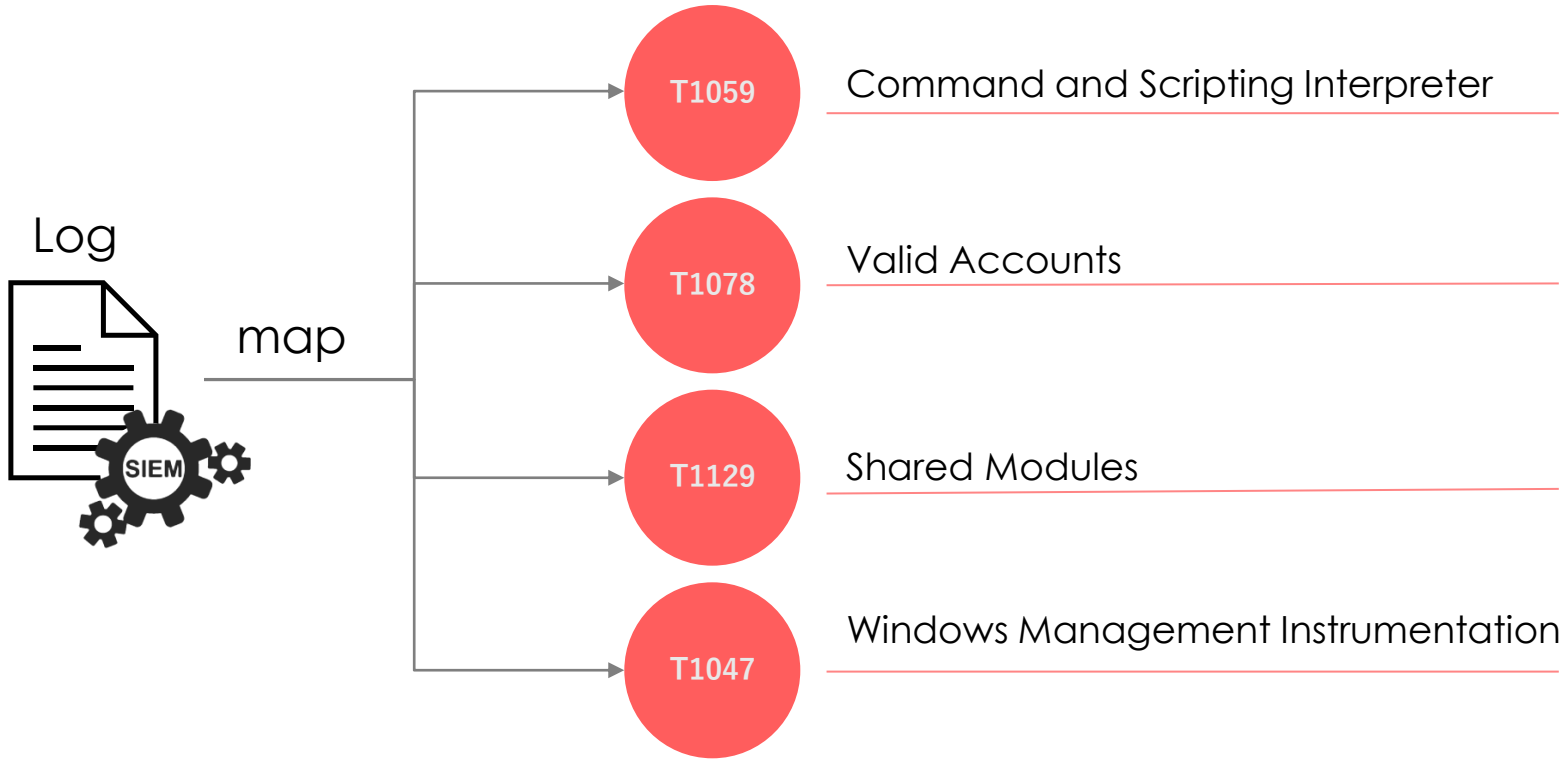
- Groups and techniques from the ATT&CK data are used as training data for collaborative filtering.
- The input is log file
- Recommended techniques can be considered as attack predictions and visualized as a graph database.
- Search query of SIEM mapped from technique is outputted



We refer to the ongoing attacker as “Adversary”

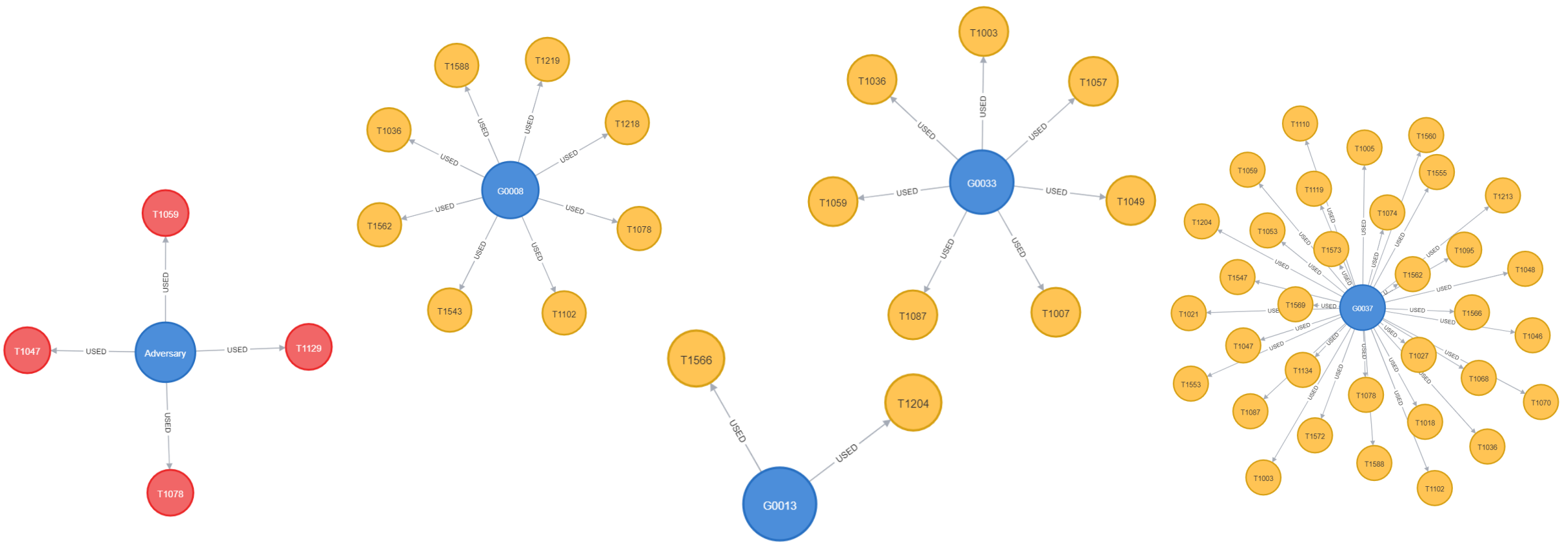
Our Tool

- Step 1 : Mapping from Logs to ATT&CK Techniques
  - Map from SIEM log to ATT&CK technique using database created based on Atomic Red Team



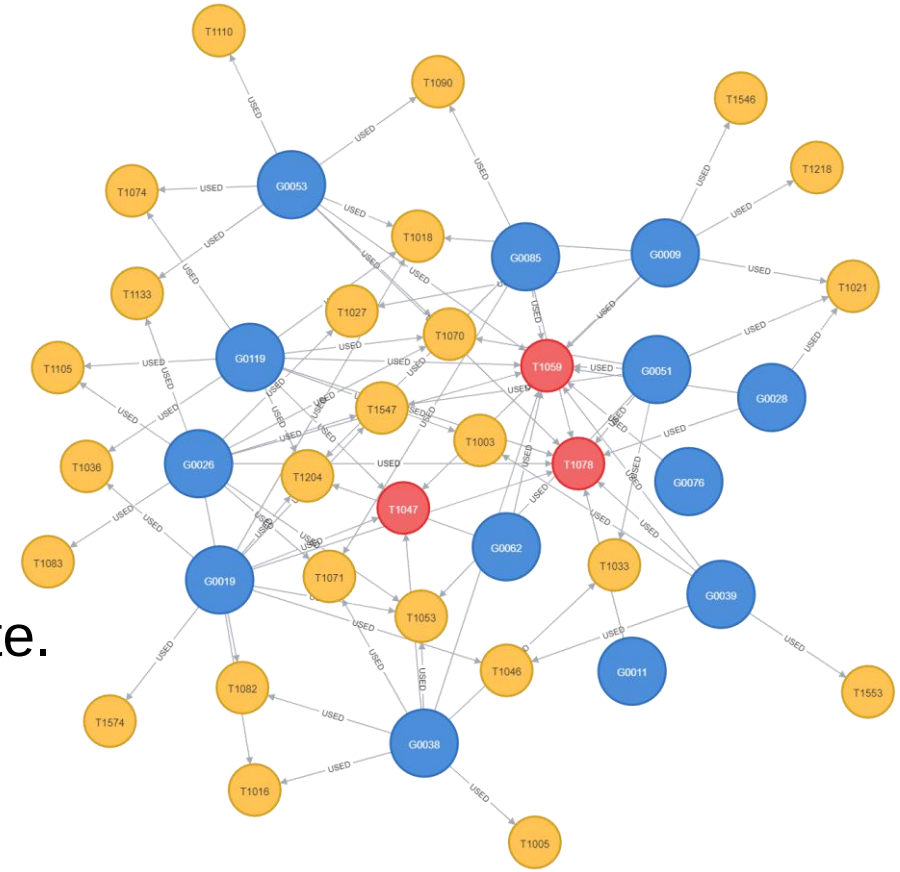
Our Tool

- Step 2 : Recommendation on Graph Databases
  - There are technique usage history for each of these groups as graph database.
  - Create the Adversary data from the technique in Step 1



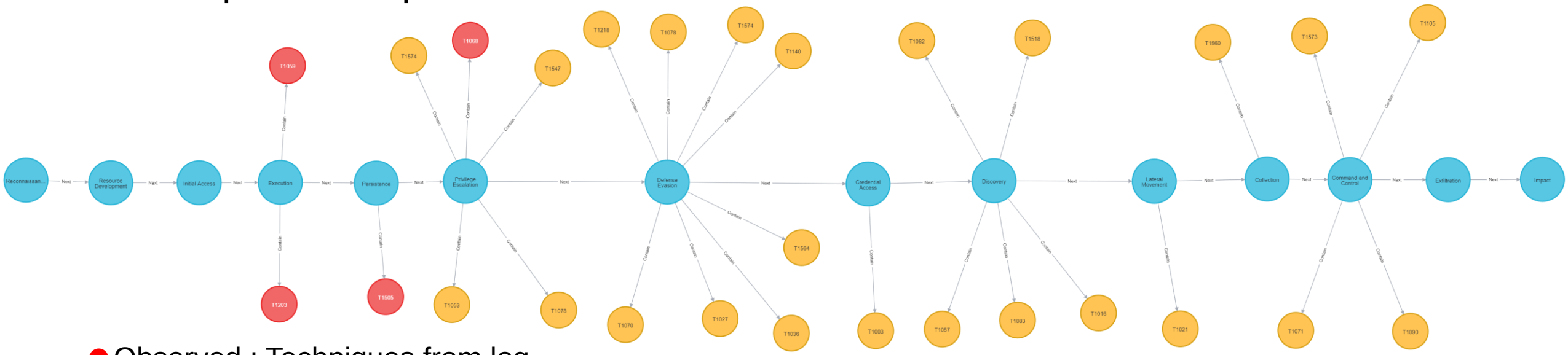
Our Tool

- Step 2 : Recommendation on Graph Databases
  - Calculate the similarity between the adversary and each group.
  - Consider the top *k* groups with high similarity are similar to the adversary
  - Calculate support rate considering similarity among the *k* groups
  - Recommend techniques above a certain support rate.



Our Tool

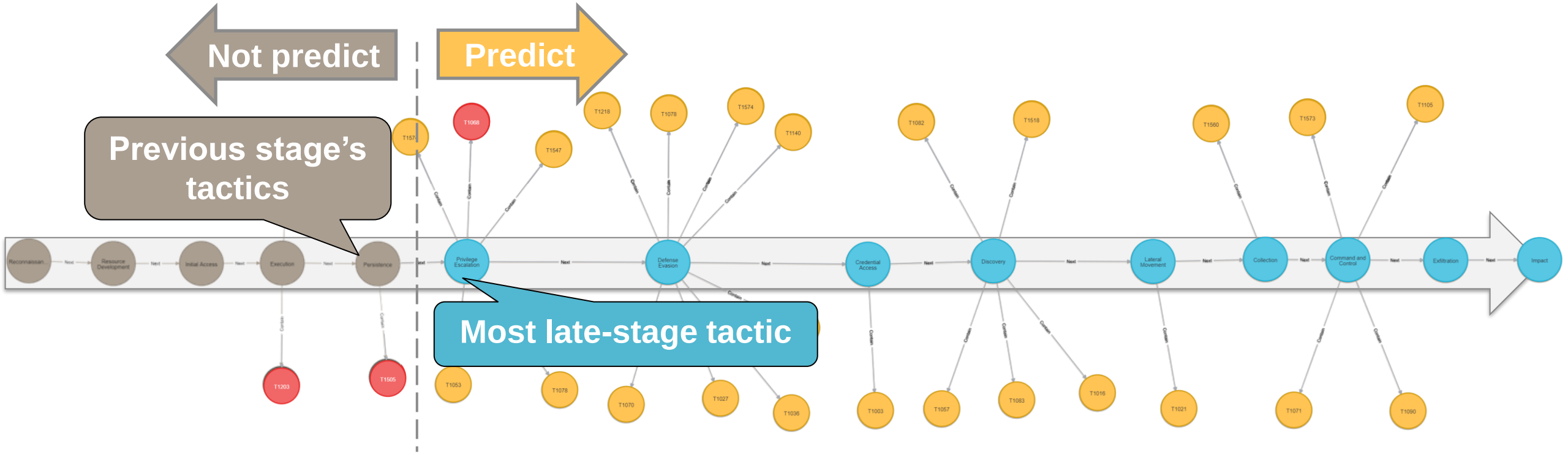
- Step 3: Visualization
  - This is attack predictions.
  - In the figure below, 22 techniques were predicted to be used later, and 128 techiques were predicted not to be used.



- Observed : Techniques from log
- Forecast : Techniques that may be used by the Adversary
- Tactics : Stages of Attack

Our Tool

- Important to note:
  - Predicting techniques in the previous stage's tactics doesn't help analysis
  - Predict only techniques included after the most late-stage tactic



Our Tool

- Step 4 : Mapping from predicted ATT&CK techniques to Search Query of SIEM
  - In the form of technique, SOC analysts cannot use the forecasting results effectively
  - So, re-map the predicted techniques to search query of SIEM