ALIBABA SECURITY

Solving The Last Mile Problem Between Machine Learning and Security Operations

Xiangyu Liu, Xinyue Shen

阿里安全 ALIBABA SECURITY

hackinthebox Keeping Knowledge Free for Over a Decade

Whoami

- Xiangyu Liu
  - Senior Algorithm Engineer @Alibaba Security
  - CUHK PhD (2016)
  - Academic: IEEE S&P, ACM CCS
  - Industry: DEF CON, Black Hat Asia
  - Interests: Machine Learning, Cybersecurity
- Xinyue Shen
  - Algorithm Engineer Intern @Alibaba Security
  - Interests: Cybersecurity, NLP, Knowledge Graph
- Special Thanks
  - Tao Zhou, Quan Lu, Security Operation Team @Alibaba Security

What is Security Operations?

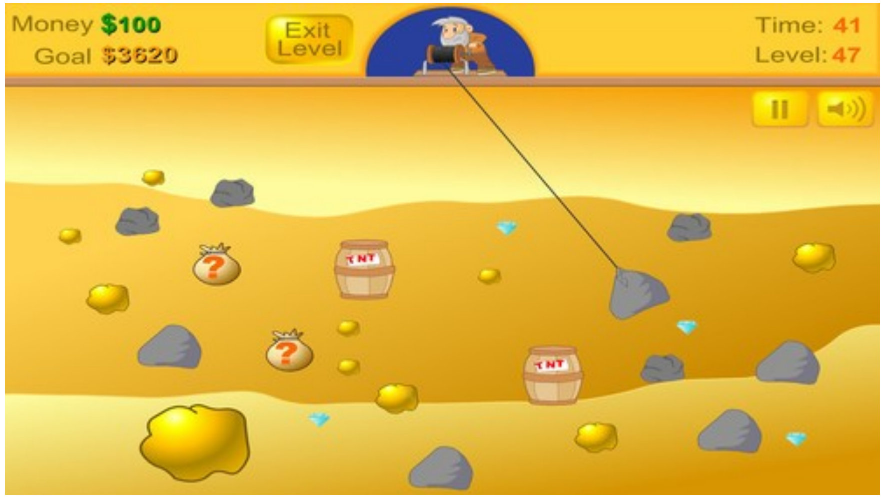
A security operations center (SOC) is a centralized unit that deals with security issues on an organizational and technical level.

——WikiPedia

What is Security Operations ?

What others think I do

What I think I do



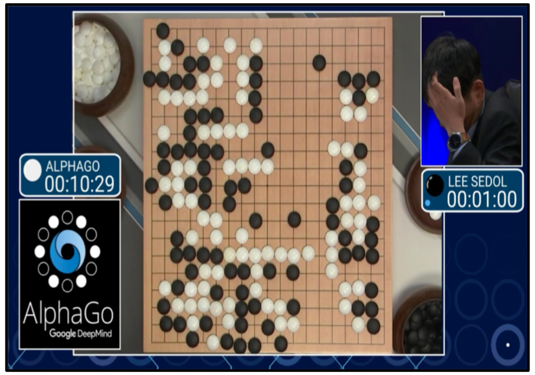
What I really do

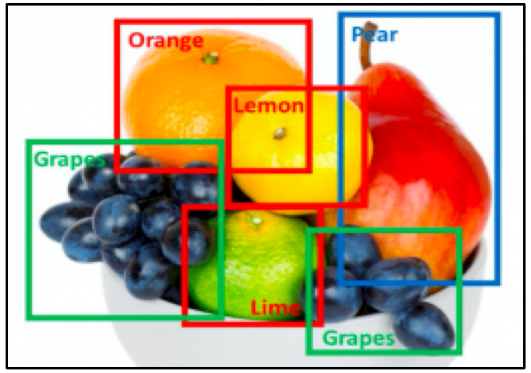
Why not introduce **Machine Learning** in **SOC** ?

Challenges

Partially Observable	Uncertainty	Correlation	Strong Interpretability
Hard to collect all security-related data	Depend on attackers and environment	Current decisions affect subsequent	Security needs strong interpretability

≠





Challenges

Partially Observable

Hard to collect all security-related data

Uncertainty

Depend on attackers and environment

Correlation

Current decisions affect subsequent

Strong Interpretability

Security needs strong interpretability

≈

What ML can do in Security

- Data + Close Domain+ Quantitative Expert Experience

Application Security	Spam Email	Porn Identification	Credit Card Fraud	……
Web Security	Phishing	Botnet	XSS	……
System Security	PUF	PBS	Device Authentication	……

Application of Machine Learning in Cyberspace Security Research. Lei Zhang, Yong Cui, Jing Liu, Yong Jiang, Jianping Wu. Chinese Journal of Computers, 2017.

Is there anything wrong when they meet SOC?

The Gap Between Machine Learning and Security Operations

Data Scientists

“The Accuracy Rate of Our Model is 99.9%!”

Security Operation Experts

“Sounds good. But our data scale is enormous. Over 100 million every day.”

“So, even the accuracy is high, your model will still produce 100000 alerts every day….”

“Well …. How many alerts can you handle?”

“only 100 alerts per day!"

The Gap Between Machine Learning and Security Operations

Data Scientists

Produce 100000 alerts per day

Security Operation Experts

Handle 100 alerts per day

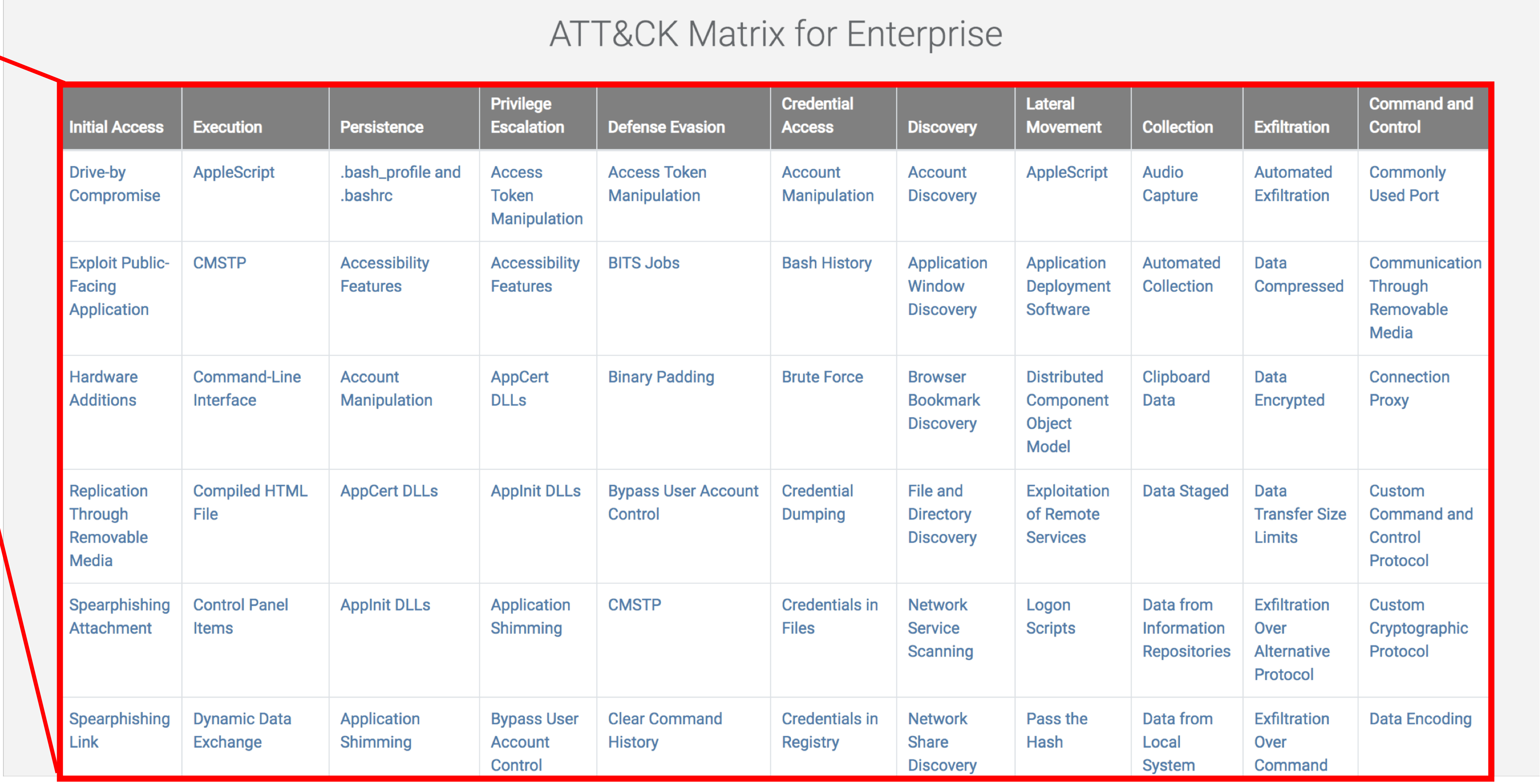
“And this is only one model."

The Gap Between Machine Learning and Security Operations

Produce 100000 alerts per day

Handle 100 alerts per day

“How many attack types we may meet in reality?"



ATT&CK Matrix for Enterprise

Initial Access	Execution	Persistence	Privilege Escalation	Defense Evasion	Credential Access	Discovery	Lateral Movement	Collection	Exfiltration	Command and Control
Drive-by Compromise	AppleScript	.bash_profile and .bashrc	Access Token Manipulation	Access Token Manipulation	Account Manipulation	Account Discovery	AppleScript	Audio Capture	Automated Exfiltration	Commonly Used Port
Exploit Public-Facing Application	CMSTP	Accessibility Features	Accessibility Features	BITS Jobs	Bash History	Application Window Discovery	Application Deployment Software	Automated Collection	Data Compressed	Communication Through Removable Media
Hardware Additions	Command-Line Interface	Account Manipulation	AppCert DLLs	Binary Padding	Brute Force	Browser Bookmark Discovery	Distributed Component Object Model	Clipboard Data	Data Encrypted	Connection Proxy
Replication Through Removable Media	Compiled HTML File	AppCert DLLs	AppInit DLLs	Bypass User Account Control	Credential Dumping	File and Directory Discovery	Exploitation of Remote Services	Data Staged	Data Transfer Size Limits	Custom Command and Control Protocol
Spearphishing Attachment	Control Panel Items	AppInit DLLs	Application Shimming	CMSTP	Credentials in Files	Network Service Scanning	Logon Scripts	Data from Information Repositories	Exfiltration Over Alternative Protocol	Custom Cryptographic Protocol
Spearphishing Link	Dynamic Data Exchange	Application Shimming	Bypass User Account Control	Clear Command History	Credentials in Registry	Network Share Discovery	Pass the Hash	Data from Local System	Exfiltration Over Command	Data Encoding

ATT&CK（Adversarial Tactics, Techniques, and Common Knowledge）is a globally-accessible knowledge base of adversary tactics and techniques based on real-world observations.

——MITRE

Ref. https://attack.mitre.org/

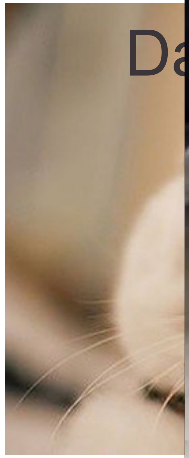
The Gap Between Machine Learning and Security Operations

Produce 100000 alerts per day

Handle 100 alerts per day

“So actually the number of alerts is 100000 × 300 + per day…"

The Gap Between Machine Learning and Security Operations



Produce 1

ts per day

Can we bridge the gap and solve this awkward thing?

Our Solutions

- Behavior analysis
- Feature based sorting
- Ensemble risks
- Knowledge graph
- White list
- …

Best Practices: Large-Scale Data

Porn Identification

- Labeling is easy
- Labeling is relatively cheap
- Lots of samples

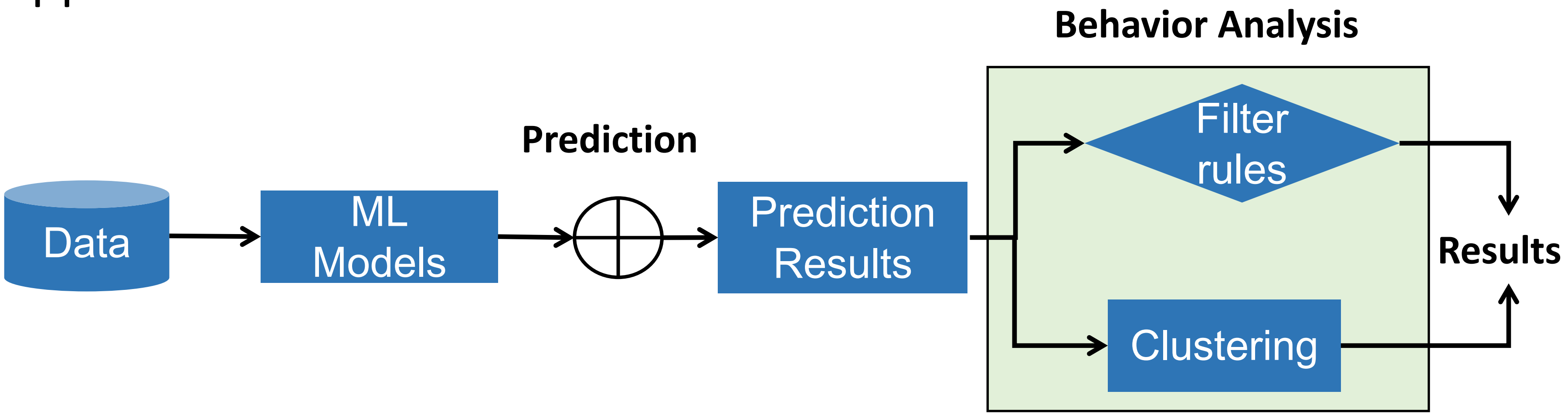
Intrusion detection

- Depend on experience and time consuming
- Security experts are expensive
- Few samples

Best Practices: Behavior Analysis

- A cyber-security problem can be taken as consisting of several subproblems
  - Machine learning can be applied in some part
  - The malicious behaviors can be distinguished by rules or can be clustered

- Our Approach

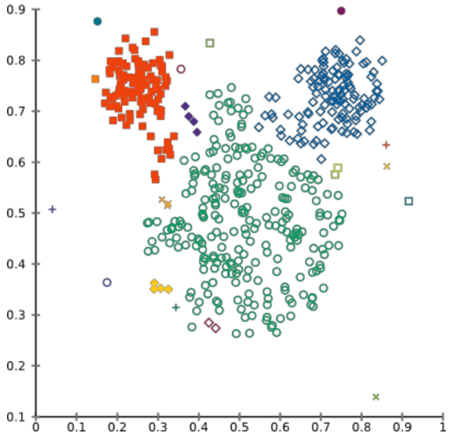


Best Practices: Behavior Analysis

- Example
  - Domain generating algorithm (DGA) detection
  - A DGA is a program that provides malware with new domains
  - **Mistakes**: Using ML to detect DGAs directly

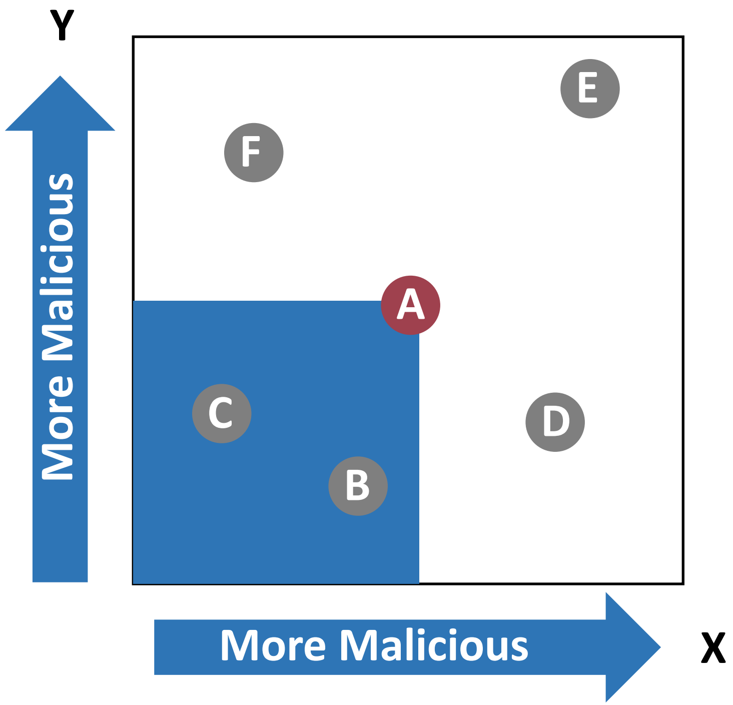
- Approach
  - ML is used to detect the randomness of domains
    - LSTM, Ngram, and etc.
  - Filter rules
    - IP relationship, number of requests, number of subdomains, and etc.
  - Clustering
    - The features described above, and/or embedding techniques

earnestnessbiophysicalohax.com
kwtoestnessbiophysicalohax.com
rvcxestnessbiophysicalohax.com
hjbtestnessbiophysicalohax.com
txmoestnessbiophysicalohax.com
agekestnessbiophysicalohax.com
dbzwestnessbiophysicalohax.com
sgjxestnessbiophysicalohax.com
……



Best Practices: Feature Based Sorting

- Focus on precision
- Feature extraction
  - Assume we have only two features: X and Y
- Scoring:
  - if *A* is more malicious than *B* in every dimension, Increment *A*'s score by one
  - Can be customized
- Sorting:
  - Let N denote all the elements, K as the budget of SOC
  - Sort *N* by each element's score, and select top K elements



Best Practices: Feature Based Sorting

- Compare with historical data
  - Extract features per day/hour/…
  - Sort the data in a longer time window, e.g. one week

- Application
  - Phishing detection, *Usenix Security'17*
  - UEBA
  - …

- Limitations
  - At the expense of recall
  - What features to extract is very hard to determine

Ho, G., Javed, A. S. M., Paxson, V., & Wagner, D. (2017). Detecting Credential Spearphishing Attacks in Enterprise Settings. USENIX Security'17

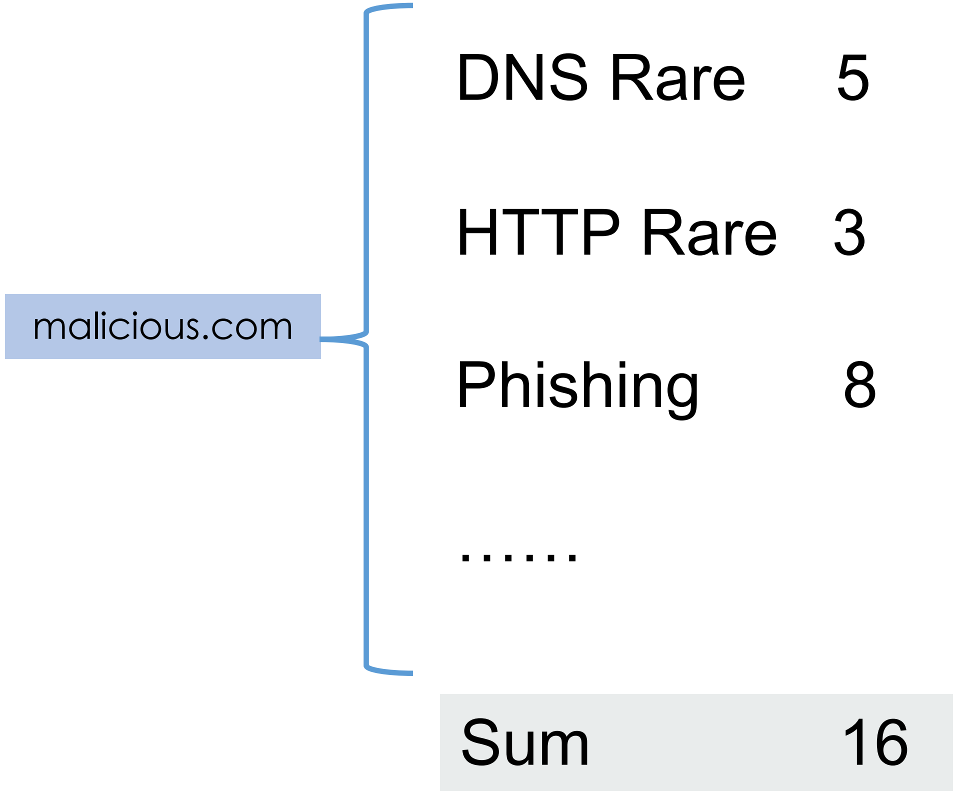
Best Practices: Accumulation Risk

Alerts Pool

1. xxx
2. xxx
3. xxx
4. xxx

Best Practices: Accumulation Risk

Traditional Way:



malicious.com

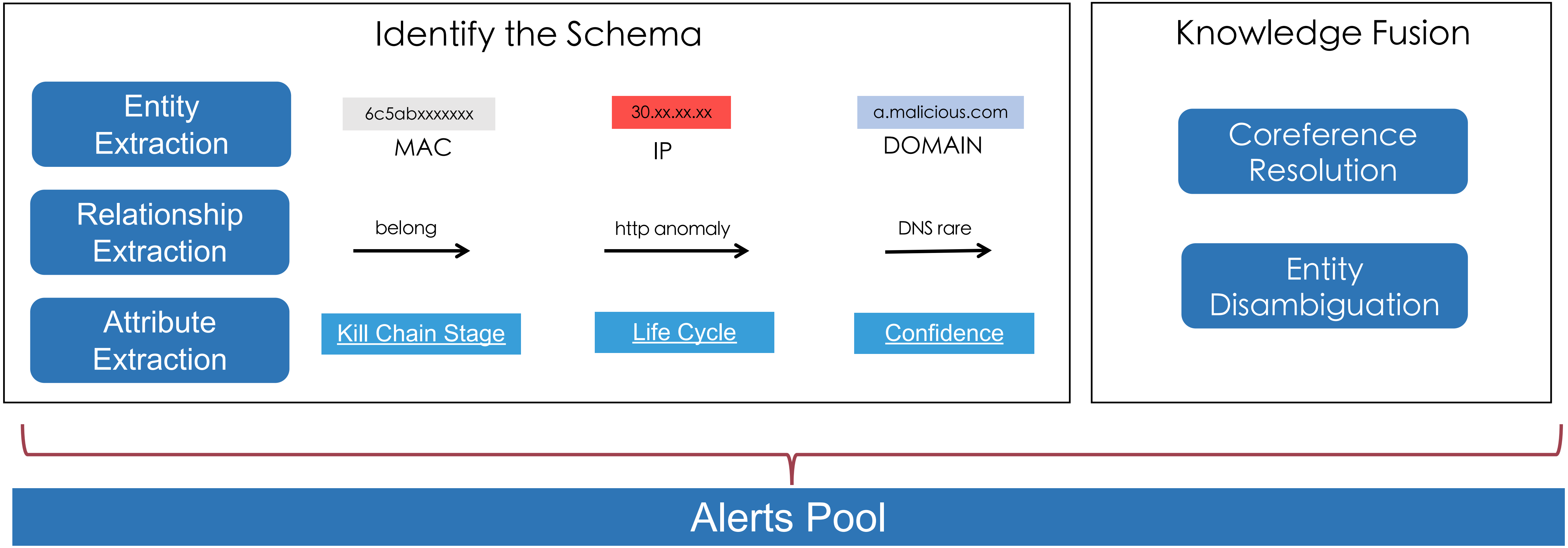
DNS Rare	5
HTTP Rare	3
Phishing	8
……	
Sum	16

Problems behind it:

1. Not all related alerts can be produced.
2. Lateral movement is common.

Best Practices: Knowledge Graph

Alerts Pool Construction



Identify the Schema

Entity Extraction

Relationship Extraction

Attribute Extraction

6c5abxxxxxxx
MAC

30.xx.xx.xx
IP

a.malicious.com
DOMAIN

belong →

http anomaly →

DNS rare →

Kill Chain Stage

Life Cycle

Confidence

Knowledge Fusion

Coreference Resolution

Entity Disambiguation

Alerts Pool

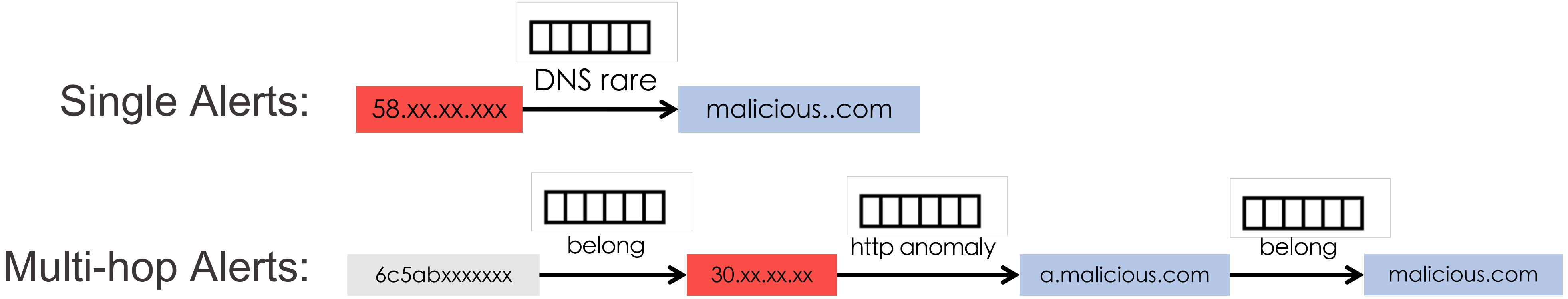
Best Practices: Accumulation Risk

Some attributes
- Kill chain stage
- Life cycle
- Confidence
- ……

After identify the Schema, every alert is a Triple(entity-relationship-entity).

Alerts Pool

1. xxx
2. xxx
3. xxx
4. xxx



Single Alerts: 58.xx.xx.xxx → DNS rare → malicious..com

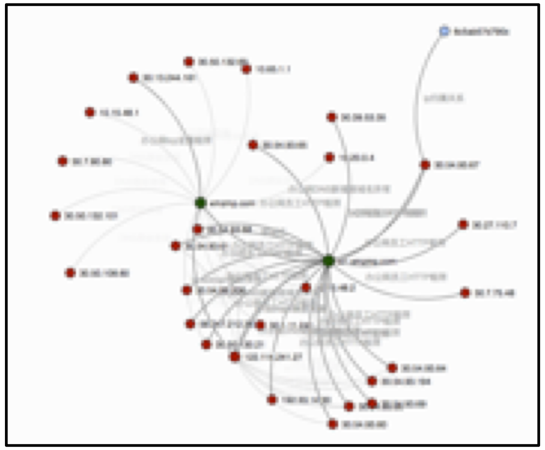
Multi-hop Alerts: 6c5abxxxxxxx → belong → 30.xx.xx.xx → http anomaly → a.malicious.com → belong → malicious.com

An intrusion case is usually combined by **many multi-hop alerts**!

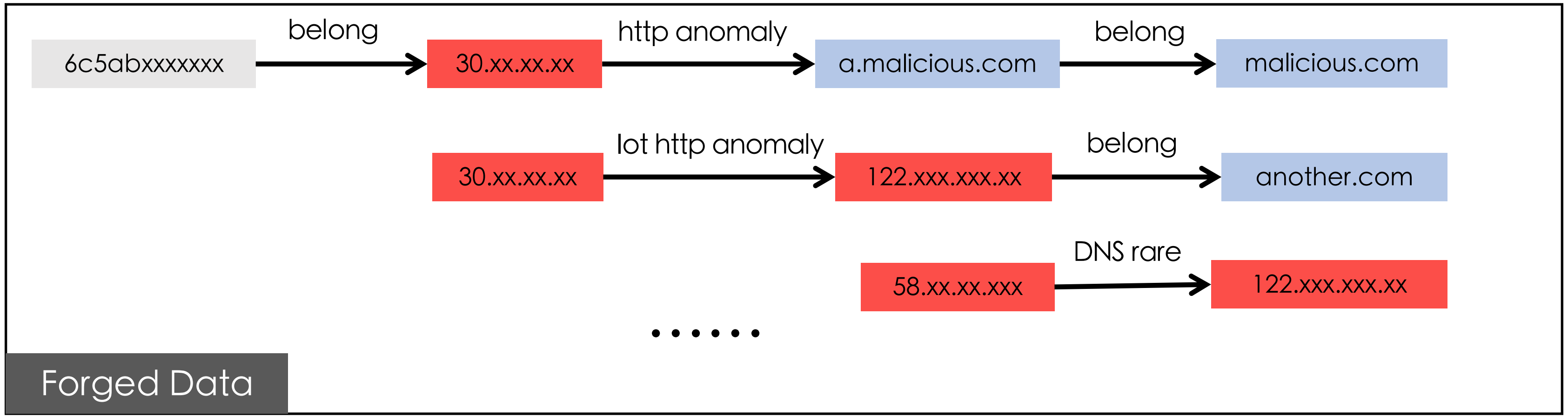
Best Practices: Accumulation Risk

An intrusion case is usually combined by **many multi-hop alerts**!

Eg.



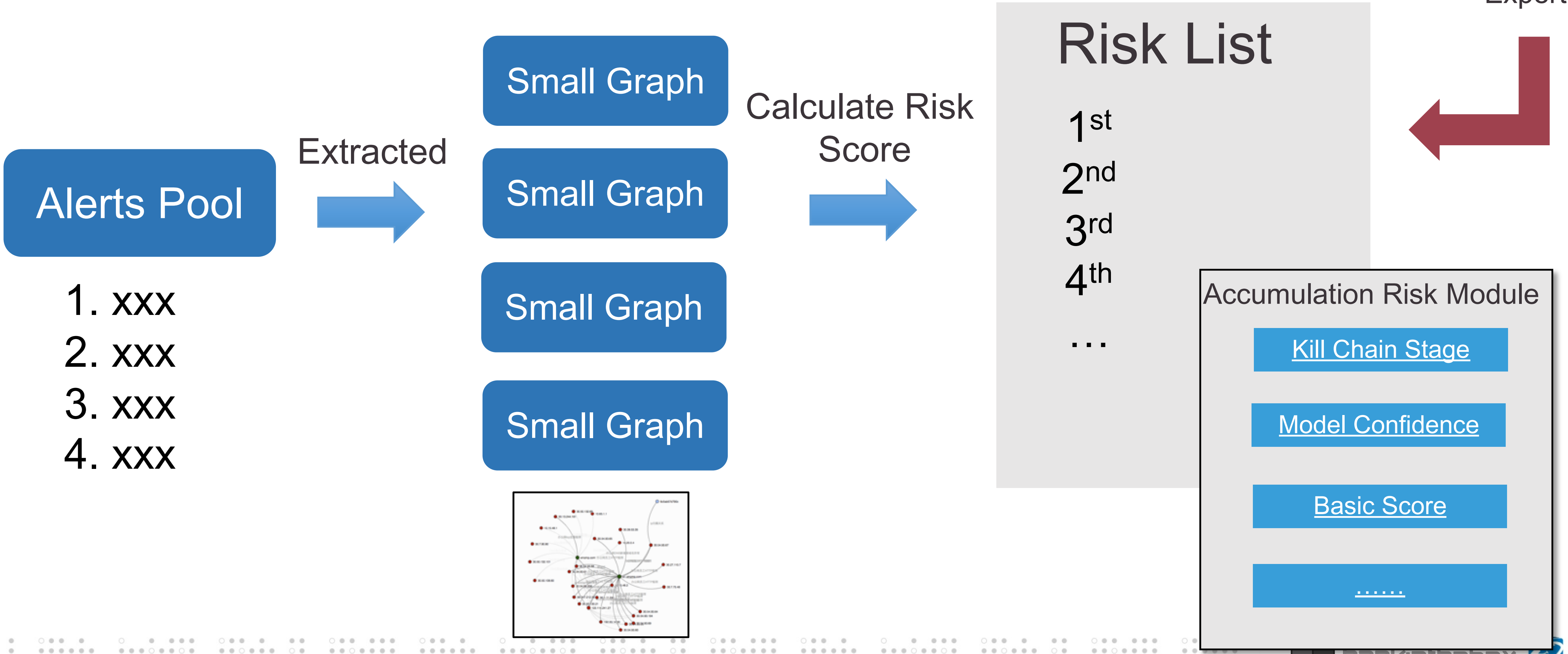
An intrusion graph



Multi-hop alerts

Best Practices: Accumulation Risk

Security Operation Experts



Alerts Pool

1. xxx
2. xxx
3. xxx
4. xxx

Extracted

Small Graph

Small Graph

Small Graph

Small Graph

Calculate Risk Score

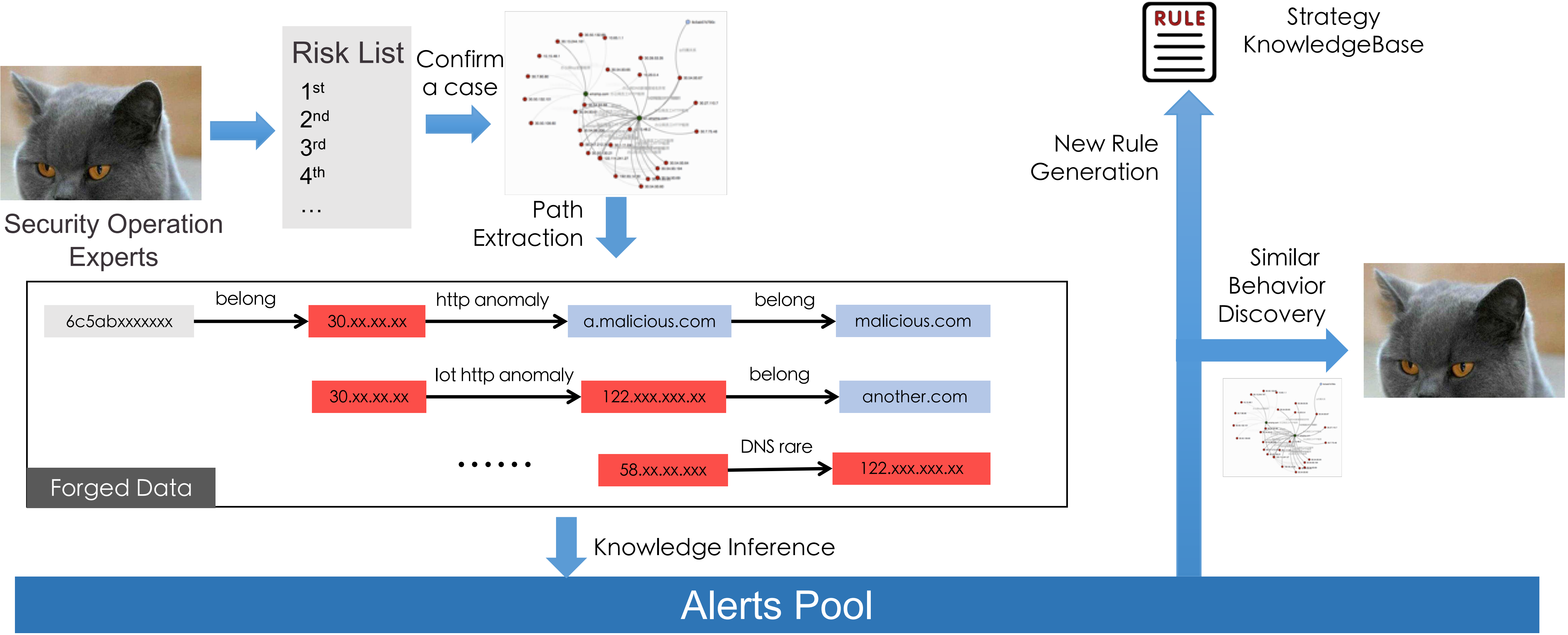
Risk List

1st
2nd
3rd
4th
…

Accumulation Risk Module

- Kill Chain Stage
- Model Confidence
- Basic Score
- ……

Best Practices: Knowledge Graph



Security Operation Experts

Risk List
1st
2nd
3rd
4th
…

Confirm a case

Path Extraction

6c5abxxxxxxx	belong	30.xx.xx.xx	http anomaly	a.malicious.com	belong	malicious.com
		30.xx.xx.xx	Iot http anomaly	122.xxx.xxx.xx	belong	another.com
		·······		58.xx.xx.xxx	DNS rare	122.xxx.xxx.xx

Forged Data

Knowledge Inference

Alerts Pool

New Rule Generation

RULE

Strategy KnowledgeBase

Similar Behavior Discovery

Summary

- An in-depth analysis on state-of-the-art security operations and machine learning techniques, reveals the gap between them.

- Several strategies are proposed to solve the last mile problem.

- As showcases, we demonstrate how to implement these approaches in practice.

THANKS