Innovative R&D by NTT

Extracting & Exploring Threat Intel on Open Sourced Documents using Natural Language Processing

Mayo YAMASAKI
NTT-CERT, NTT Secure Platform Labs

Overview

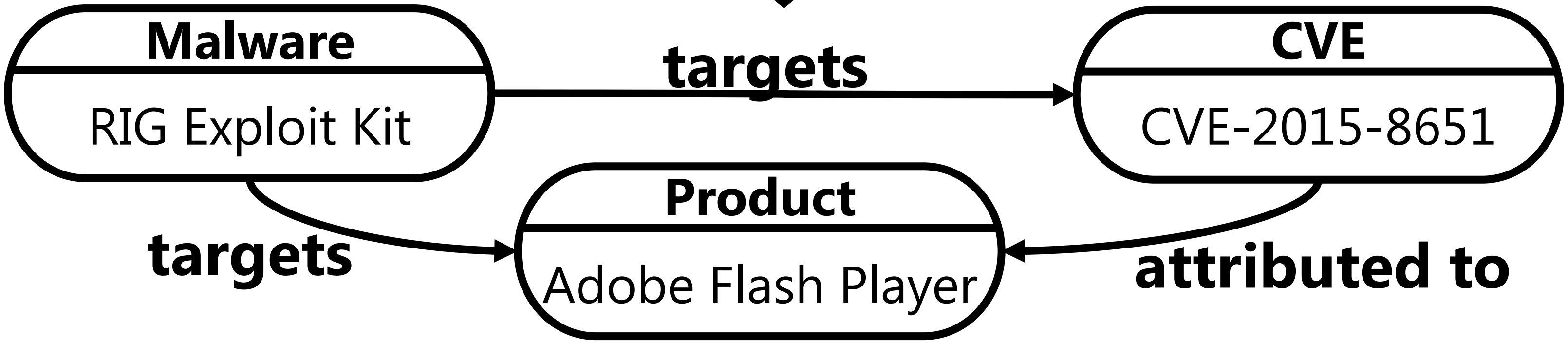
Developing a Threat Knowledge Extraction System by Using NLP.

Set of Unstructured Documents

RIG Exploit Kit targets Adobe Flash Player exploit (CVE-2015-8651).

⋮

Knowledge Graph



About NTT-CERT

- The CSIRT of NTT group
- Department of NTT Secure Platform Labs
- Activities
  - Incident Response
  - Product Evaluation
  - Forensics & Malware Analysis
  - Vulnerability Reporting
  - Training & Education
  - **OSINT**
  - Etc.

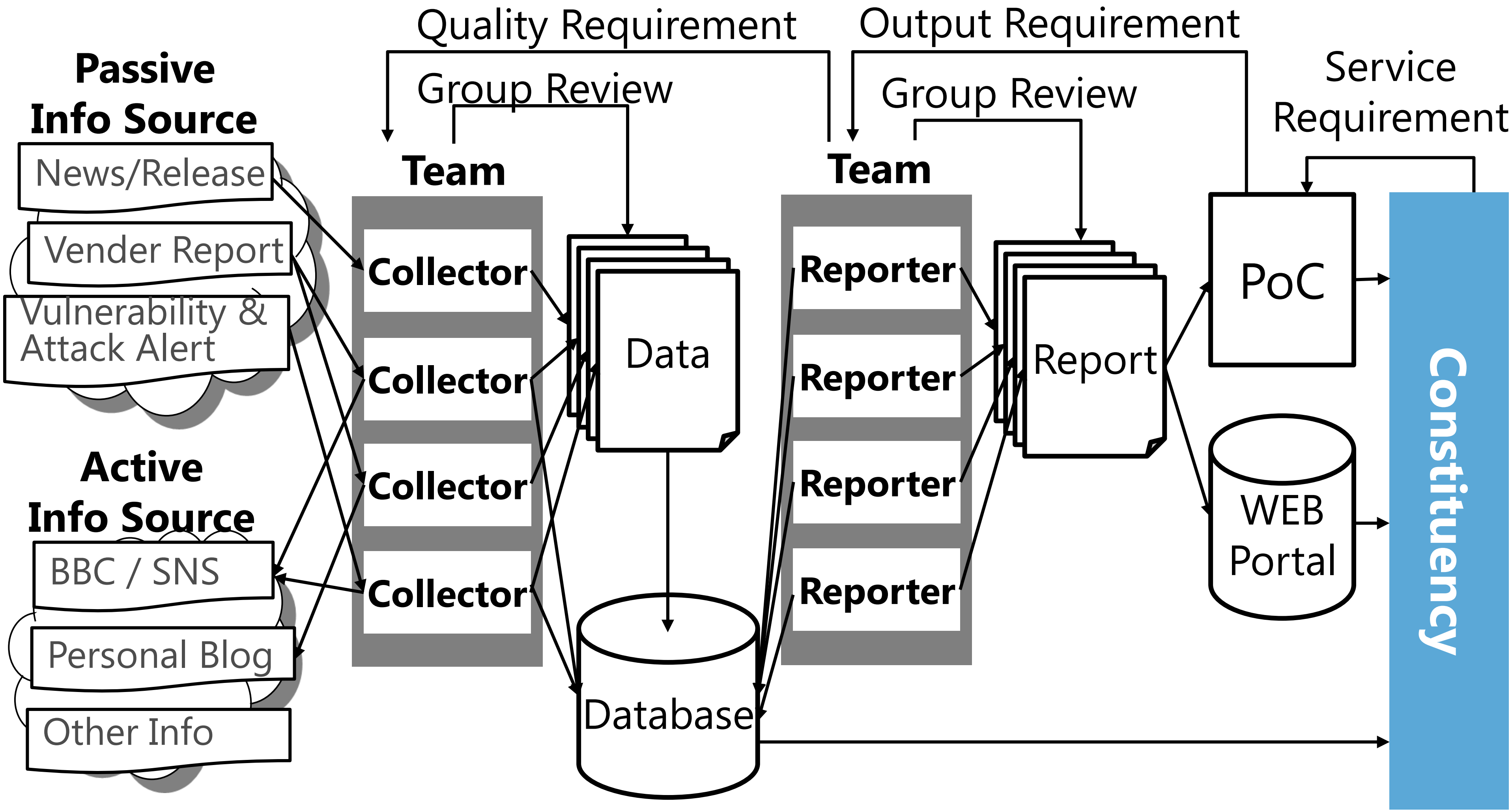
www.ntt-cert.org

What is OSINT ?

"Open-source intelligence (OSINT) is intelligence that is ***produced from publicly available information*** *and is collected, exploited, and disseminated in a timely manner to an* ***appropriate audience for the purpose of addressing a specific intelligence requirement****"*

- United States Department of Defense

Our OSINT Activities



Problem

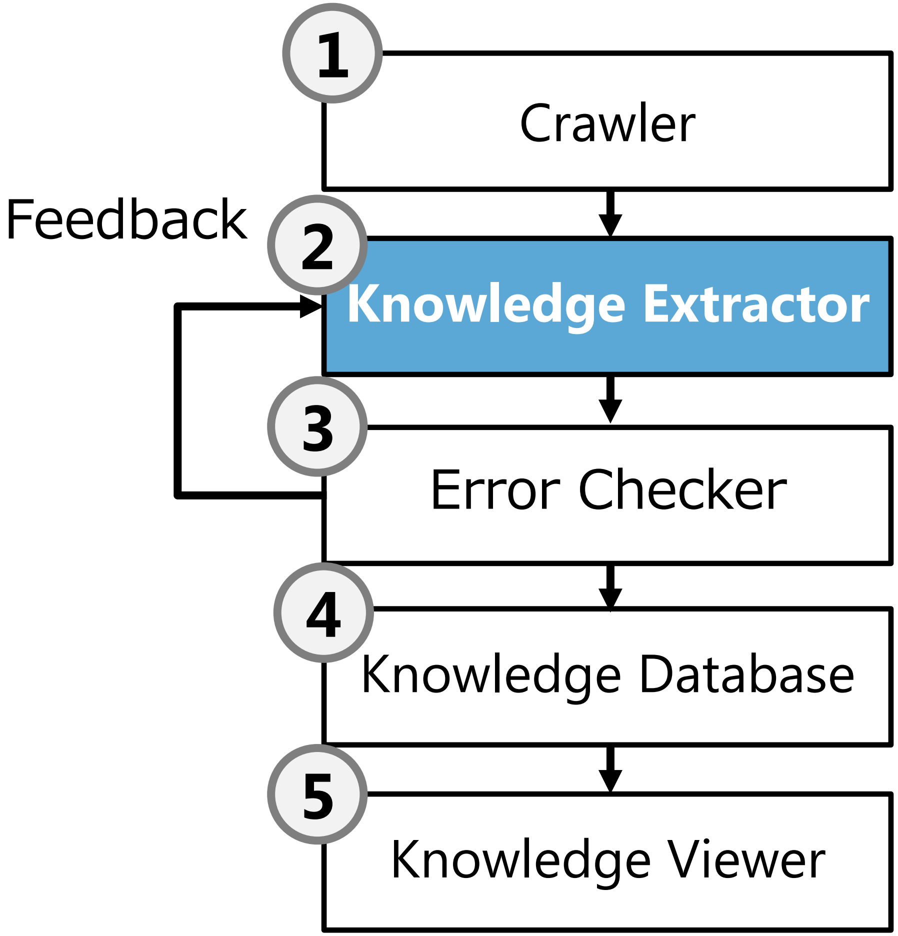
- Collecting & Storing **Unstructured Documents**

⬇

- Hard to Search and Understand Threat Intel
- Dependency on Knowledge of Team Members

Solution

- Automatic Threat Knowledge Extraction System



1 Crawler

2 Knowledge Extractor

3 Error Checker

4 Knowledge Database

5 Knowledge Viewer

Feedback

Collect pages & extract text

Extract structure using NLP

Check extraction error by human

Store knowledge in graph DB

Search & show knowledge

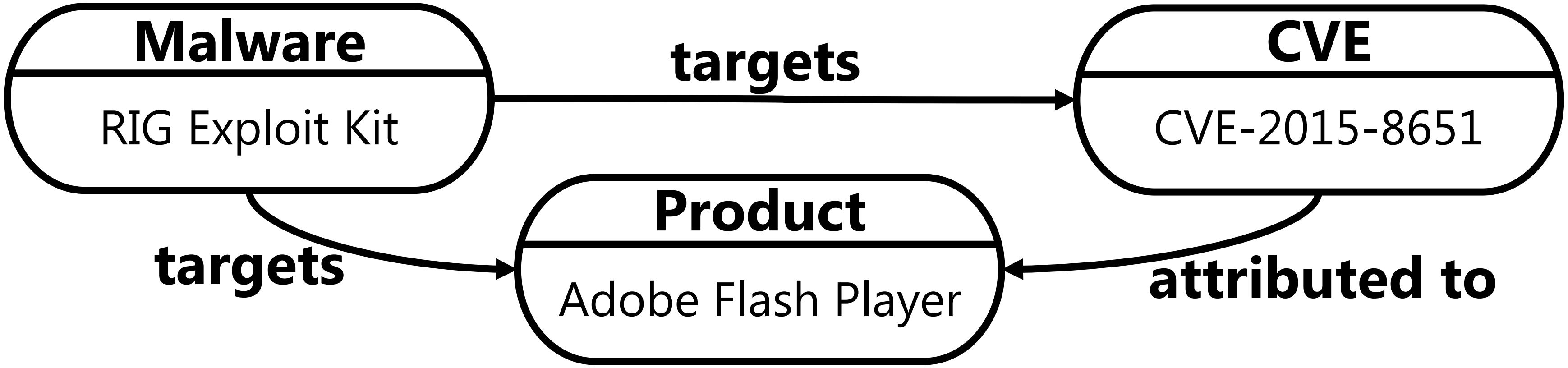
Knowledge Extraction in NLP

Set of Unstructured Documents

RIG Exploit Kit targets Adobe Flash Player exploit (CVE-2015-8651).

⋮

Knowledge Graph



Related Work

Knowledge Extraction Tasks

🟥 Semi Supervised
🟦 Supervised

Tasks of Non Security Domain

🟦 MUC-4 '92	Attack, Kidnapping, Bombing, Arson, etc.	Terrorism
🟦 CoNLL '03	Person, Organization, Location, MISC.	News
🟦 BioNLP '11	Gene Expression, Protein catabolism, etc.	Medical
🟦 ScienceIE '17	Process, Task, Material.	Resarch

Tasks of Security Domain

🟦 Joshi+ '13	Software, Hardware, Attack mean, etc.	Vulnerability
🟥 Jones+ '15	Software, Vender, CVE, Version, etc.	Vulnerability
🟥 Ramnani+ '17	Vulnerability, Thret Actor, IoC, TTP, etc.	Threat

Our Approach

Previous Threat Knowledge Extraction

- × Low accuracy
- × Hard to evaluate accuracy

⬇

Our Approach by using Supervised Learning

- ✓ High accuracy
- ✓ Easy to evaluate accuracy

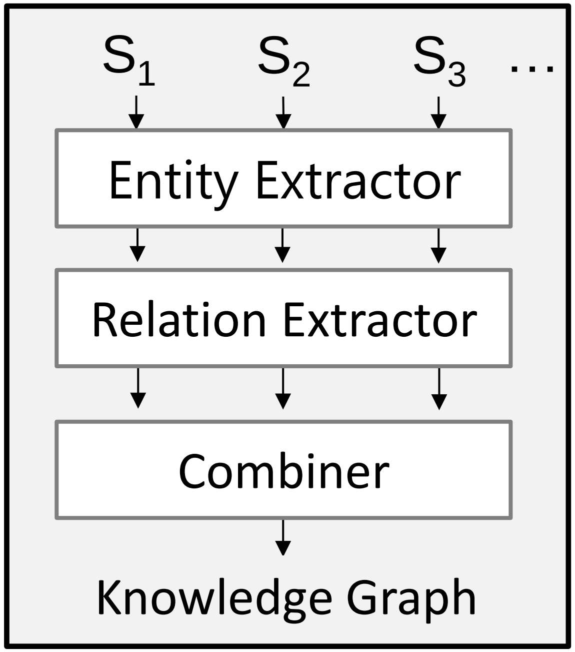
Our Approach

Our Task Overview

- Input
  - Set of Sentences
- Output
  - Knowledge Graph
- Three Sub Tasks
  - Entity Extraction
  - Relation Extraction
  - Combining Graphs

S_i is a sentence

Task Overview



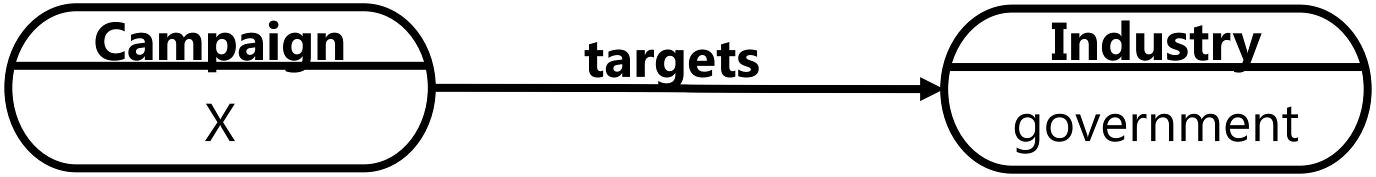
How to Define Threat Structure

- Using STIX 2.0 for Knowledge Extraction
  - Adaptation for Ambiguous Structure in Text
    - ✓ **Missing Data**(e.g. Unknown Identity)
    - ✓ **Required Binary Relation**(e.g. Unstructured Property)

Input Sentence:

X campaign **targets** a **governmental organization**

Desirable Output:



Entity Extraction

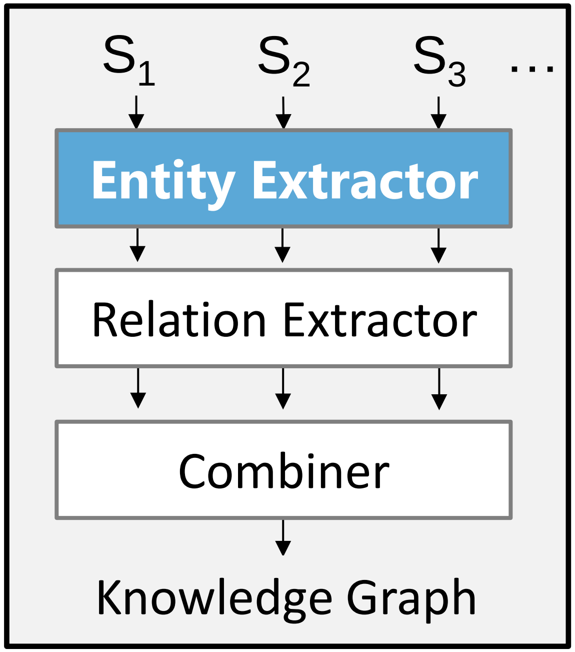
- Extracting Subsequences of Words as Entities

RIG	EK	targets	Adobe	Flash	Player	exploit	(	CVE-2015-8651	)	.
Malware	Malware	O	Product	Product	Product	O	O	Cve	O	O

- Multiclass Classification
- Entity Classes:

  { AttackPattern, Campaign, Cve, Domain, Hash, Identity, Industry, Ip, Malware, Product, Region, Role, ThreatActor, Time, Version, O }

Task Overview



Relation Extraction

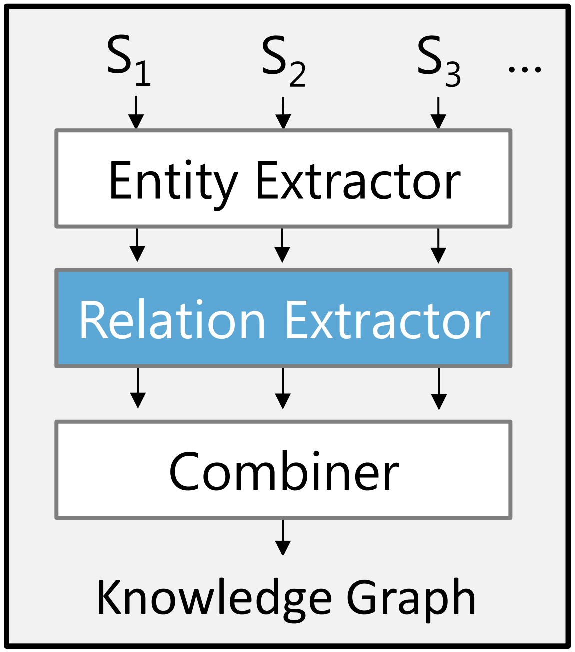
- Extracting Relation between Entities

entity1	entity2	relation
Rig Exploit Kit	Adobe Flash Player	**targets**
Rig Exploit Kit	CVE-2015-8651	**targets**
Adobe Flash Player	Rig Exploit Kit	O
Adobe Flash Player	CVE-2015-8651	O
CVE-2015-8651	Rig Exploit Kit	O
CVE-2015-8651	Adobe Flash Player	**attributed-to**

- Multiclass Classification
- Relation Classes:

  {attributed-to, aliases, indicate, observed-in, uses, targets }

Task Overview



Combining Graphs

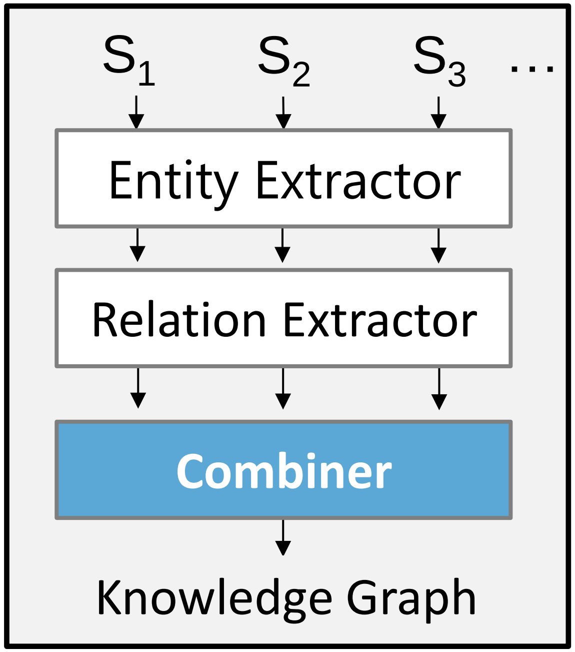
- Combining Graphs Extracted from each Sentences

ID	relation(arg1, arg2)
1	targets(Rig Exploit Kit, Adobe Flash Player)
2	targets(Rig Exploit Kit, CVE-2015-8651)
3	attributed-to(CVE-2015-8651, Adobe Flash Player)

arg1	arg2	Is combining?
ID1-arg1	ID2-arg1	**YES**
ID1-arg1	ID2-arg2	NO
…	…	…

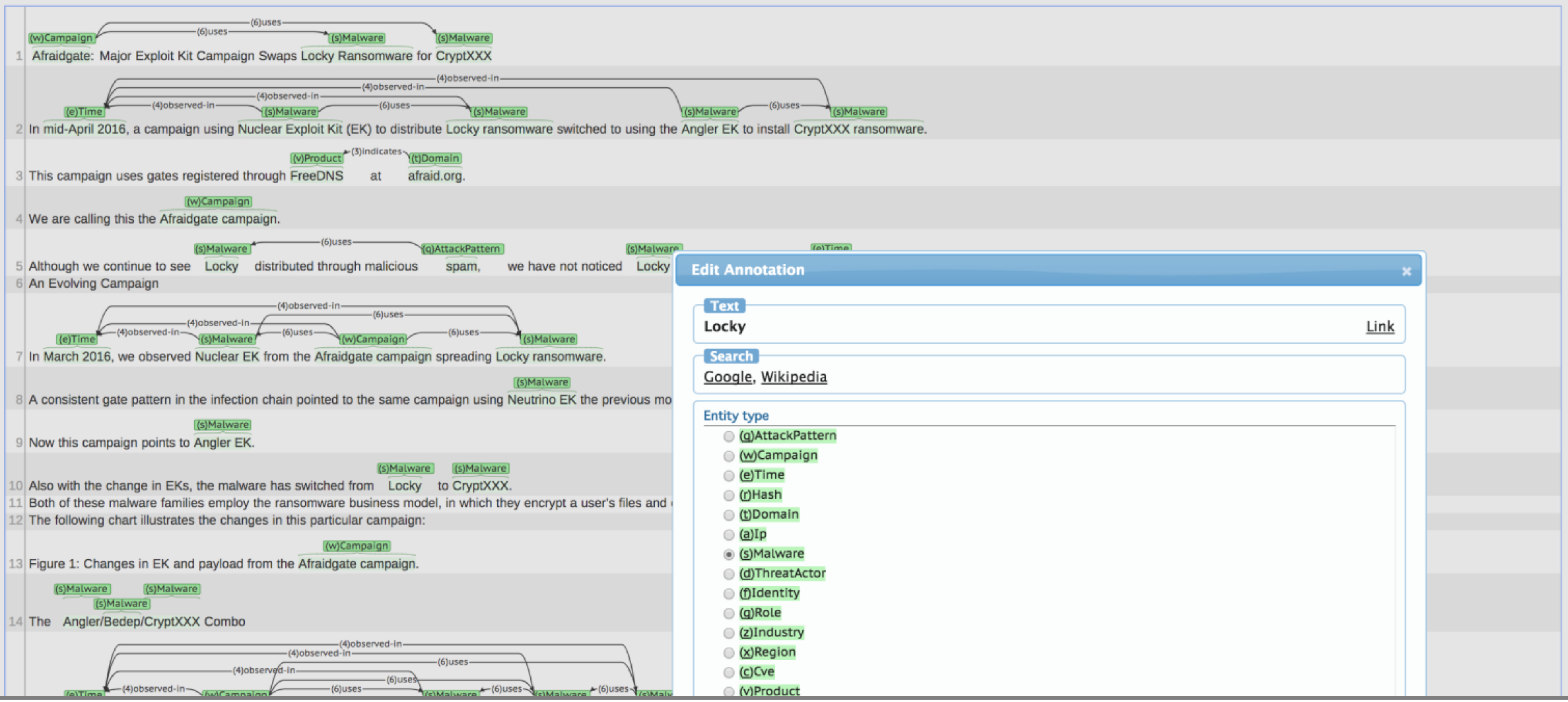
- Binary Classification
- Classes: Same Entity or Not

Task Overview



Developing Labeled Data

- Labeling 200 WEB documents(about 10,000 sentences)
- Labeling documents by 5 peoples using a tool



1 Afraidgate: Major Exploit Kit Campaign Swaps Locky Ransomware for CryptXXX

2 In mid-April 2016, a campaign using Nuclear Exploit Kit (EK) to distribute Locky ransomware switched to using the Angler EK to install CryptXXX ransomware.

3 This campaign uses gates registered through FreeDNS at afraid.org.

4 We are calling this the Afraidgate campaign.

5 Although we continue to see Locky distributed through malicious spam, we have not noticed Locky

6 An Evolving Campaign

7 In March 2016, we observed Nuclear EK from the Afraidgate campaign spreading Locky ransomware.

8 A consistent gate pattern in the infection chain pointed to the same campaign using Neutrino EK the previous mo

9 Now this campaign points to Angler EK.

10 Also with the change in EKs, the malware has switched from Locky to CryptXXX.

11 Both of these malware families employ the ransomware business model, in which they encrypt a user's files and

12 The following chart illustrates the changes in this particular campaign:

13 Figure 1: Changes in EK and payload from the Afraidgate campaign.

14 The Angler/Bedep/CryptXXX Combo

Edit Annotation

Text: Locky

Search: Google, Wikipedia

Entity type
- ○ (q)AttackPattern
- ○ (w)Campaign
- ○ (e)Time
- ○ (r)Hash
- ○ (t)Domain
- ○ (a)Ip
- ● (s)Malware
- ○ (d)ThreatActor
- ○ (f)Identity
- ○ (g)Role
- ○ (z)Industry
- ○ (x)Region
- ○ (c)Cve
- ○ (v)Product

Policy for Labeling Documents

- Creating a Guideline Document with Case Studies
  - E.g.1 Masked domains are labeled as domain.

    ```
    /reallstatistics[.]info/Domain
    ```

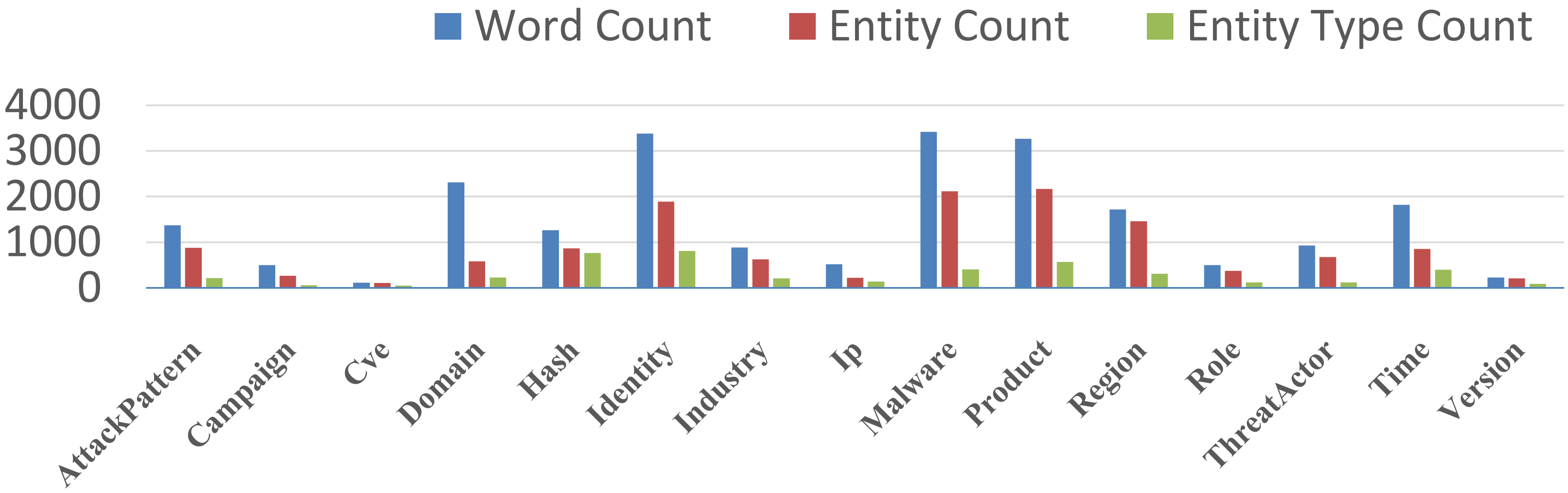
  - E.g.2 Malware types aren't attack pattern.

    ```
    /Key logging/AttackPattern
    /Keylogger/O
    ```

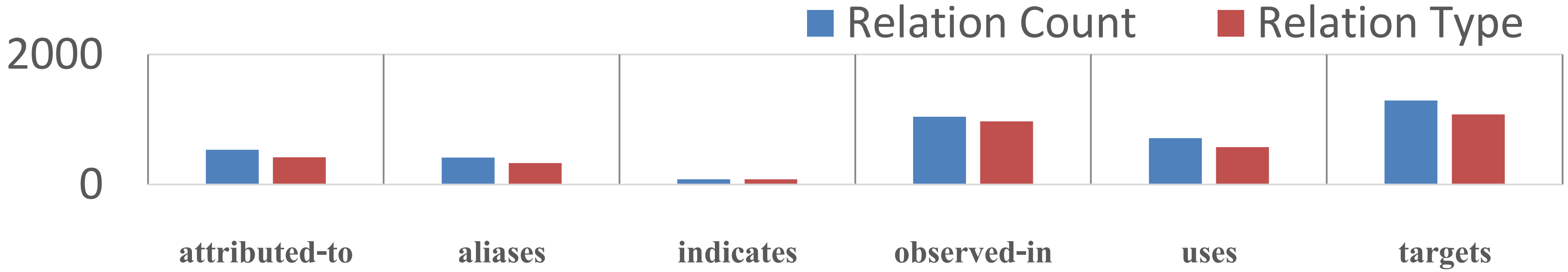
- Force Restriction
  - E.g. Relations are defined only between specific entities by "brat" annotation tool.
- Checking All Labeled Data by Supervisor
- Hiring Cyber Security Domain Experts

Stats of Labeled Data

Labeled Data for Entity Extraction



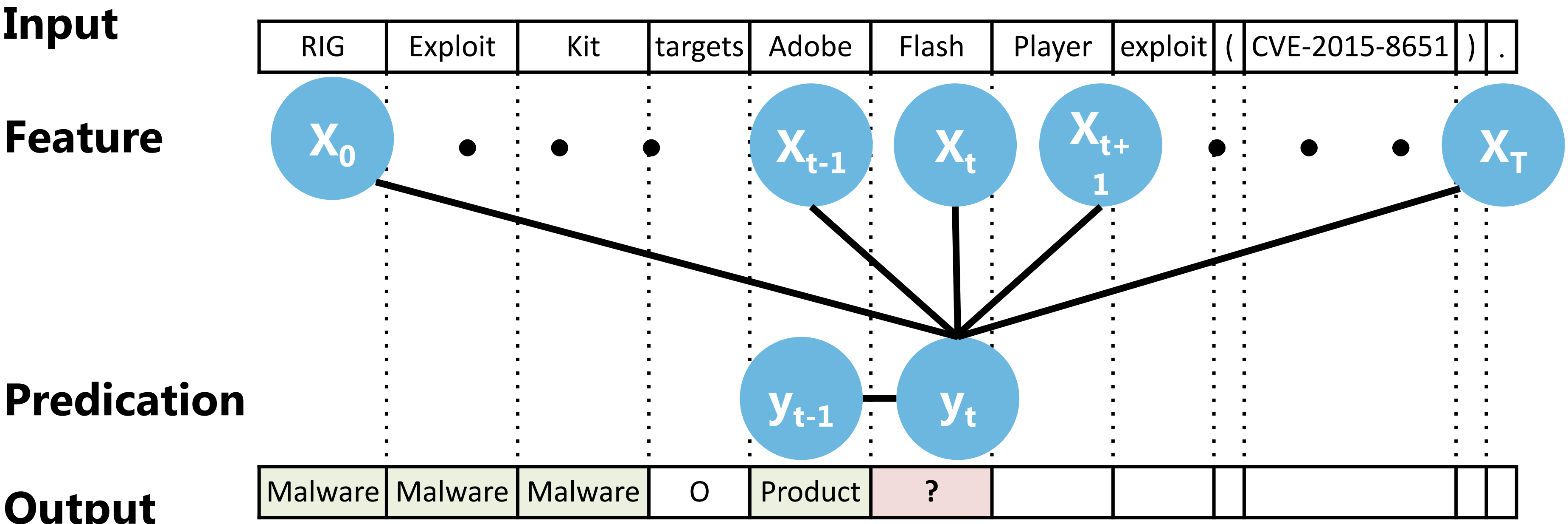
Labeled Data for Relation Extraction



Experiments & Results

Experiment of Entity Extraction

■ Extraction by CRF(Conditional Random Field)

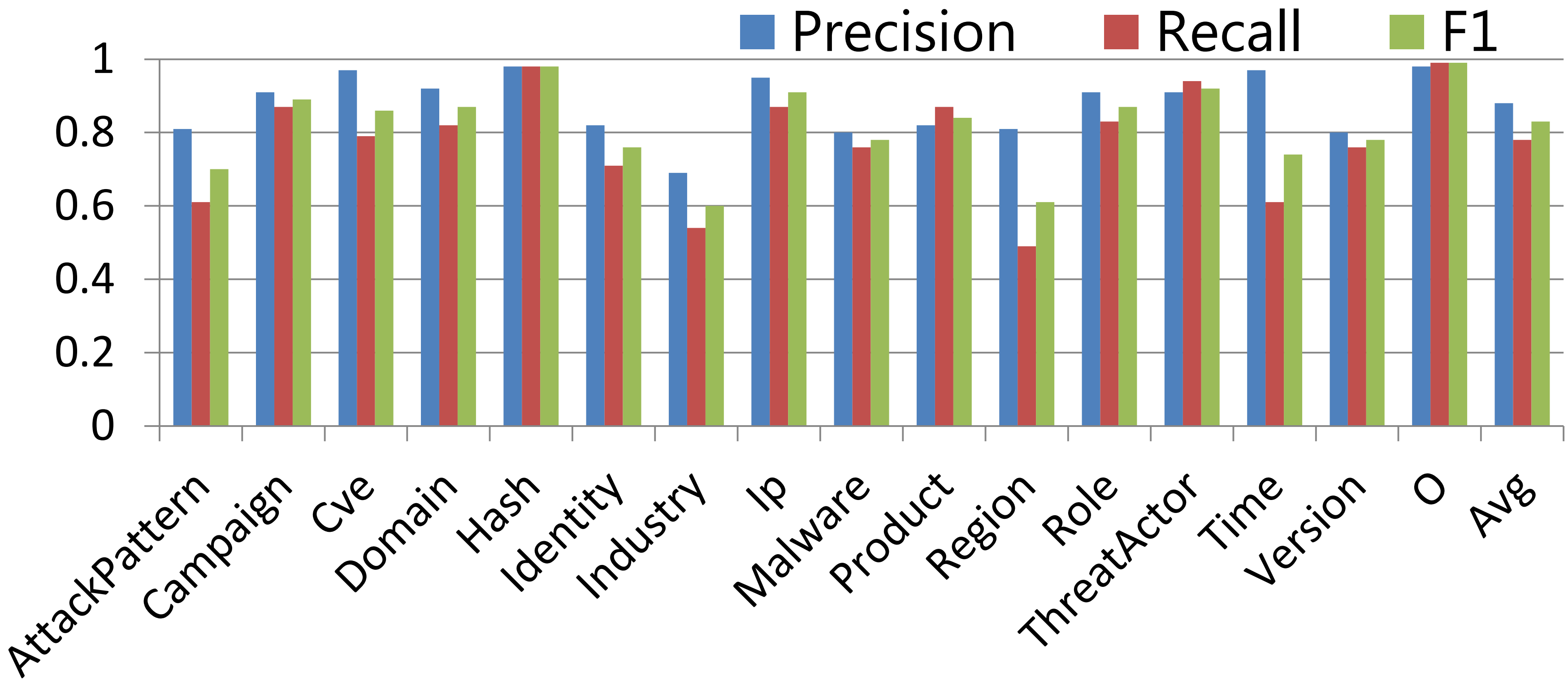


Input	RIG	Exploit	Kit	targets	Adobe	Flash	Player	exploit	(	CVE-2015-8651	)	.
Output	Malware	Malware	Malware	O	Product	?						

- Features(Ratinov+ 2009)：Form, POS, Entity Labels for News, BoW Character N-gram, Brown Cluster, Wikipedia & demonyum Lexicon,.
- Hyper Parameters： Decision by Random Search

Result of Entity Extraction

- Average of 3 F Scores of Predicting Labels for each Words
  - Training and Developing Model by 80% of Dataset
  - Testing Model by Rest 20% of Dataset



Experiments of Relation Extraction

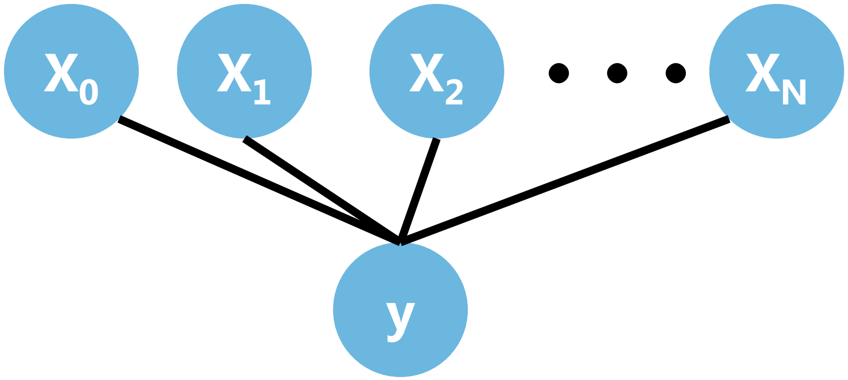
- Extraction by Linear SVM(Support Vector Machine)

Sentence

RIG	Exploit	Kit	targets	Adobe	FLash	Player	exploit	(	CVE-2015-8651	)	.

Input Pair (Rig Exploit Kit, Adobe Flash Player)

Feature



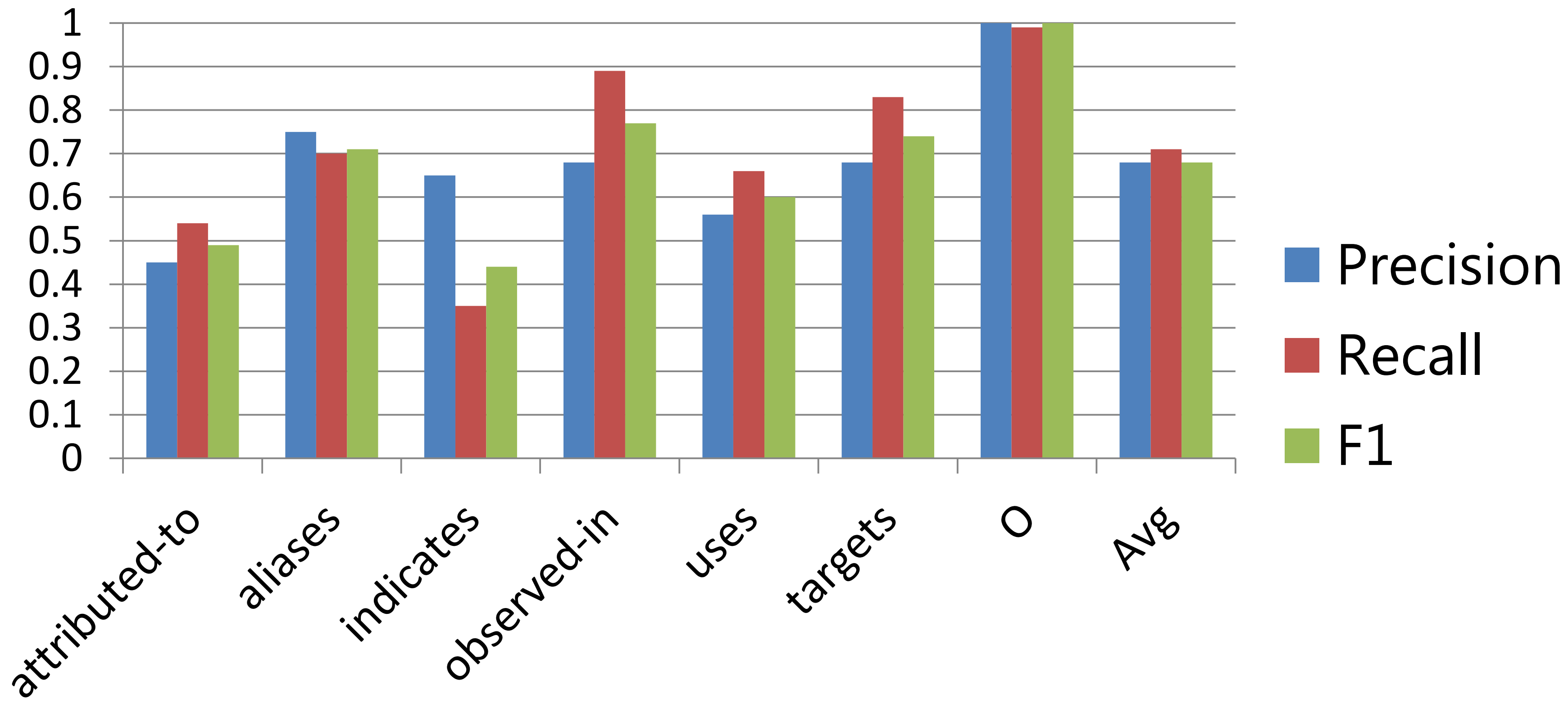
Predication

Output **targets**

- Features(Rink + 2010)：Our Entity Labels, Form, POS, Entity Labels for News, Hypernym on WordNet, Distance, Dependency Tree.
- Hyper Parameters ： Decision by Grid Search.

Results of Relation Extraction

- Average of 3 F Scores of Predicting Relation
  - Training and Developing Model by 80% of Dataset
  - Testing Model by Rest 20% of Dataset



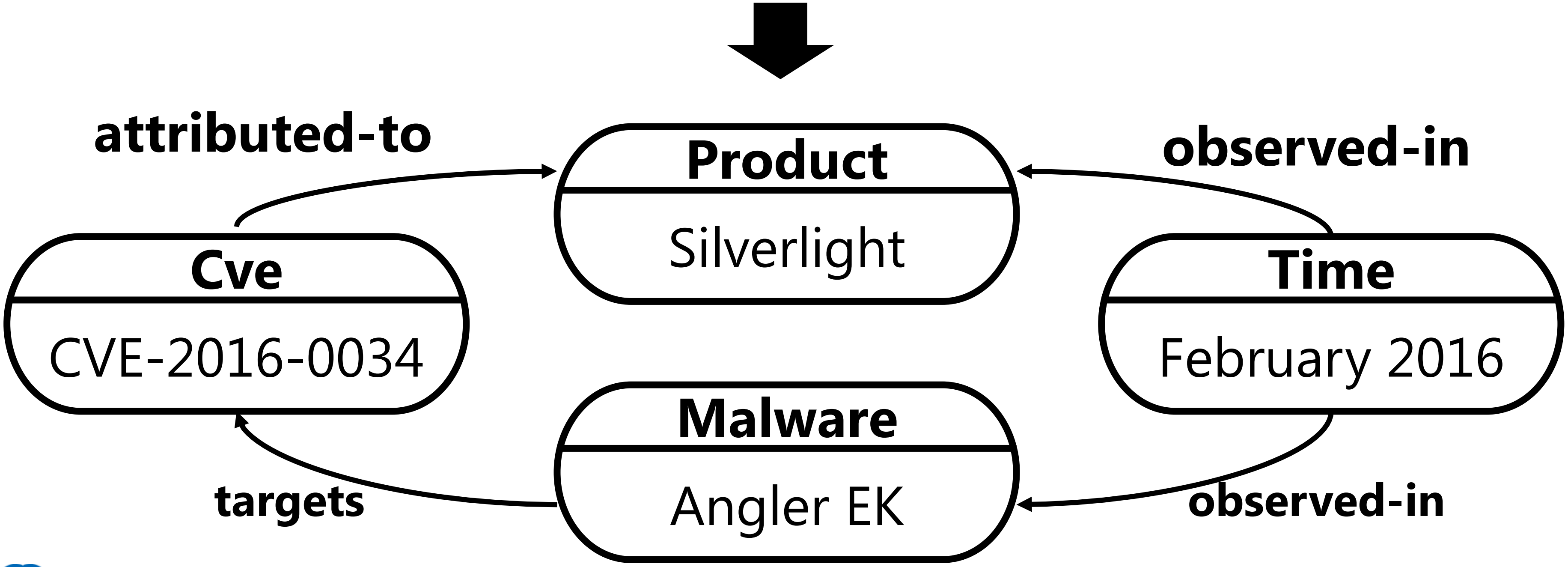
Experimental Result of Combining

- Average of 3 F Scores of Extracting Entity & Relation
  - Combining Results with Naive Rule
    - ✓ Rule: If words and labels of two entities are same, we define these entities are same.
  - Training and Developing Model by 80% of Dataset
  - Testing Model by Rest 20% of Dataset

	Precision	Recall	F1
Entity Extraction	0.85	0.74	**0.79**
Relation Extraction	0.66	0.76	**0.71**

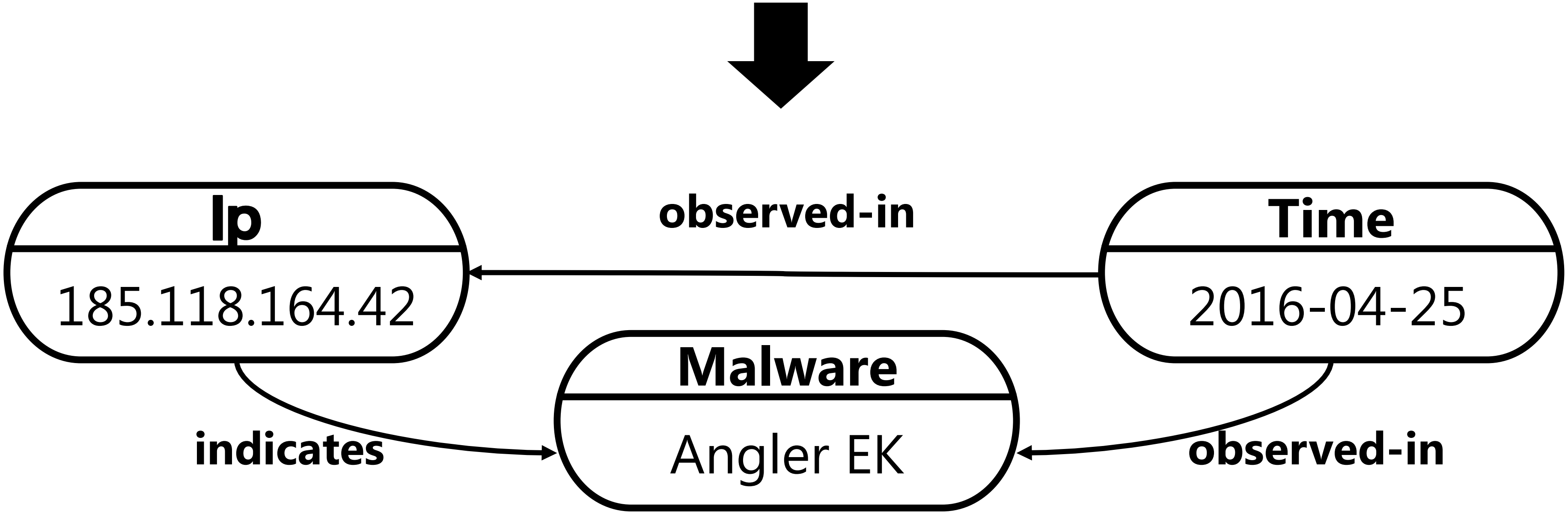
Example of Extraction Result 1

In **February 2016** , exploits for **Silverlight** based on **CVE-2016-0034** found their way into **Angler EK** a little more than a month after Microsoft issued a patch for the vulnerability .



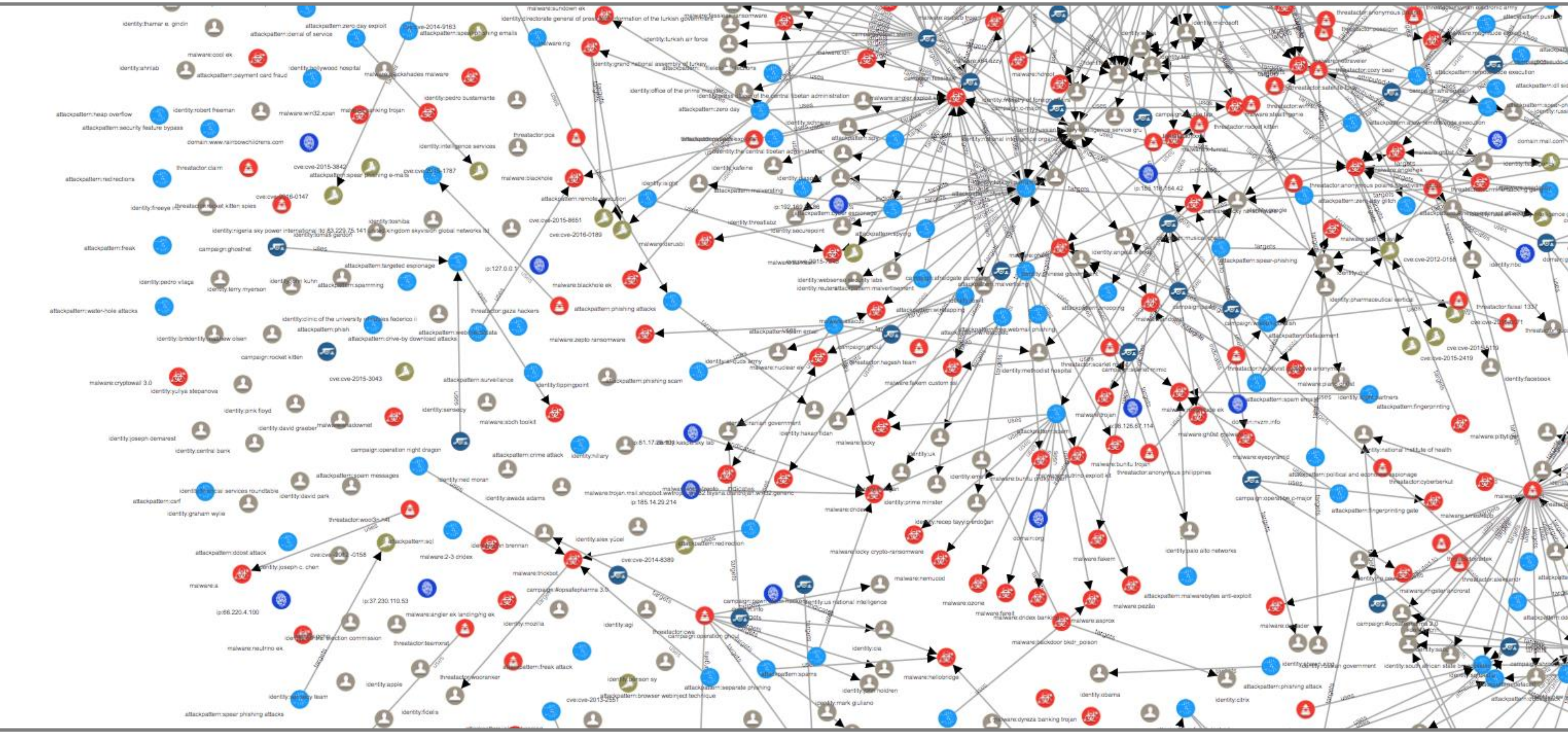
Example of Extraction Result 2

Similar gate on **185.118.164.42** leads to more **Angler EK** traffic on **2016-04-25** .



Exploring Threat Intel on the WEB

- Collect About 25,000 WEB Documents by Crawling
- Extract and Convert STIX Data(995 SDOs & 684 SROs)



Conclusion

Conclusion

- Summary
  - Developing **Threat Knowledge Extraction System** using **Supervised Learning** and **Labeled Dataset**
  - Developing **Entity Extractor** of about **80% F Score**
  - Developing **Relation Extractor** of about **70 % F Score**

- Future Work
  - Examination of Baseline Score for Production
  - Analysis of Massive Knowledge Graph