Looking Beyond IoCs: Automatically Extracting Attack Patterns from External CTI

Md Tanvirul Alam
Rochester Institute of Technology
Rochester, New York, USA
tanvirul.alam@mail.rit.edu

Dipkamal Bhusal
Rochester Institute of Technology
Rochester, New York, USA
db1702@rit.edu

Youngja Park
IBM Research
Yorktown Heights, New York,
USA young_park@us.ibm.com

Nidhi Rastogi
Rochester Institute of Technology
Rochester, New York, USA
nxrvse@rit.edu

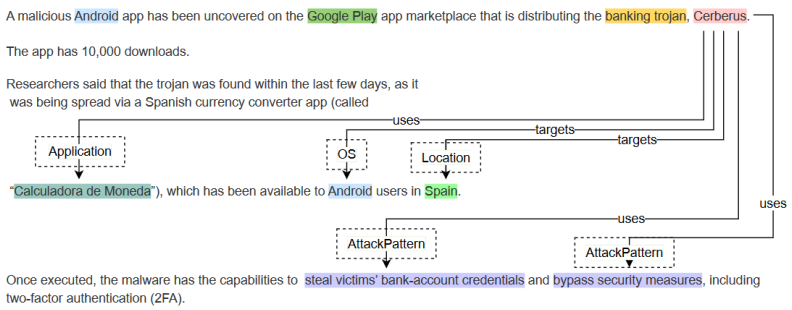
International Symposium on Recent Advances in Intrusion Detection 2023

Outline

- Problem Statement
  - CTIs
  - Attack Pattern
- Overview of LADDER
- System Design
  - Datasets
  - Entity Extraction
  - TTPClassifier
  - Relation Extraction
  - KG
- Experiment
- Case Study
- Conclusion
- Extension

Problem Statement

- Given
  - CTIs
  - MITRE ATT&CK IDs

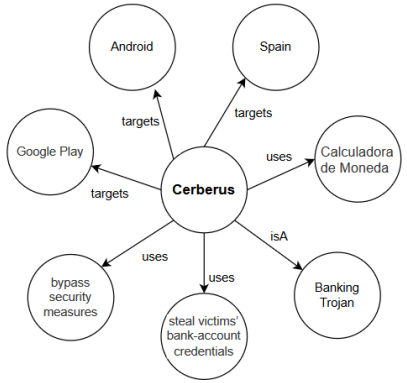


- Information Extracting
  - Extract Attack Patterns from CTIs
  - Extract Entities and Relationships from CTIs

Steal victim's bank account credentials (class:Attack Pattern)

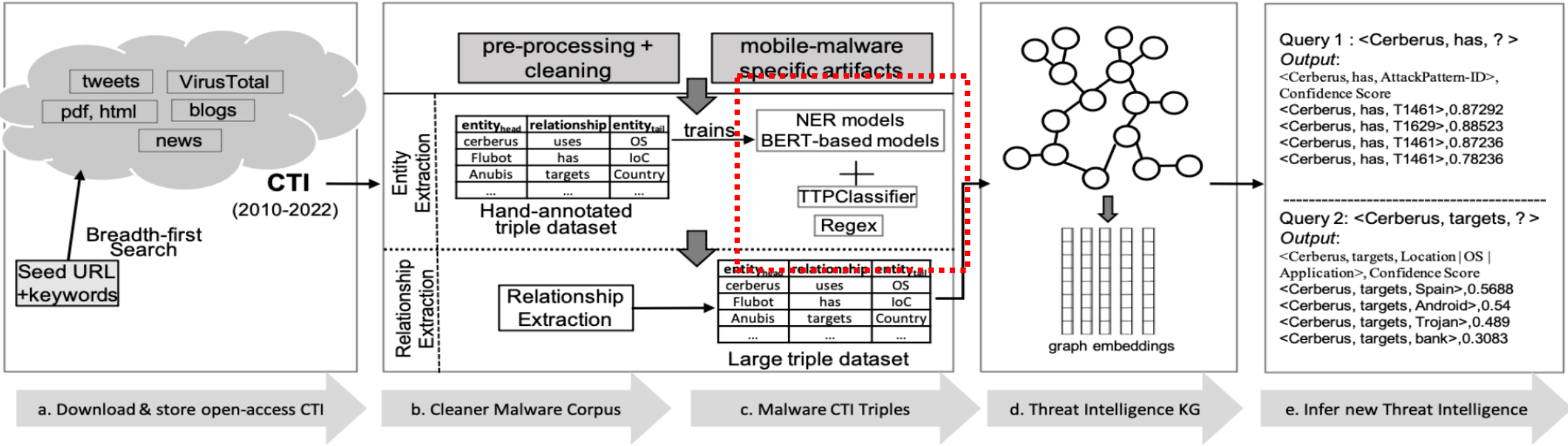
Bypass security measures (class:Attack Pattern)

- Mapping and Generating KG
  - Map Attack Patterns into MITRE ATT&CK IDs
  - Generate a KG for link prediction



Overview of LADDER

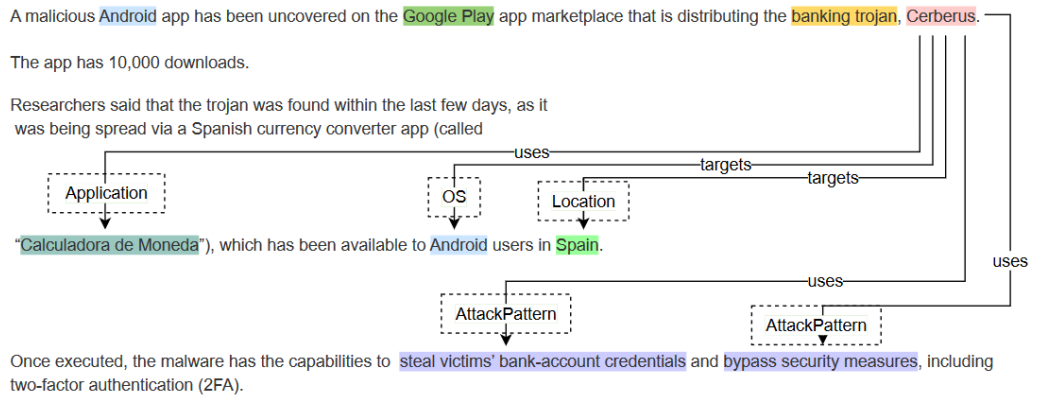
- A framework to extract text-based Attack Patterns
  - Extracts CTI using crawlers
  - Pre-processes and prepares for entity and relationship extraction
  - Generates data in the form of triples
  - Creates a knowledge graph for link prediction



System Design

- Datasets
  - Employed *BART* to manually annotate threat concepts and relationships
  - Scraped over 12,000 relevant unstructured open-access CTI reports

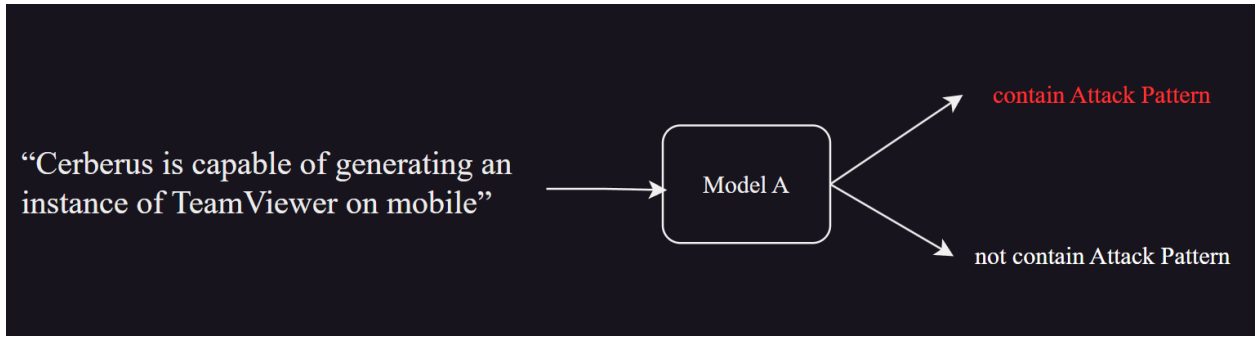
- Entity Extraction
  - Classes including Malware, Application …
  - Transformer-based models
  - Pattern matching to extract URL , IP …

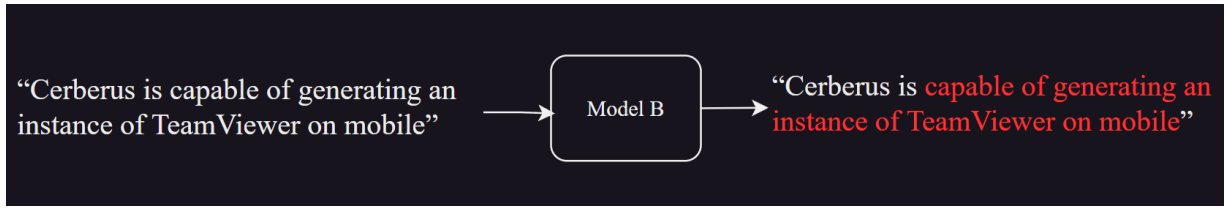


Entity Types	Regular Expression
FilePath	r'[a-zA-Z]:\\([0-9a-zA-Z]+)', r'(\/[^\s\n]+)+'
Email	r'[a-z][_a-z0-9-.]+@[a-z0-9-]+[a-z]+'
SHA256	r'[a-f0-9]{64}|[A-F0-9]{64}'
SHA1	r'[a-f0-9]{40}|[A-F0-9]{40}'
CVE	r'CVE—[0-9]{4}—[0-9]{4,6}'
IPv4	r'^((25[0-5]|(2[0-4]|1\d|[1-9]|)\d)(\.(?!)|)){4}$'

System Design

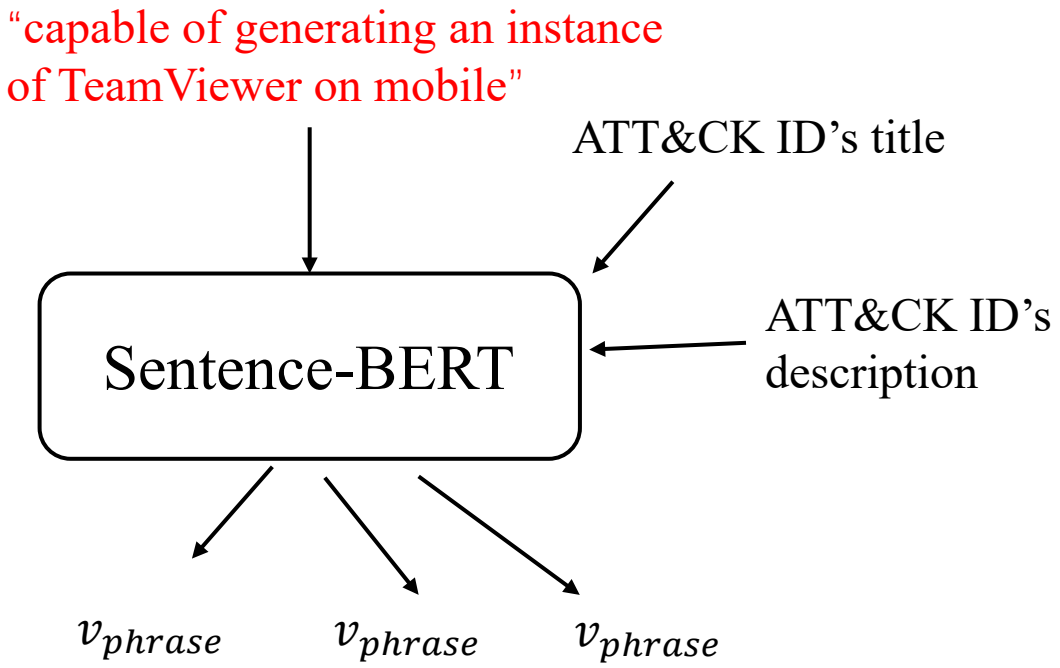
- Attack Pattern Extraction
  - Challenges: a larger block of text
    - Maybe include other entity types
    - “Cerberus is capable of generating an instance of TeamViewer on mobile”
  - TTP Classifier
    - 1 relevant sentence extraction
    - 2 attack phrase identification & extraction
    - 3 mapping attack patterns to the MITRE ATT&CK





System Design

- TTP Classifier
  - 3 mapping attack patterns to the MITRE ATT&CK



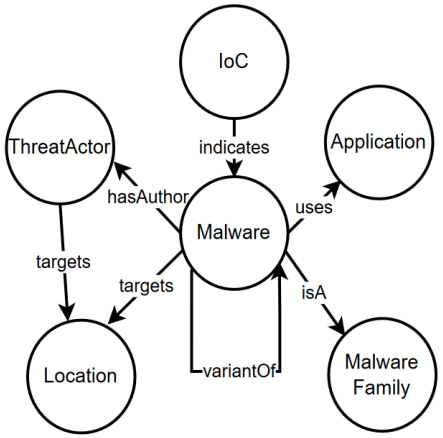
$$d_i = w_t \cos(v_{phrase}, v^i_{title}) + (1 - w_t)\cos(v_{phrase}, v^i_{desc})$$

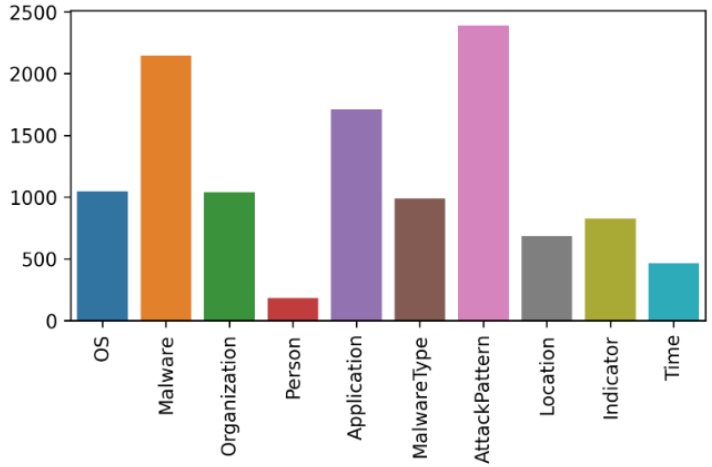
$$\cos(u, v) = 1 - \frac{u \cdot v}{||u||_2||v||_2}$$

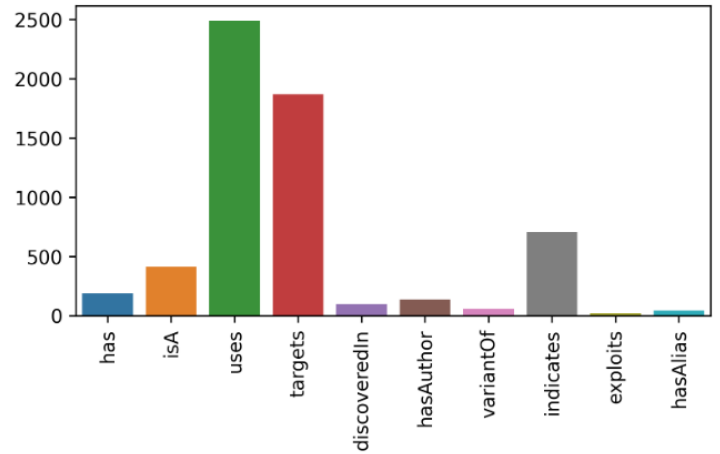
- Enumerates 66 attack patterns for mobile platforms
  - $d_i < \tau$

System Design

- Relation Extraction
  - Got All kinds of Entities include Attack patterns
  - *An Ontology-driven Knowledge Graph for Android Malware*
    - *has uses targets ....*
- Train a relation classification model
  - Input: [CLS] ⟨e1⟩ Cerberus ⟨/e1⟩ is capable of generating an instance of ⟨e2⟩ TeamViewer ⟨/e2⟩
  - Output: uses

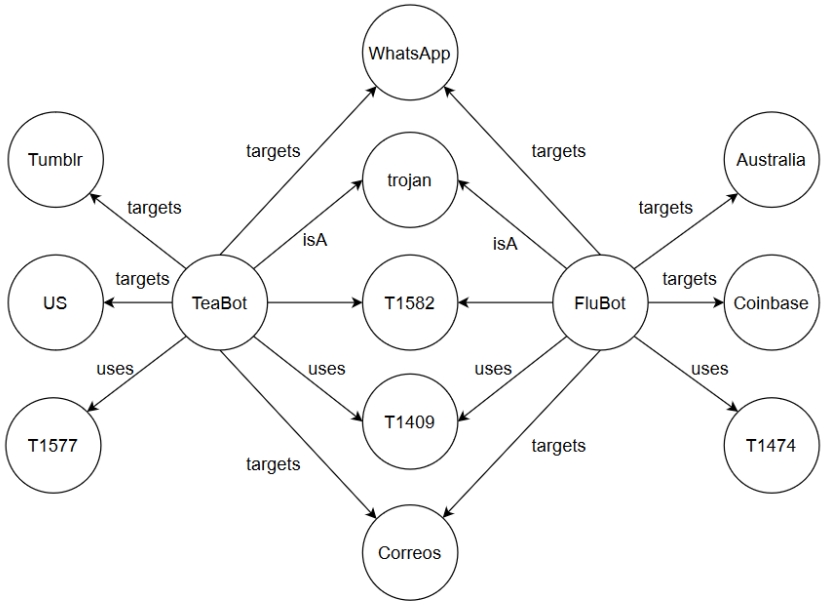






System Design

- Knowledge Graph
  - KG = {E, R, T}
  - E, R and T indicate the sets of entities, relations, and triples
  - *<ehead,r,etail> ∈ T* indicates that there is a relationship *r ∈ R*



Entity prediction for KG follows *TuckER [5]*

<ehead,r, ? > → <ehead,r, etail_by_prediction >

Query 1 : <Cerberus, has, ? >
Output:
<Cerberus, has, AttackPattern-ID>, Confidence Score
<Cerberus, has, T1461>,0.87292
<Cerberus, has, T1629>,0.88523
<Cerberus, has, T1461>,0.87236
<Cerberus, has, T1461>,0.78236

Query 2: <Cerberus, targets, ? >
Output:
<Cerberus, targets, Location | OS | Application>, Confidence Score
<Cerberus, targets, Spain>,0.5688
<Cerberus, targets, Android>,0.54
<Cerberus, targets, Trojan>,0.489
<Cerberus, targets, bank>,0.3083

Experiments

- Entity Extraction

Model	Precision	Recall	F1-score
BERT-base	73.34	77.88	75.14
BERT-large	75.30	79.23	77.12
RoBERTa-base	41.55	41.01	40.84
RoBERTa-large	35.95	36.23	35.49
XLM-RoBERTa-base	75.32	79.06	76.98
XLM-RoBERTa-large	**76.97**	**81.57**	**78.98**

Class	Precision	Recall	F1-score
Malware	78.45	83.08	80.70
MalwareType	65.64	87.18	74.89
Application	70.13	73.26	71.66
OS	89.95	96.24	92.99
Organization	73.68	74.12	73.90
Person	88.24	75.00	81.08
ThreatActor	58.33	37.84	45.90
Time	85.51	89.39	87.41
Location	93.55	89.92	91.70
Average	76.97	81.57	78.98

- TTP Classifier

Model	Precision	Recall	F1-score
BERT-base	86.50	85.20	85.84
BERT-large	86.06	85.80	85.93
RoBERTa-base	87.42	86.10	86.76
RoBERT-large	**89.22**	**90.03**	**89.62**
XLM-RoBERTa-base	83.00	88.52	85.67
XLM-RoBERTa-large	84.73	88.82	86.73

Model	Precision	Recall	F1-score
BERT-base	87.67	90.55	89.09
BERT-large	87.74	87.81	87.78
RoBERTa-base	88.53	90.12	89.32
RoBERT-large	**89.19**	**92.14**	**90.64**
XLM-RoBERTa-base	86.82	90.72	88.73
XLM-RoBERTa-large	88.55	91.77	90.13

- Relationship Extraction

Model	Precision	Recall	F1-score
BERT-base	93.75	92.22	92.60
BERT-large	**93.78**	**92.46**	**92.62**
RoBERTa-base	81.66	83.88	82.27
RoBERT-large	77.47	81.06	77.97
XLM-RoBERTa-base	81.77	83.86	81.74
XLM-RoBERTa-large	72.64	78.69	75.29

Class	Precision
noRelation	100.0
isA	98.6
targets	96.3
uses	46.2
hasAuthor	87.5
has	100.0
variantOf	100.0
hasAlias	26.3
indicates	75.9
discoveredIn	100.0
exploits	100.0

Experiments

- KG1: from less CTIs
- KG2: from more CTIs

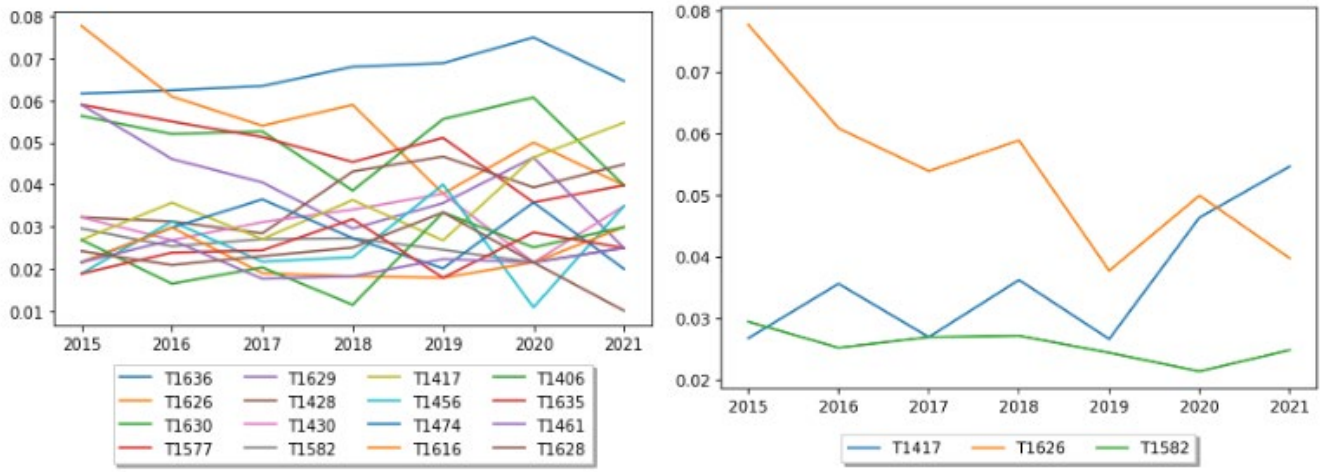
KG	$TestSet_1$				$TestSet_2$			
	Hits@3	Hits@10	Hits@30	MRR	Hits@3	Hits@10	Hits@30	MRR
KG_1	0.209	0.365	0.497	0.186	0.090	0.195	0.322	0.093
KG_2	0.221	0.353	0.516	0.211	0.215	0.359	0.501	0.203

– Knowledge GraphComparison with state-of-the-art for TTP Classifier
  – Match extracted phrase against MITRE ATT&CK for enterprise 4
  – Create ground truth annotation from threat reports listed on the MITRE ATT&CK website for five different malware containing 9360 tokens.

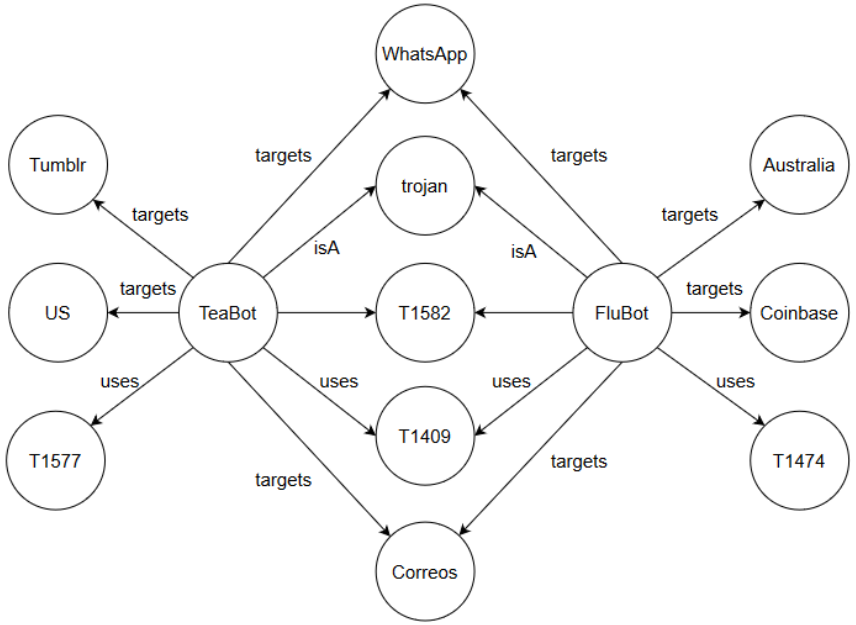
Method	TP	FN	FP	Precision	Recall	F1-score
MITRE	38	27	0	**1.00**	0.58	**0.74**
TTPDrill[19]	22	43	231	0.09	0.34	0.14
AttackKG[29]	12	53	85	0.12	0.18	0.15
TTPClassifier	41	24	22	0.65	**0.63**	0.64

Case-studies

- Trend Analysis



- Identifying Similar Malware and APT Groups

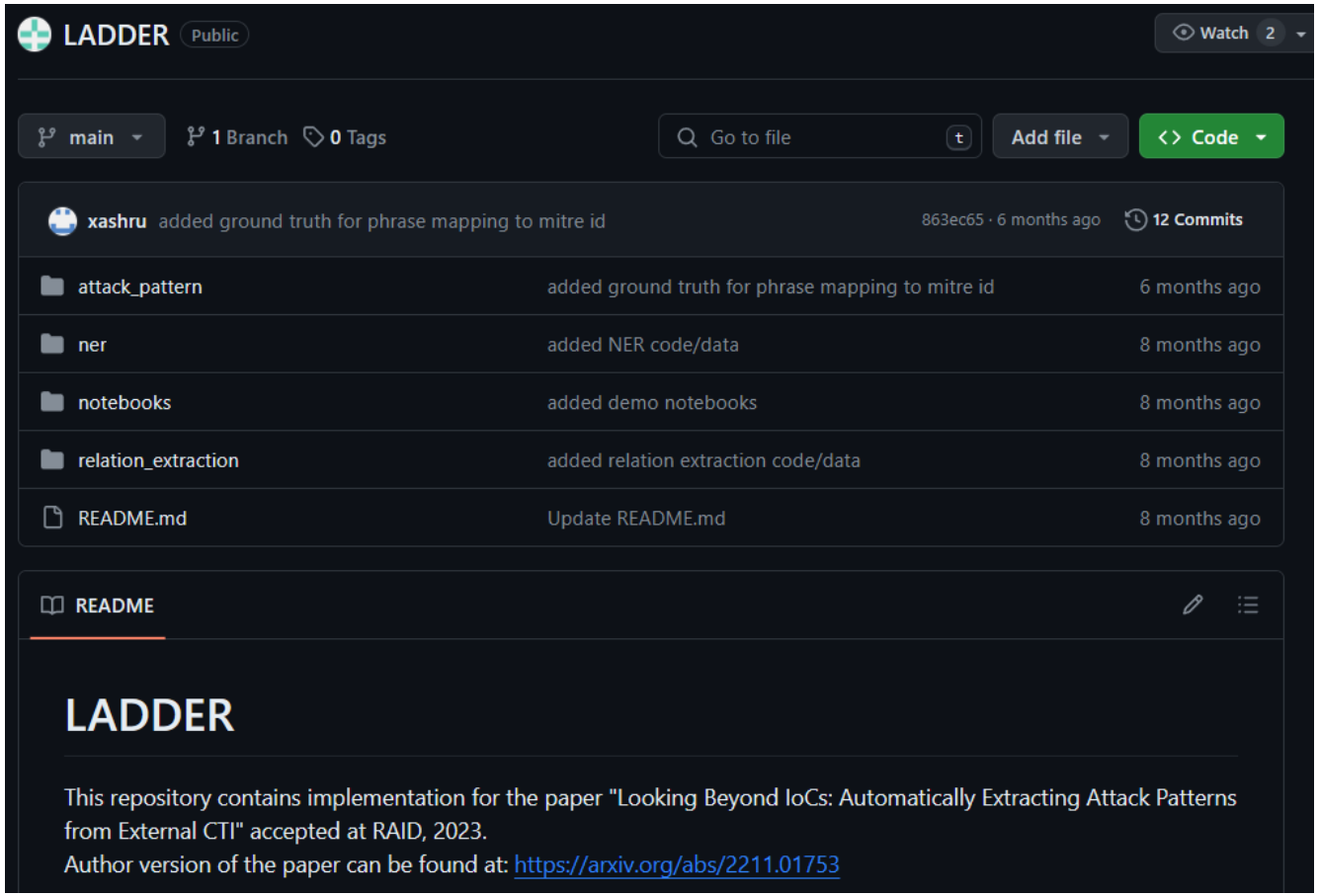


Related Work

Areas	Paper	Time	From
TTP Extraction	Paragraph-based Estimation of Cyber Kill Chain Phase from Threat Intelligence Reports	2020	Journal of Information Processing 28 (2020)
	TTPDrill: Automatic and Accurate Extraction of Threat Actions from Unstructured Text of CTI Sources	2017	ACSAC
	AttacKG: Constructing Technique Knowledge Graph from Cyber Threat Intelligence Reports	2022	ESORICS
NER	A comparative study of deep learning based named entity recognition algorithms for cybersecurity	2020	IEEE International Conference on Big Data
Relation Extraction	RelExt: Relation Extraction using Deep Learning approaches for Cybersecurity Knowledge Graph Improvement	2019	ASONAM
	Open-CyKG: An Open Cyber Threat Intelligence Knowledge Graph	2021	Knowledge Based System 233
Threat KG	Creating Cybersecurity Knowledge Graphs From Malware After Action Reports	2020	IEEE Access 8

Conclusion

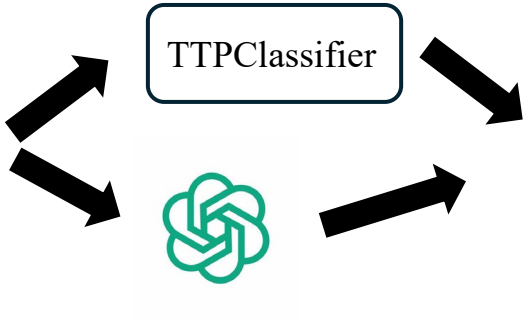
- An Entity Extraction Model
- A TTP Classifier for Attack Pattern
  - A sentence classification model
  - A token classification model
  - A mapping algorithm
- A Relation Extraction Model
- A Link Prediction Model
- A Manually Annotations Dataset
- A Web Crawler



Extension

- TTPClassifier → LLM/GPT

Cerberus is capable of generating an instance of TeamViewer on mobile



capable of generating an instance of TeamViewer on mobile