A Decade of Mal-Activity Reporting: A Retrospective Analysis of Internet Malicious Activity Blacklists

Benjamin Zhao, **Muhammad Ikram,** Hassan Asghar, Mohamed Ali (Dali) Kaafar, Abdelberi Chaabane, and Kanchana Thilakarathna

MACQUARIE University

DATA 61 CSIRO

UNSW SYDNEY

MACQUARIE University SYDNEY·AUSTRALIA

UNIVERSITY OF MICHIGAN

THE UNIVERSITY OF SYDNEY

Agenda

MACQUARIE University

1. **Why we need enriched cybersecurity dataset?**
2. Large Scale Cybersecurity Data Collection and Enrichment Process
3. Insights
    1. Characterization
    2. Temporal Analysis
4. Way forward: How can we leverage this dataset to improve detection systems

Why we need enriched cybersecurity dataset? Recent Attacks

Why we need enriched cybersecurity dataset? Recent Attacks

Sep 2007
Nov 2007
Jan 2014
Dec 2015
Oct 2016
Apr 2017
May 2017
Jun 2017

DHS Generator
Natanz nuclear power
North Carolina
Ukraine power grid
IOT attack on DYN
Dallas sirens
WannaCry ransomware
Notpetya ransomware Maersk

Why we need enriched cybersecurity dataset? Costs

MACQUARIE University

- In 2015 cyber attacks cost businesses as much as $400 billion a year
- From 2013 to 2015 the cyber crime costs quadrupled

- Cost of data breaches will increase to $2.1 trillion globally by 2019

- The average cost of one cyber breach
  - $4 million globally
  - $7 million in the United States

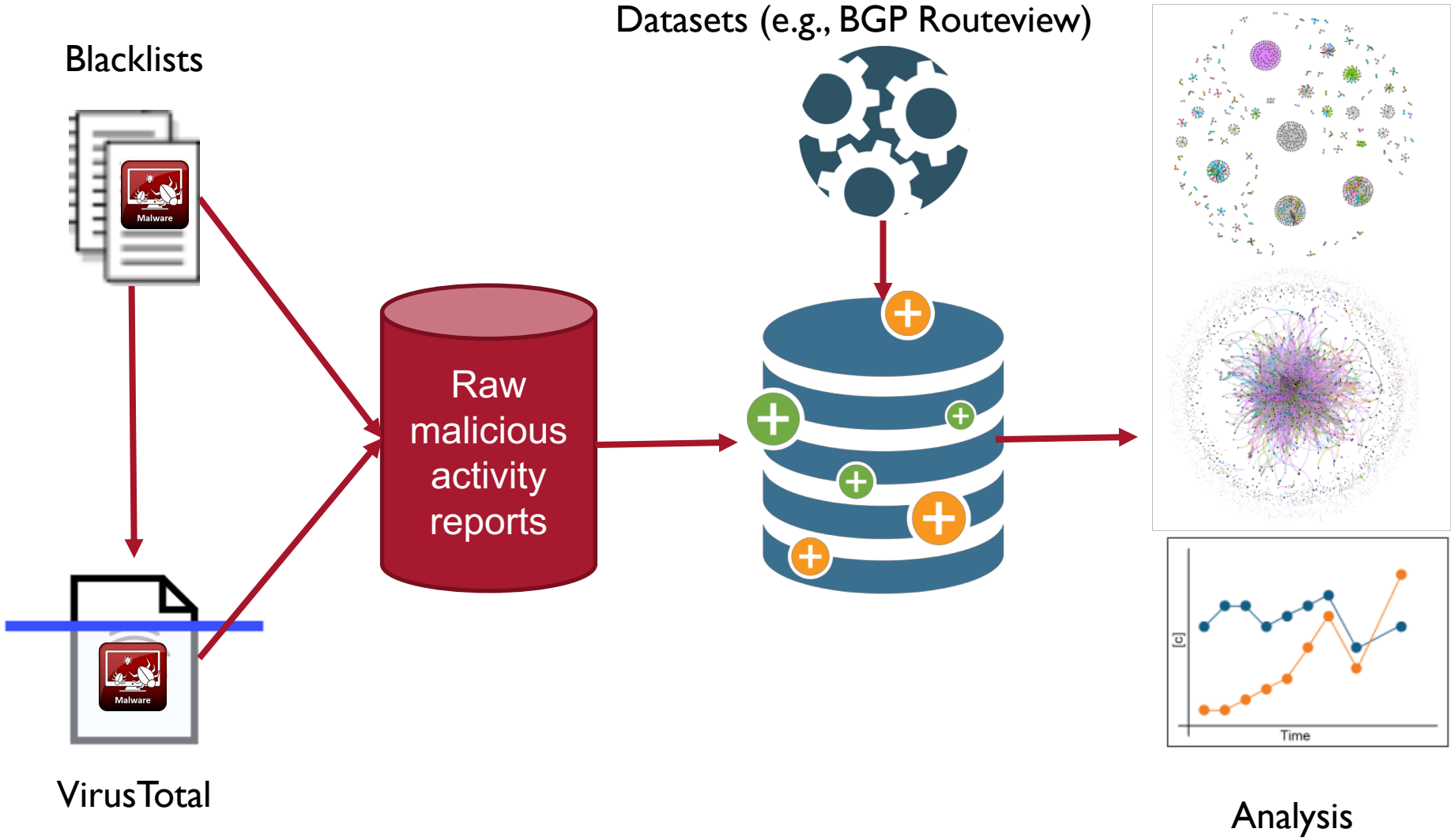
- One cyberattack can result in millions of dollars in expenses:
  - < 30 days to contain a cyberattack, the average cost is $7.7 million
  - > 90 days, the average cost is $12.2 million

Agenda

MACQUARIE University

1. Why we need enriched cybersecurity dataset?
2. **Large Scale Cybersecurity Data Collection and Enrichment Process**
3. Insights
    1. Characterization
    2. Temporal Analysis
4. Way forward: How can we leverage this dataset to improve detection systems

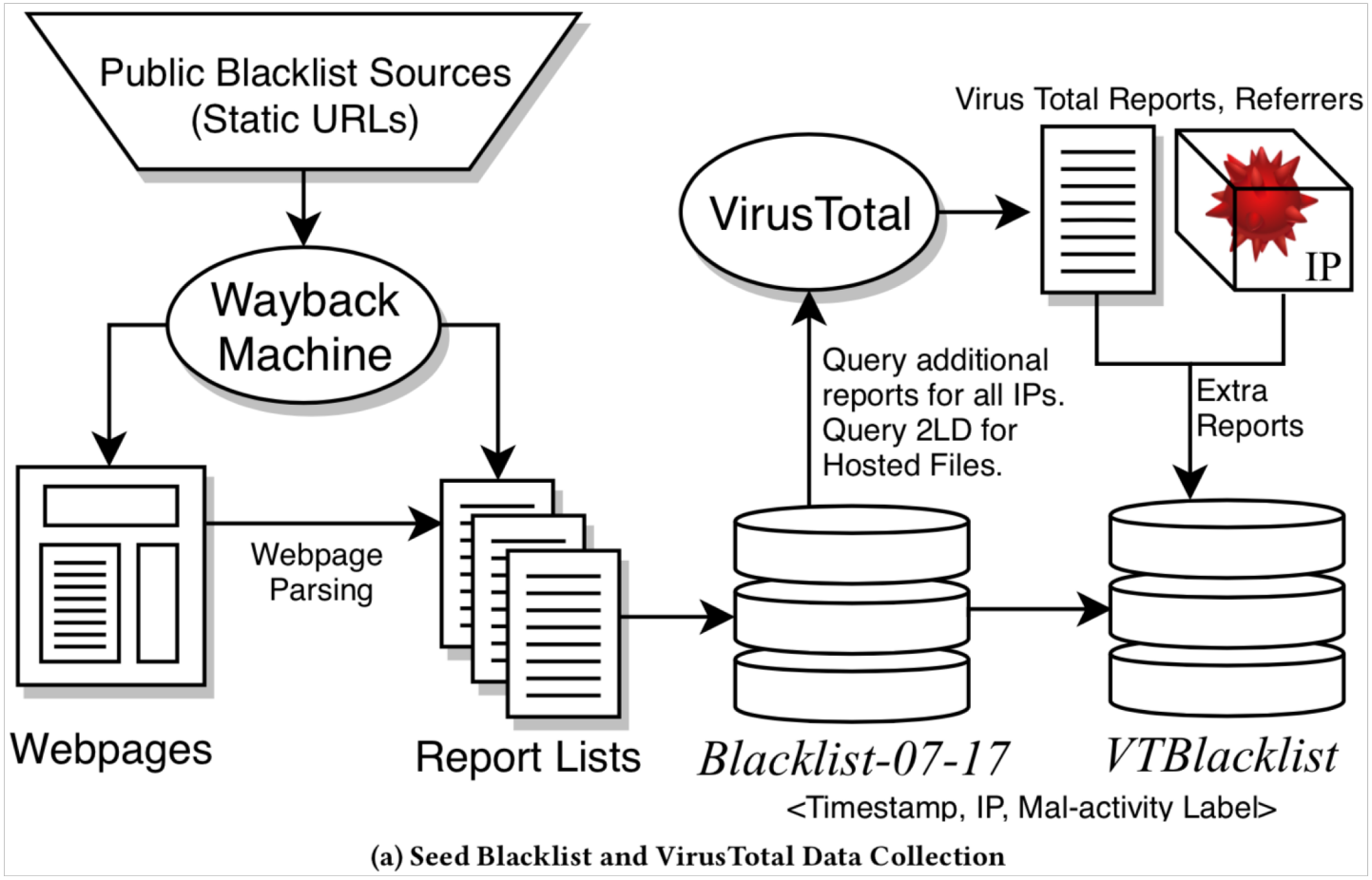
How we collect cybersecurity data at scale?



Insights: How we collect cybersecurity data at scale?

MACQUARIE University

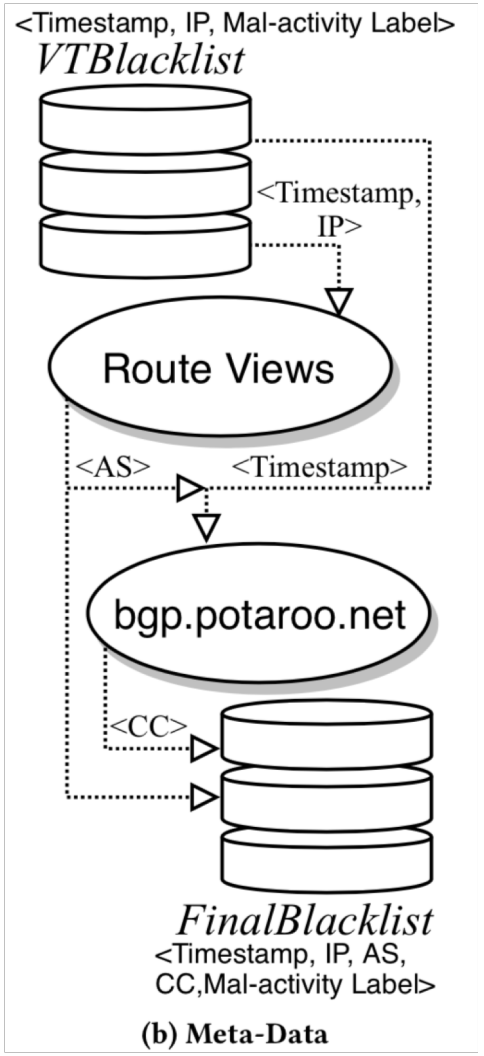
Step 1. Collecting reported malicious activities from seed blacklists and VirusTotal



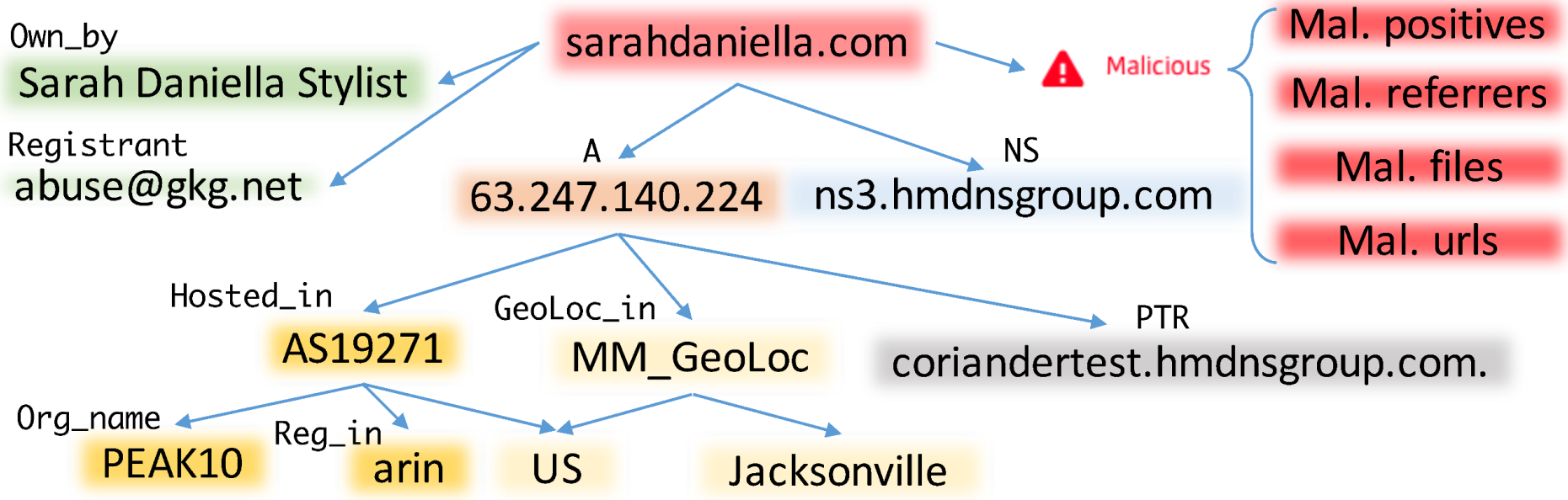
(a) Seed Blacklist and VirusTotal Data Collection

How we enrich cybersecurity data at scale?

Step 2. Leverage additional datasets (e.g., BGP Routeview and Potaroo) to enrich malicious activities dataset

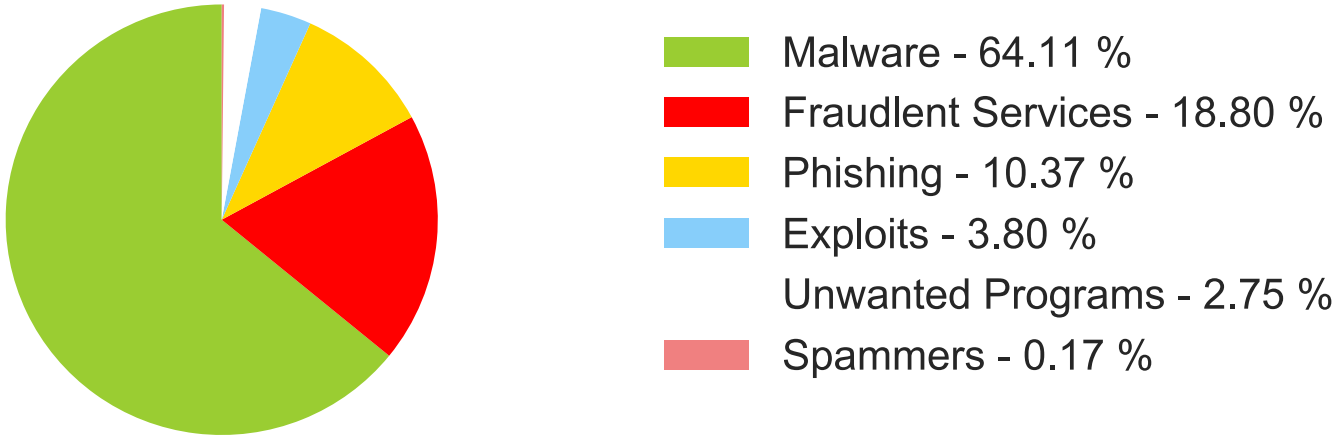


What do we have in our enriched cybersecurity data? Example



Classification Challenge

We collect 51.6M malicious activities with 15% (7.6M) of them are **labeled** by their respective data sources, and the remaining 85% (44M) **unlabeled**



15% (7.6M) Labeled Dataset

Challenge: How to label the unlabeled dataset?

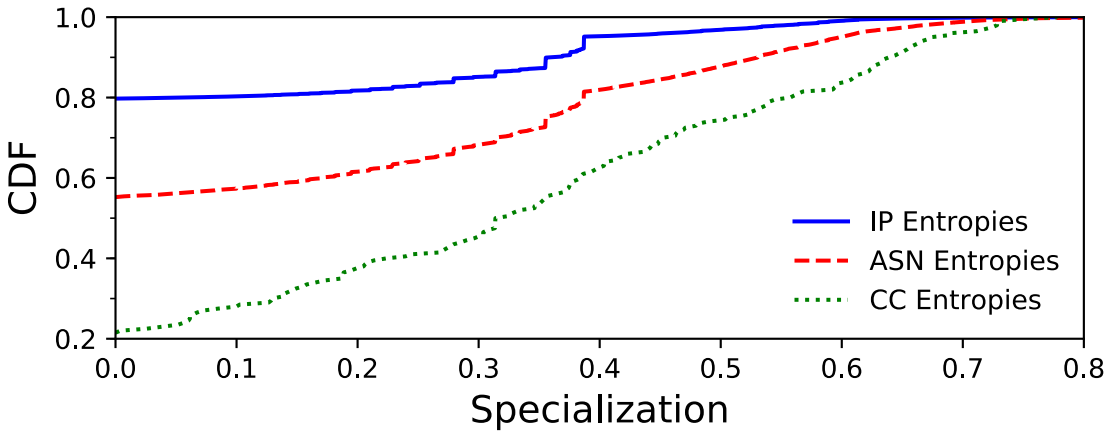
Classification Challenge

MACQUARIE University

1. Host Specialization, *S(h)*

$$S(h) = (-\sum_{\forall a} P(h) \log_2 P(h))/\log_2 k,$$

$$P(h) = \frac{\#\text{ of reports from host } h \text{ with activity } a}{\text{Total }\#\text{ of reports for host } h}$$



80% of the IPs exclusively participate in one class of malicious-activity

Classification Challenge

MACQUARIE University

2. Machine Learning Approach
   - Use features of labeled dataset to classify 44M malicious activities

Average accuracies of Malware, Phishing, Exploits, Fraudulent Services, PUP, Spammers is 93%, 94%, 79%, 92%, 96%, 83%.

92.45% Weighted Classification Accuracy

Features used in Classification Task

Feature	Data Type
Day	integer
Month	integer
Year	integer
IP bits (0-7)	integer
IP bits (8-15)	integer
IP bits (16-23)	integer
IP bits (24-31)	integer
AS	integer
Country	One-Hot encoding
Organization	One-Hot encoding

Class	# Reports	# U. IP	# U. ASes	# U. CC
Malware	46,932,466 (90.9%)	427,745 (65%)	11,435 (88%)	196 (99%)
Phishing	2,450,247 (4.74%)	133,072 (20%)	4,402 (34%)	139 (70%)
FS	1,141,377 (2.21%)	87,508 (13%)	3,264 (25%)	118 (60%)
PUP	895,494 (1.73%)	165,465 (25%)	2,200 (17%)	81 (41%)
Exploits	218,791 (0.42%)	39,854 (6%)	2,966 (23%)	112 (57%)
Spammers	7,620 (0.01%)	2,209 (0.3%)	561 (4%)	60 (30%)
Total	51,645,995 (100%)	662,409 (100%)	12,950 (100%)	198 (100%)

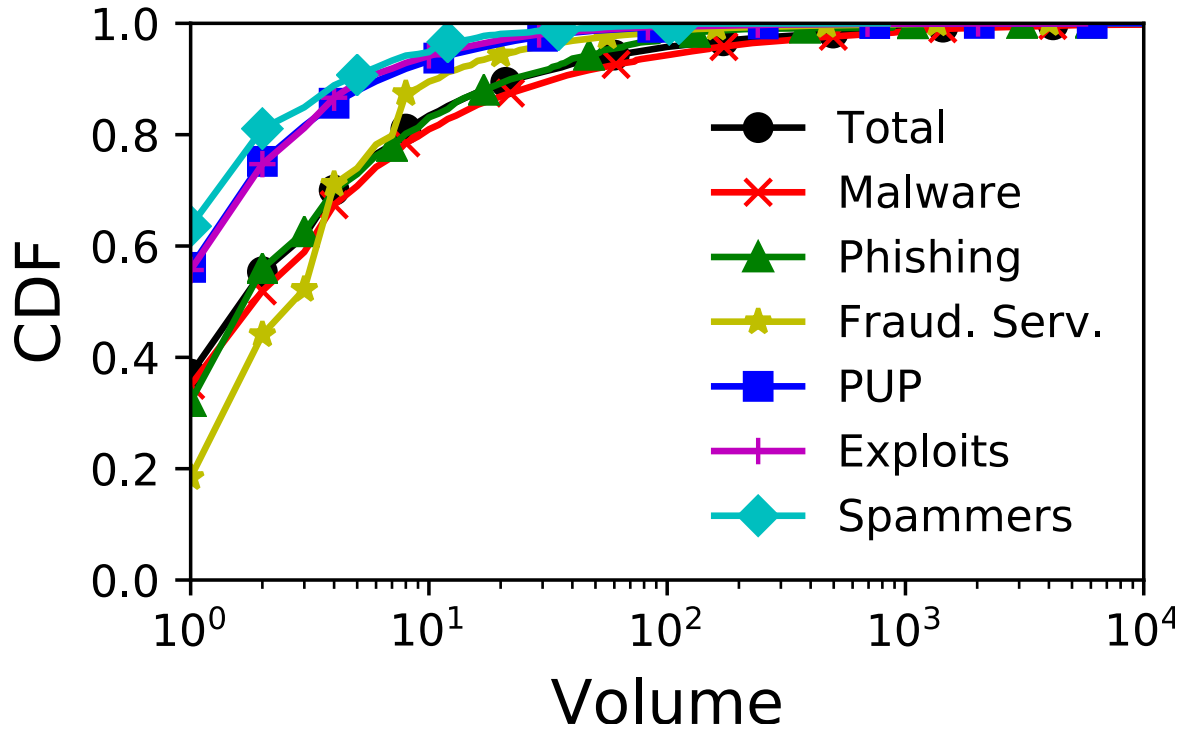
Labeled dataset from 15% (7.6M) to 92.49% (47.72M)

Agenda

MACQUARIE University

1. Why we need enriched cybersecurity dataset?
2. Large Scale Cybersecurity Data Collection and Enrichment Process
3. **Insights**
    1. **Characterization**
    2. Temporal Analysis
4. Way forward: How can we leverage this dataset to improve detection systems

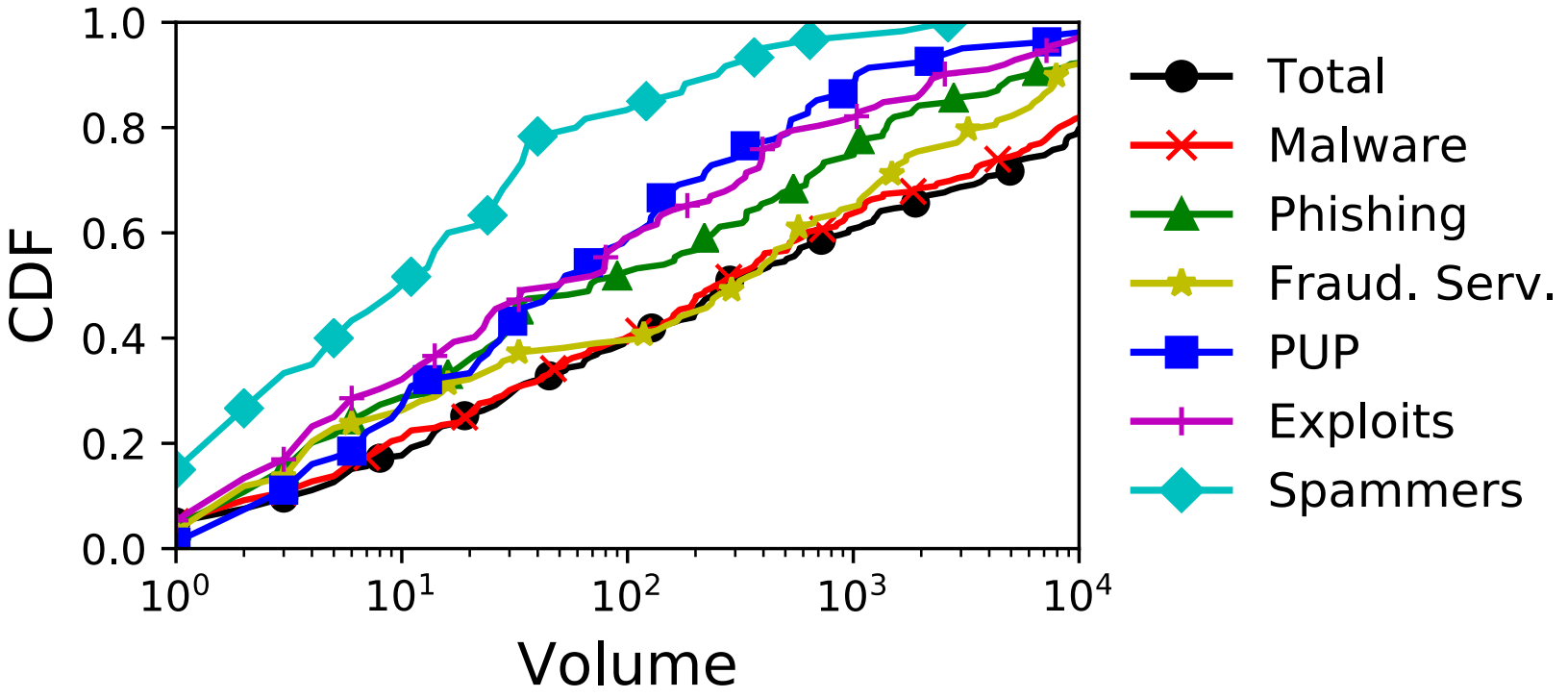
Distribution of Malicious Activities: Across IPs



IP (63.0%) are repeat offenders with FS (81.6%) and Malware (65.0%) are the most involved in more than one corresponding malicious activity

54.72.9.51 (Free AWS) is the most repeated offender with high volume of SpyEye Trojans and Exploit kits

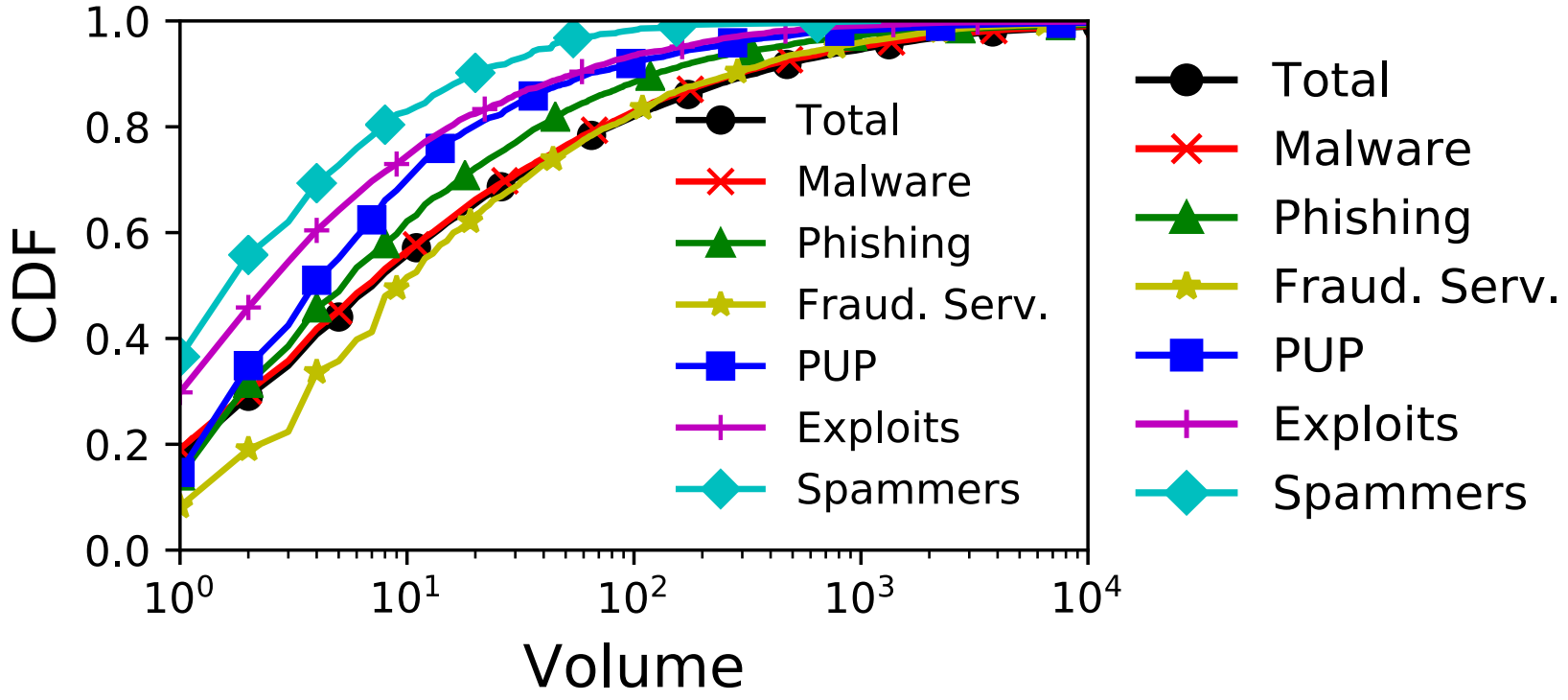
Distribution of Malicious Activities: Across Countries



Malicious activities are not evenly distributed among countries: 20.2% of countries having more than 10K malicious reports

Spamming activities: US (35%), Russia (22%), British Virgin Islands (9%), and Ukraine (5%)

Distribution of Malicious Activities: Across Autonomous Systems (ASes)



82.4% of the ASes are involved in more than one malicious activity. Spammers are distributed over the smallest proportion of ASes, only 4.33%.

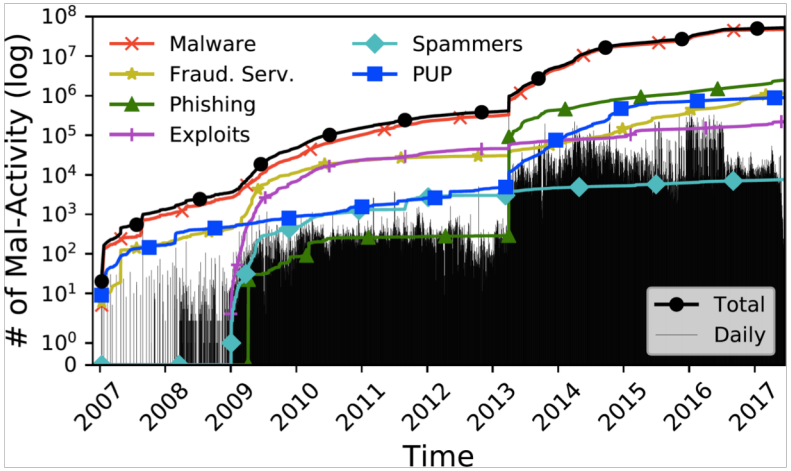
AS16509 (AMAZON-02) is most aggressive with 25.8M of all malicious reports, predominantly malware (24.5M) and phishing (463K)

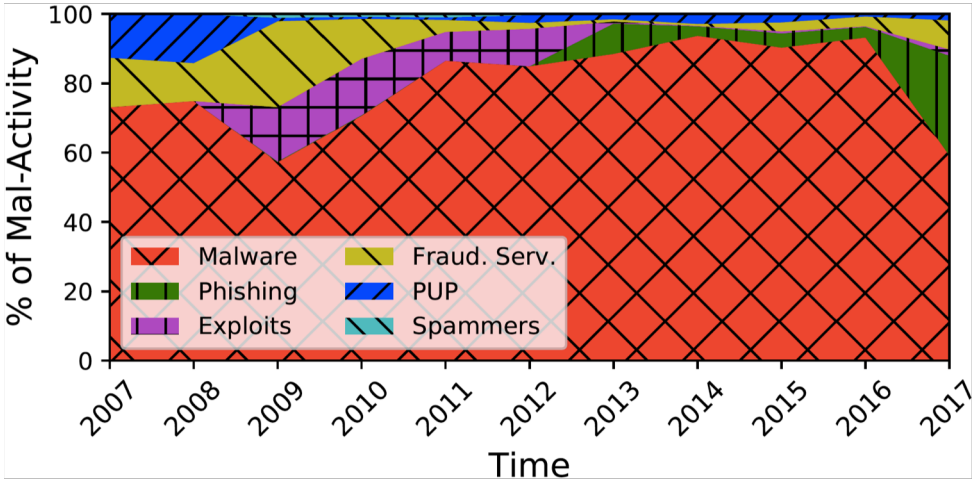
Agenda

MACQUARIE University

1. Why we need enriched cybersecurity dataset?
2. Large Scale Cybersecurity Data Collection and Enrichment Process
3. **Insights**
    1. Characterization
    2. **Temporal Analysis**
4. Way forward: How can we leverage this dataset to improve detection systems

Are Malicious activities growing? Evolution of Malicious Activities



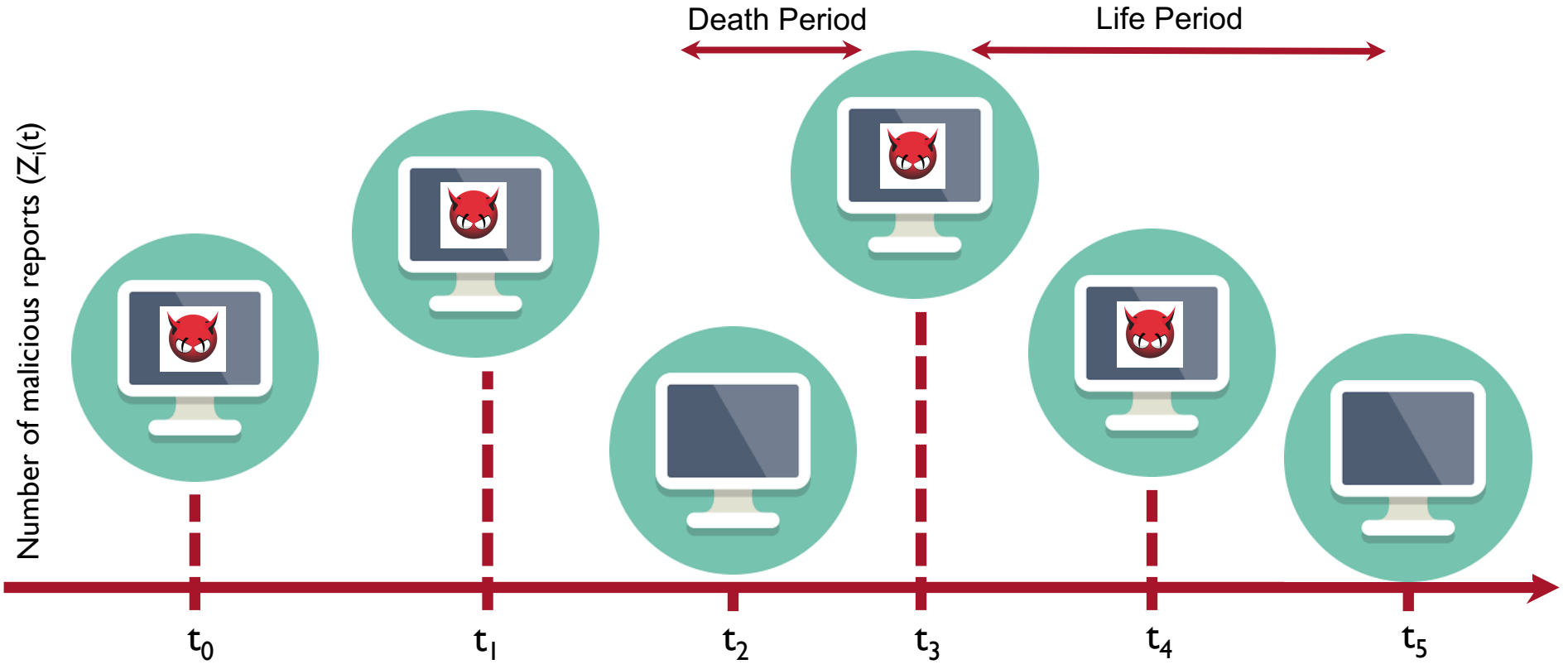


Malicious activities have been steadily increasing in volume over the last decade, with an interesting spike around 2008-2009 driven by the inception of high-profile FS and exploit kits

Phishing has recently undergone an increase in volume: 29% of all malicious activities in 2017

Do Malicious Actors Churn?
Churn, periods of presence, let's dig deeper!

- Spammers often quarantine bots for a period of time, waiting for them to be "whitelisted" again.



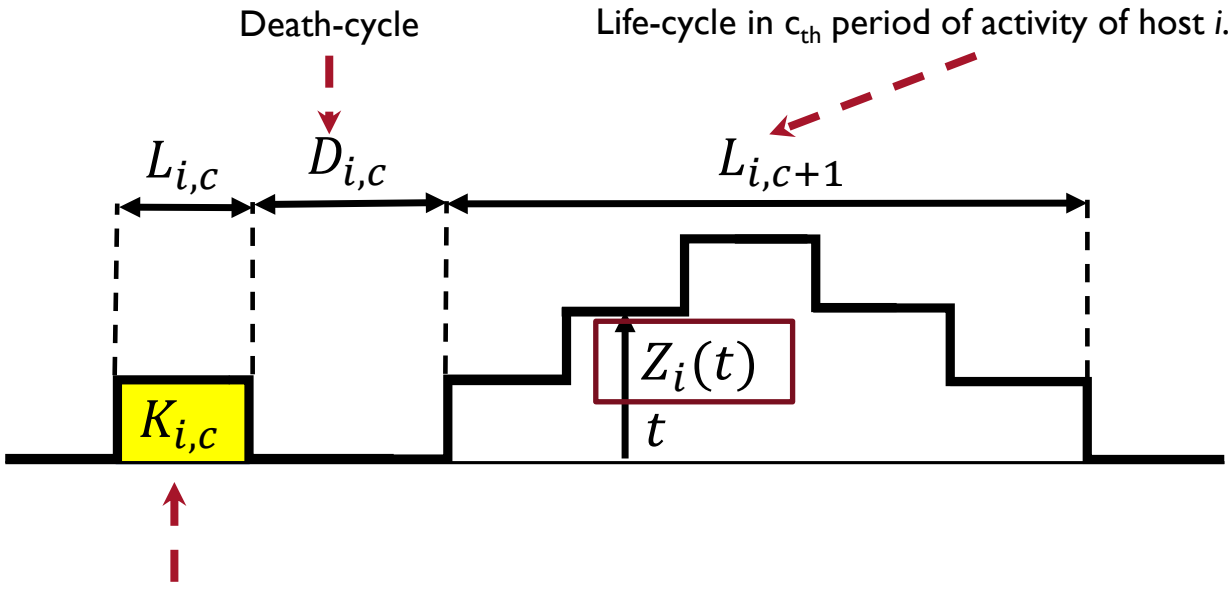
*Stone-Gross et al., The underground economy of spam: A botmaster's perspective of coordinating large-scale spam campaigns. LEET'11

Do Malicious Actors Churn?
Churn, periods of presence, let's do modeling!

MACQUARIE University

- Alternating renewal process $Z_i(t)$ for each host *h*, like peers churn model in P2P networks

$$Z_i(t) = \begin{cases} k, & \text{host } i \text{ has received} \\ & k \text{ reports at time } t \quad , 1 \le i \le n \\ 0, & \text{Otherwise} \end{cases}$$

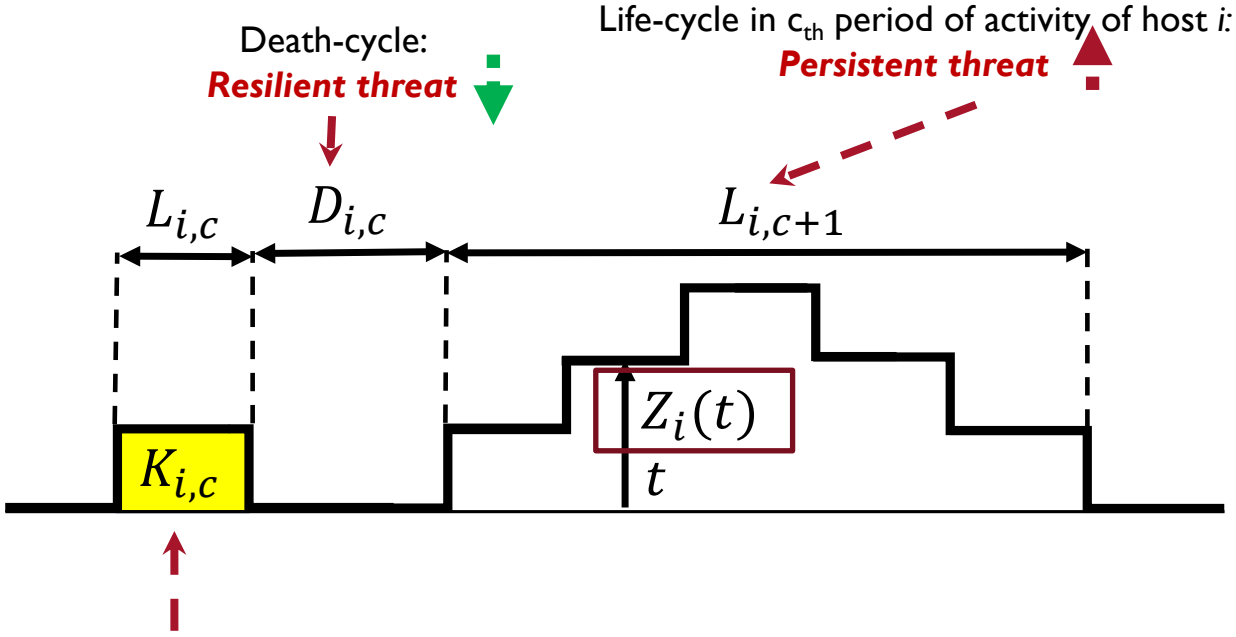


* Yao et al., Modeling heterogeneous user churn and local resilience of unstructured p2p networks, ICNP'06

Do Malicious Actors Churn?
Churn, periods of presence, let's do modeling!

- Alternating renewal process $Z_i(t)$ for each host *h*, like peers churn model in P2P networks

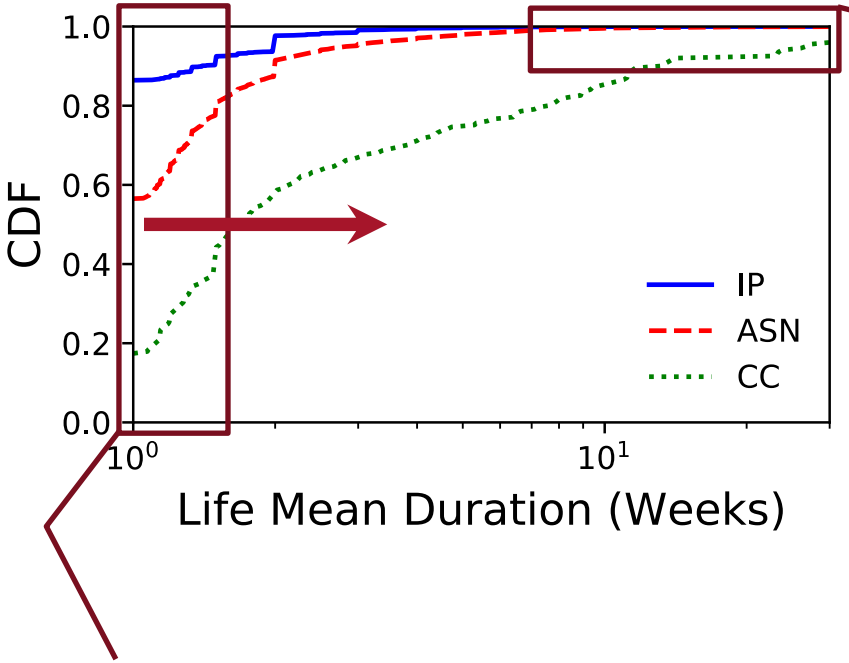
$$Z_i(t) = \begin{cases} k, & \text{host } i \text{ has received} \\ & k \text{ reports at time } t \quad , 1 \le i \le n \\ 0, & \text{Otherwise} \end{cases}$$



# of reports in the c_{th} period/cycle of activity of host *i*.

* Yao et al., Modeling heterogeneous user churn and local resilience of unstructured p2p networks, ICNP'06

Churn Analysis: Hosts
Life-cycle Time (LT) – persistency of hosts



(a) Average Lifetime - LT (Most Persistent)

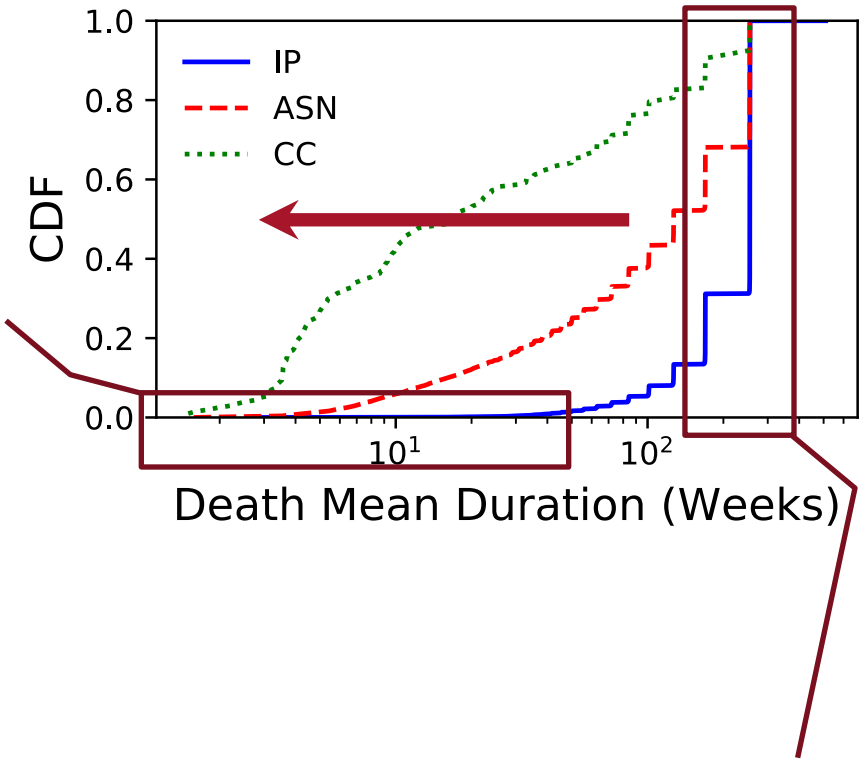
IP	LT	ASN	Organization	LT	CC	LT
209.85.200.132	62	4134	CHINANET-BACKBONE, CN	147	US	511
74.125.201.132	52	4837	CHINA169-Backbone, CN	39	CN	56
209.85.234.132	48	9800	UNICOM, CN	38	BR	55
74.125.70.132	38	32613	IWEB-AS, CA	28	CA	38
74.125.202.132	37	28753	LEASEWEB-DE-FRA-10, DE	26	GB	38

Significant portion of IPs (86.4%) are short-lived in contrast 83% of countries, mostly African or island states, are persistently participating malicious activities

Churn Analysis: Hosts
Death Time (DT) – resiliency of hosts

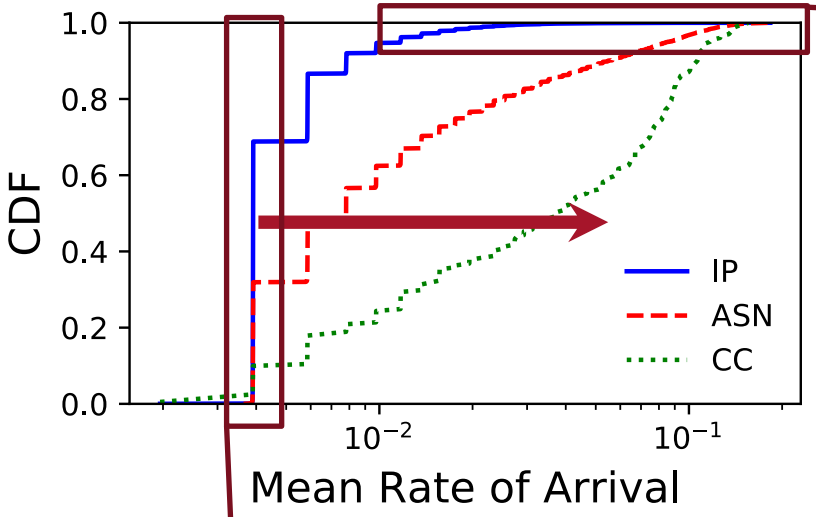
(b) Average Deathtime - DT (Most Resilient)

IP	DT	ASN	Organization	DT	CC	DT
103.224.212.222	3.0	36351	SOFTLAYER, US	1.5769	US	0
69.172.201.153	3.1	26496	GO-DADDY, US	1.6087	DE	1.5
204.11.56.48	3.7	40034	CONFLUENCE-NET., US	1.6122	VG	1.6
213.186.33.19	3.9	13335	CloudFlare, Inc. VG	1.6780	FR	1.8
208.73.211.70	4.2	14618	AMAZON-AES, US	1.8298	NA	2.0



A few IPs are recurring participants in contrast most ASes and countries are repeating offenders

Churn Analysis: Hosts
Rate of arrival – freq. of host participation, $\lambda_i = \frac{1}{L_i + D_i}$



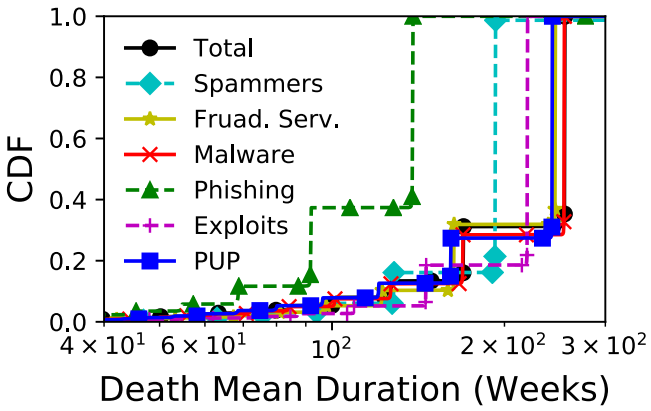
Majority of recurrent countries are African or island countries

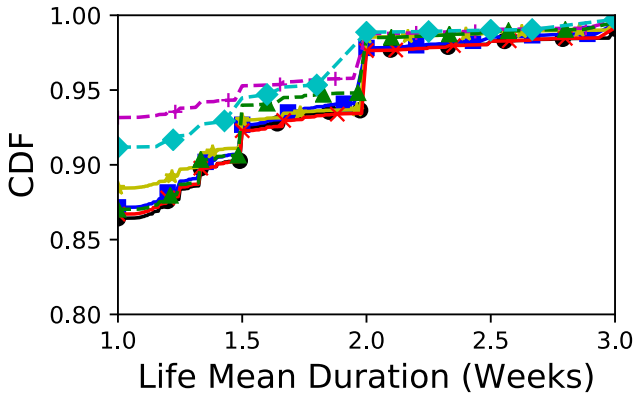
(c) Rate of Arrival - RoA (Most Frequently Active)

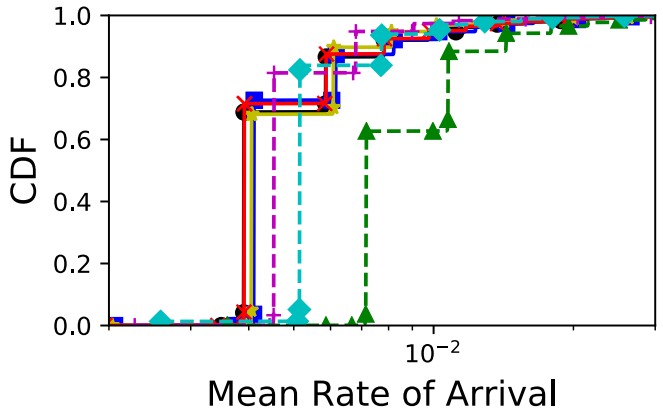
IP	RoA	ASN	Organization	RoA	CC	RoA
69.172.201.153	0.183	8001	NET-ACCESS-CORP, US	0.177	CO	0.156
103.224.212.222	0.176	9931	CAT-AP, TH	0.175	PA	0.148
208.73.211.70	0.164	46636	NATCOWEB, US	0.173	BS	0.142
213.186.33.19	0.150	13649	ASN-VINS, US	0.173	NO	0.138
213.186.33.2	0.146	31103	KEWWEB AG, DE	0.169	MX	0.138

Significant portion of countries (70%) and ASes (38%) are recurrent offenders in contrast 9% IPs

Churn Analysis
Malicious Activities Classes

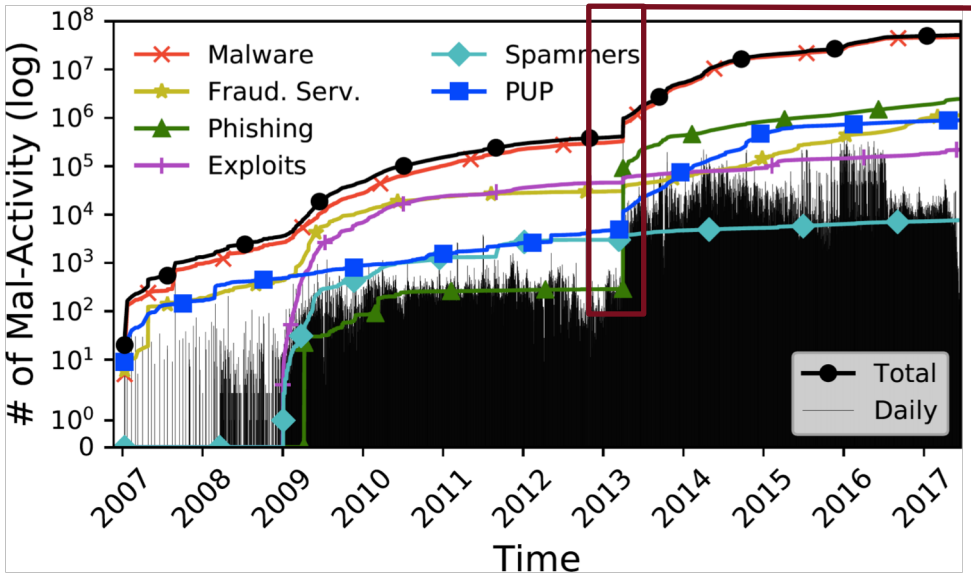






Exploits have lowest mean lifetime (least persistent) in contrast Phishing reports are resilient (lowest death duration) and recurrent (highest RoA)

Severity Metric and Analysis
Let's have a closer look at ephemeral activities



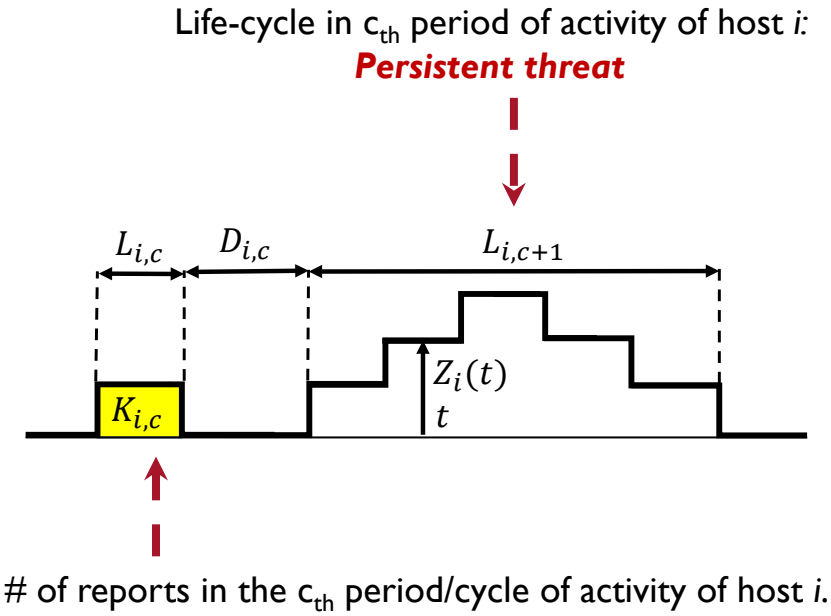
Hosts (resp. malicious activities) may be ephemeral but denser (spikes)

- How to distinguish between long-living persistent threats and short-living but denser or aggressive malicious activities (resp. hosts)?

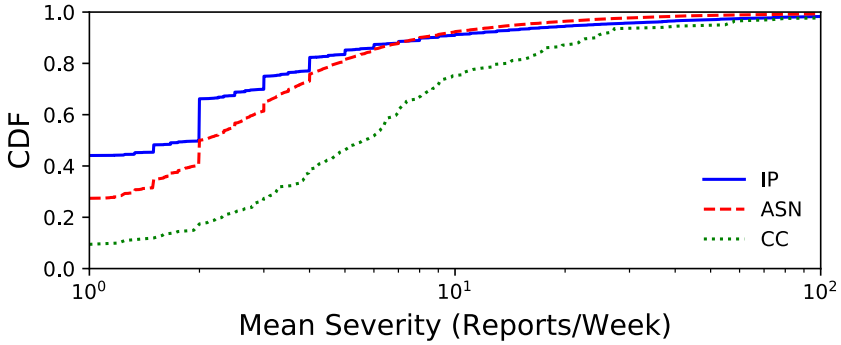
Severity Metric and Analysis

- How to distinguish between long-living persistent threats and short-living but denser or aggressive malicious activities (resp. hosts)?
- We define, Severity -- average number of reports of mal-activities per active cycle,

$$s_i = E\left[\frac{K_{i,c}}{L_{i,c}} \forall c\right],$$



Severity Metric and Analysis: Hosts



ASN	Organization	Mag.	CC	Mag.
7276	UNIVERSITY-OF-HOUSTON	2206	US	82558
6762	SEABONE-NET, IT	2153	CN	377
16509	AMAZON-02	1817	DE	212
35994	AKAMAI-AS	1707	FR	149
53684	FLASHPOINT-SC-AS	1607	UA	80

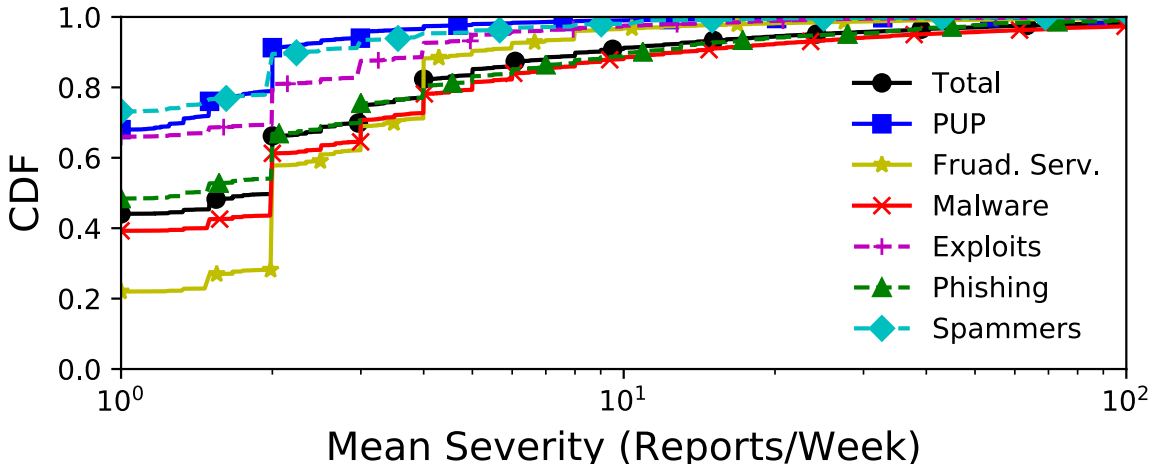
(a) Top AS, Countries (CC) magnitude offenders

27.4% of ASes and 9.45% of countries have a severity value equal to one indicating a unique malicious report per week

200 IP addresses reported to be involved in more than 10K malicious activities per

Cloud provider services (e.g., Amazon Cloud) and hosting providers unlikely to be intentionally propagating their own mal-activity;
they are potentially misused by malicious actors and

Severity Metric and Analysis: Malicious Activities Classes



Fraudulent services are reported in the "low severity" range; probably evading detection

More Insights, Dataset and Code

MACQUARIE University

A Decade of Mal-Activity Reporting: A Retrospective Analysis of Internet Malicious Activity Blacklists

Benjamin Zi Hao Zhao
benjamin.zhao@unsw.edu.au
University of New South Wales
Data61, CSIRO

Muhammad Ikram
muhammad.ikram@mq.edu.au
Macquarie University
University of Michigan

Hassan Jameel Asghar
hassan.asghar@mq.edu.au
Macquarie University
Data61, CSIRO

Mohamed Ali Kaafar
dali.kaafar@mq.edu.au
Macquarie University
Data61, CSIRO

Abdelberi Chaabane
contact@chaabane.org

Kanchana Thilakarathna
kanchana.thilakarathna@sydney.edu.au
The University of Sydney

ABSTRACT

This paper focuses on reporting of Internet malicious activity (or mal-activity in short) by public blacklists with the objective of providing a systematic characterization of what has been reported over the years, and more importantly, the evolution of reported activities. Using an initial seed of 22 blacklists, covering the period from January 2007 to June 2017, we collect more than 51 million mal-activity reports involving 662K unique IP addresses worldwide. Leveraging the Wayback Machine, antivirus (AV) tool reports and several additional public datasets (e.g., BGP Route Views and Internet registries) we enrich the data with historical meta-information including geo-locations (countries), autonomous system (AS) numbers and types of mal-activity. Furthermore, we use the initially labelled dataset of ≈ 1.57 million mal-activities (obtained from public blacklists) to train a machine learning classifier to classify the remaining unlabeled dataset of ≈ 44 million mal-activities obtained through additional sources. We make our unique collected dataset (and scripts used) publicly available for further research.

Malicious Activity Blacklists. In *ACM Asia Conference on Computer and Communications Security (AsiaCCS '19), July 9–12, 2019, Auckland, New Zealand.* ACM, New York, NY, USA, 13 pages. https://doi.org/10.1145/3321705.3329834

1 INTRODUCTION

Public reports of malicious online activity are commonly used in the form of blacklists by intrusion detection systems, spam filters and alike to determine if a host is known for suspicious activity. However very little is known about the dynamics of the reporting of malicious activities. Understanding what has been reported and how the reported activity evolves over time can be of paramount importance to help assess the efficacy of blacklist-based threat prevention systems. We conduct a longitudinal measurement study of reporting of malicious online activities (abridged to *mal-activities*), over a ten-year period (from January 2007 to June 2017). We define a mal-activity as *any activity reported by one or more public data sources* (in particular, within blacklists). The actor or entity behind each mal-activity can be reduced to a combination of IP address,

https://internetmaliciousactivity.github.io

Agenda

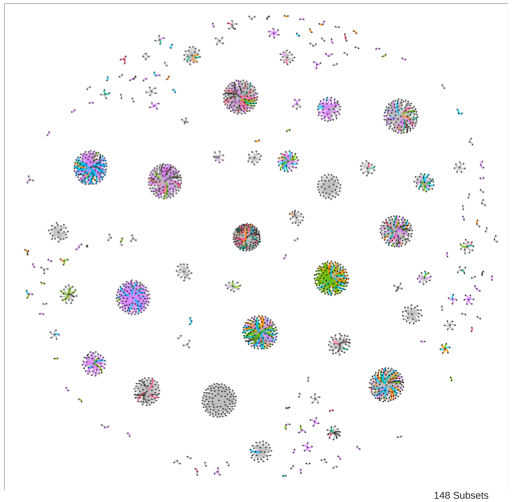
MACQUARIE University

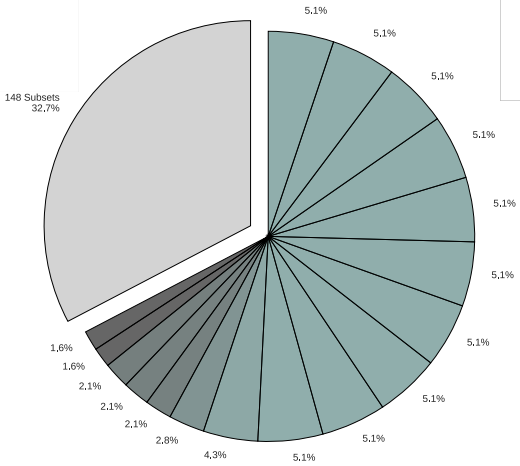
1. Why we need enriched cybersecurity dataset?
2. Large Scale Cybersecurity Data Collection and Enrichment Process
3. Insights
    1. Characterization
    2. Temporal Analysis
4. **Moving forward: How can we leverage this dataset to improve detection systems**

Cybersecurity Use-case

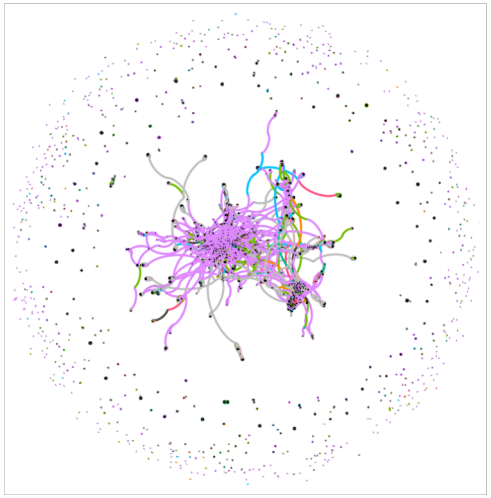
- Graph analytics/machine learning to detect and prevent subgraph of malicious actors

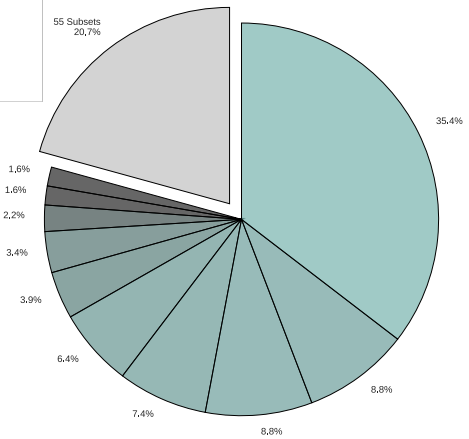
Phishing IP referrers





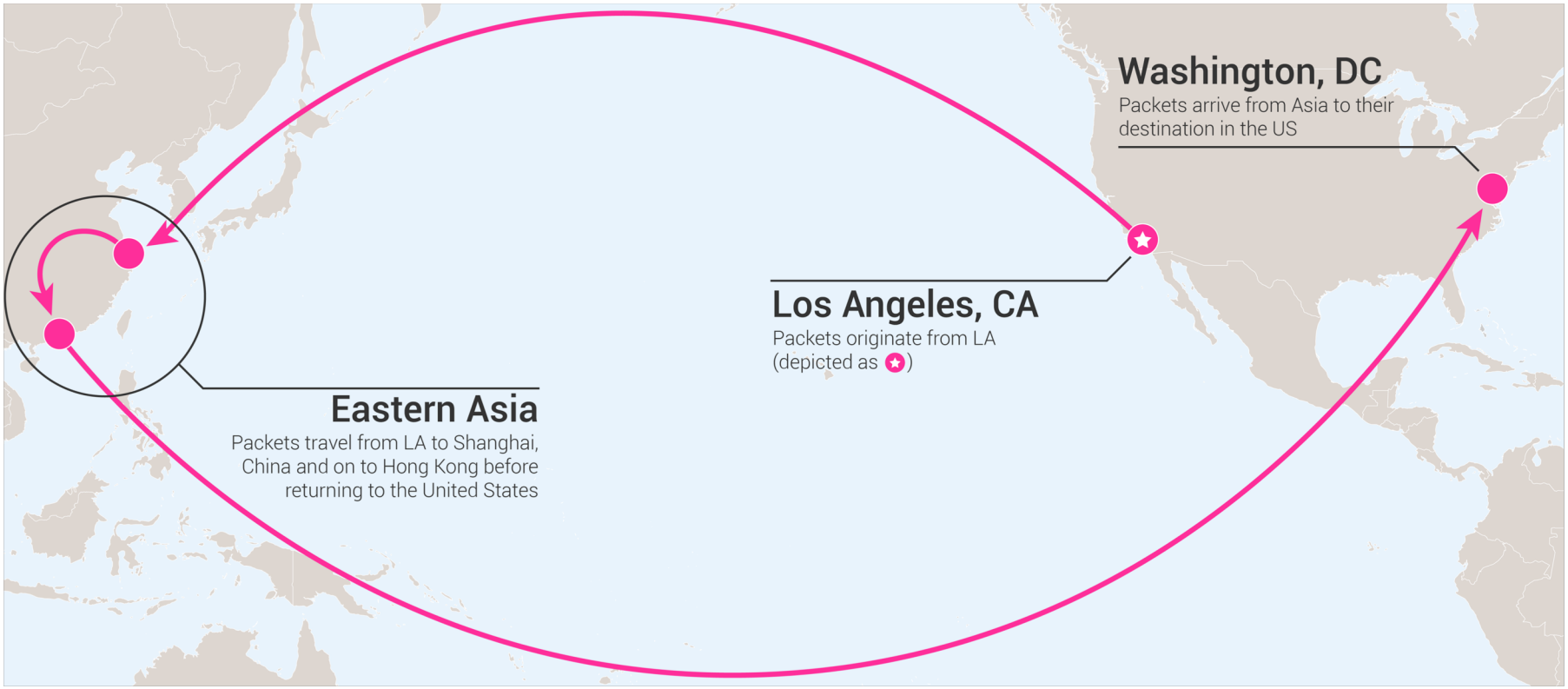
Exploits kits on IPs





Cybersecurity Use-case

- Internet traffic (mal)mis-direction prediction and malicious host behavior prediction



MACQUARIE University

Question(s)?

https://internetmaliciousactivity.github.io

For details and further info:

Muhammad Ikram
(Muhammad.Ikram@mq.edu.au)