Resident Evil: Understanding Residential IP Proxy as a Dark Service

Xianghang Mi, Xuan Feng, Xiaojing Liao
Baojun Liu, XiaoFeng Wang, Feng Qian
Zhou Li, Sumayah Alrwais, Limin Sun , Ying Liu

INDIANA UNIVERSITY

清華大學
Tsinghua University

جامعة
الملك سعود
King Saud University
1957

Background: Web Proxies

Tor

VPN

HTTP/HTTPS /SOCKS

✗ Exit nodes are constrained

✗ Exit nodes are distinguishable

✗ Exit nodes may be heavily abused

Service blocking or degradation

Background: Residential IP Proxy as a Service

proxyrack HTTP & SOCKS Proxies Residential VPN Tutorials ☎ +1 844 309 3256 Sign in SIGNUP HERE

USA & CANADA	EUROPE	ASIA/OCEANIA
230,090	345,135	115,045

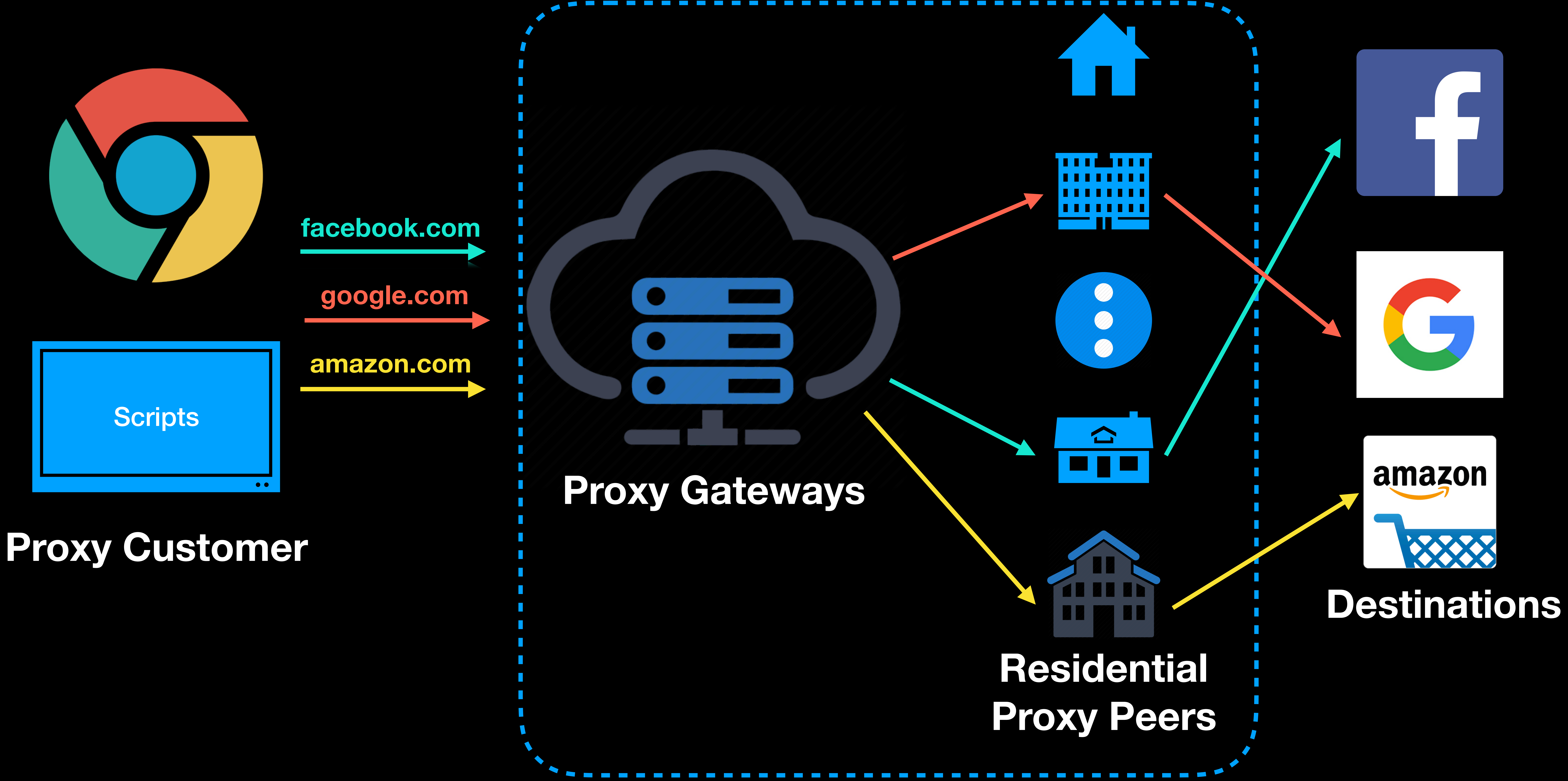
Background: Residential IP Proxy as a Service

- Millions of Residential IPs
- Clean IPs, Never Get Blocked
- Globally Distributed
- No Traffic Limits

Outline

Service Overview	Network Structure & Scale & Distribution
Residential or Not	Are proxy peers authentically residential IP addresses?
Evasiveness	How well can proxy peers evade traffic detection or blocking?
Recruitment	How can millions of proxy peers get recruited?
Usage	What are those proxies used for, in the real world?
Misc. Findings	Collusion, Local traffic, etc.

Service Overview: How it works



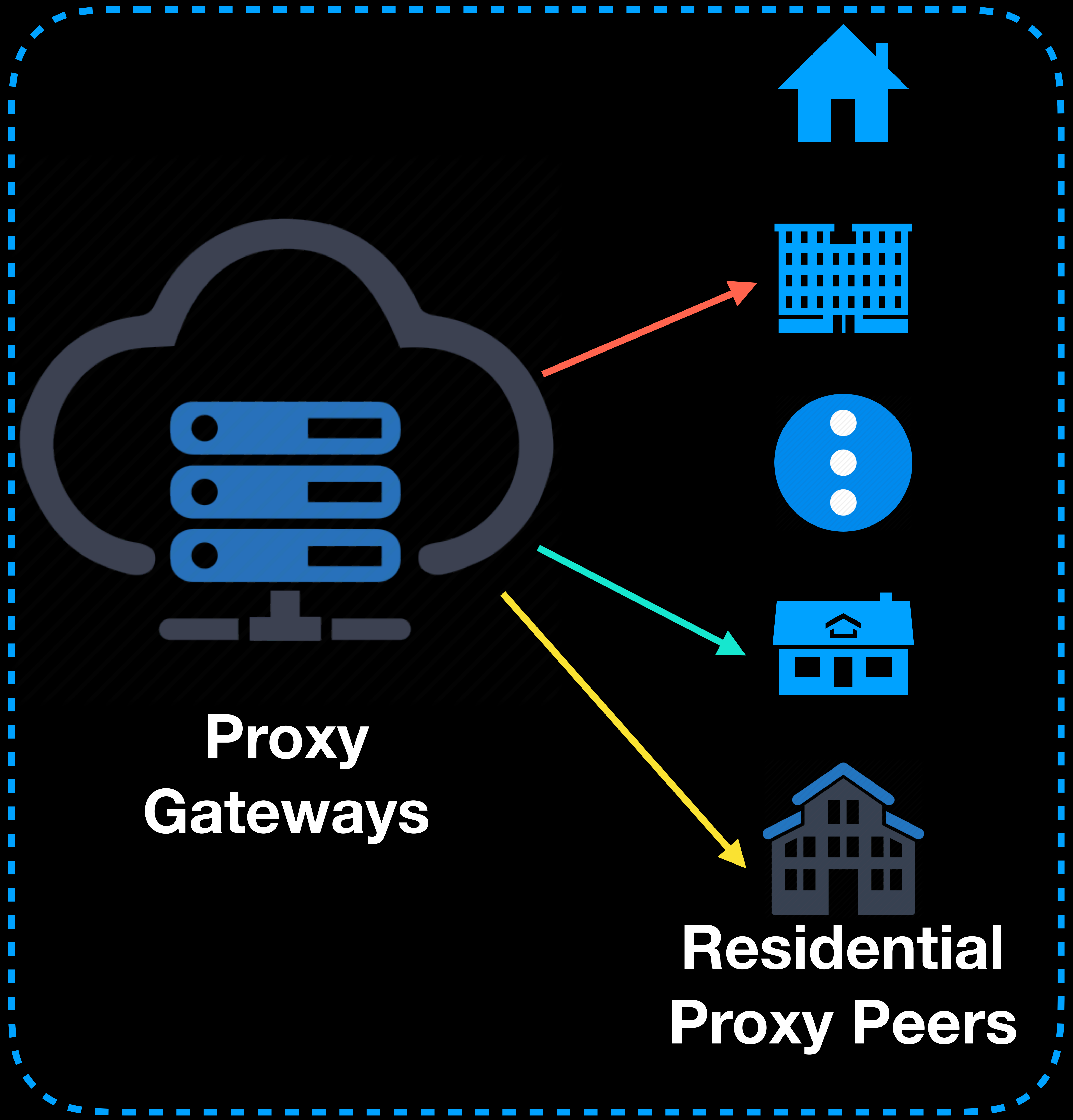
Proxy Customer

Proxy Gateways

Residential Proxy Peers

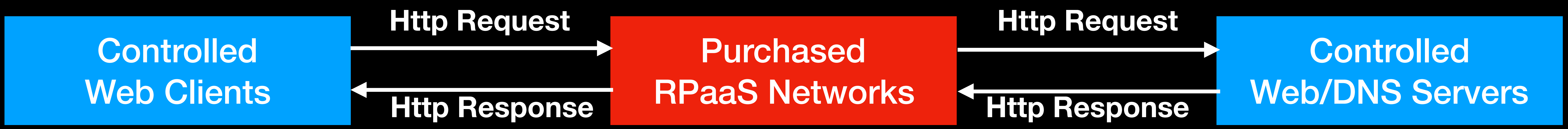
Destinations

Service Overview: How it works



- Back-connect proxy model, proxy peers are hidden from customers
- HTTP/HTTPS/SOCKS
- Multiple rotating strategies: sticky & non-sticky
- Allow customers to customize location of proxy peers

Service Overview: Scale

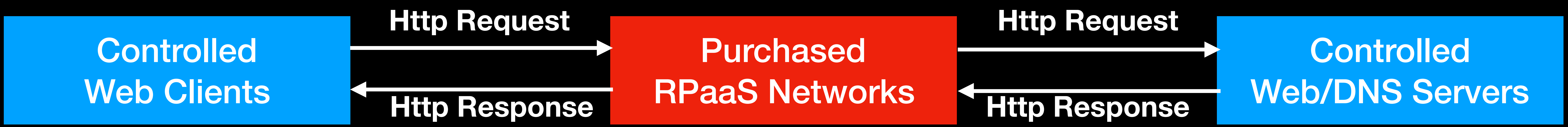


✓ Each request is identified by a unique subdomain

✓ Each request/response has payload encrypted and signed

Provier	Price	Payment	Infiltration Period
Proxies Online	$25/GB	Paypal	07/06/2017 - 11/24/2017
Geosurf	$300/month	Paypal	09/17/2017 - 10/22/2017
ProxyRack	$40/month	Bitcoin	09/18/2017 - 11/24/2017
Luminati	$500/month	Paypal	09/25/2017 - 11/01/2017
IAPS Security	$500/month	Bitcoin	09/23/2017 - 11/01/2017

Service Overview: Scale



✓ Each request is identified by a unique subdomain

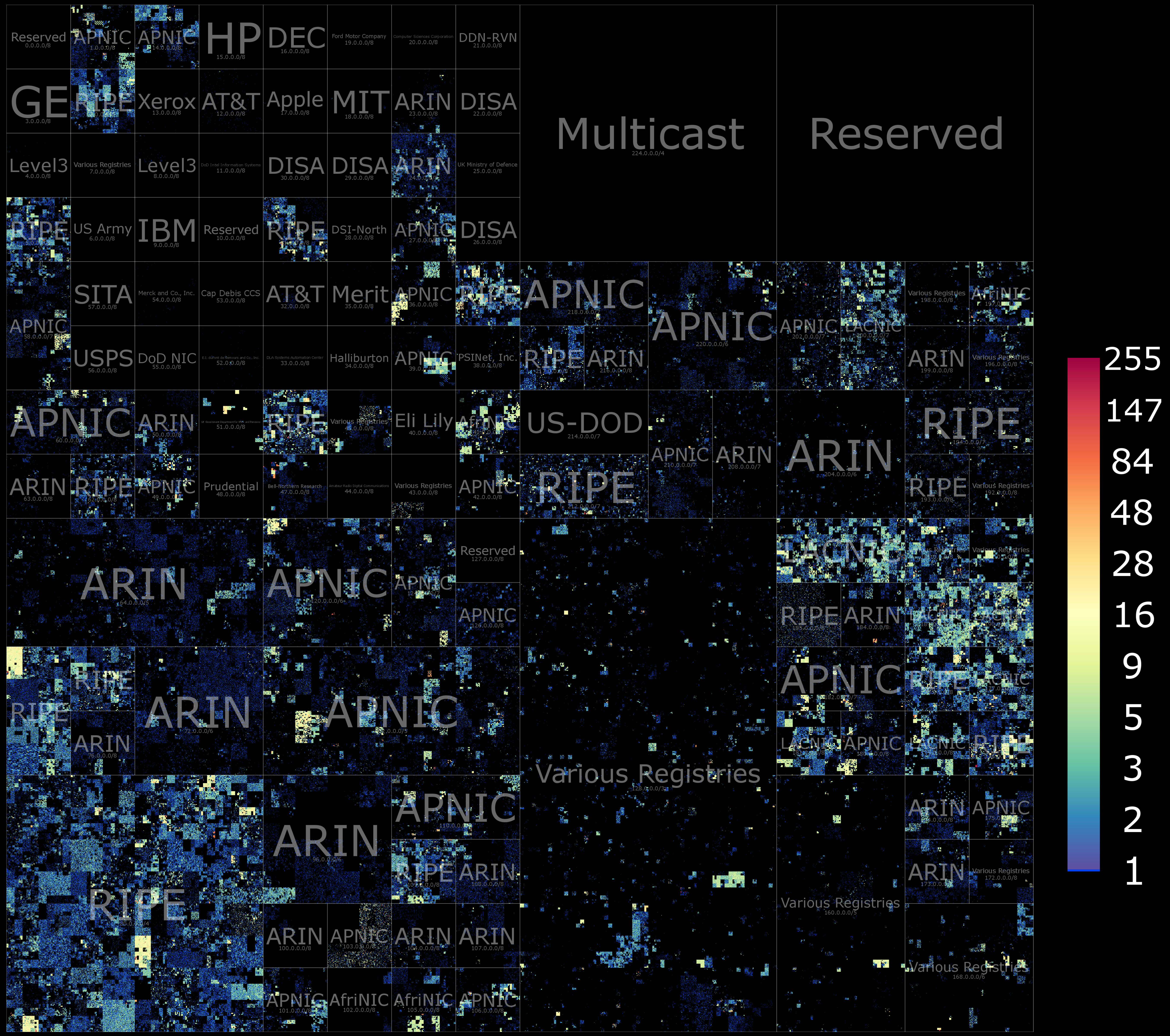
✓ Each request/response has payload encrypted and signed

60+ millions of successful probes

6.2 millions of unique IPv4 addresses

238 countries/regions, 52K+ ISPs.

Service Overview: Distribution

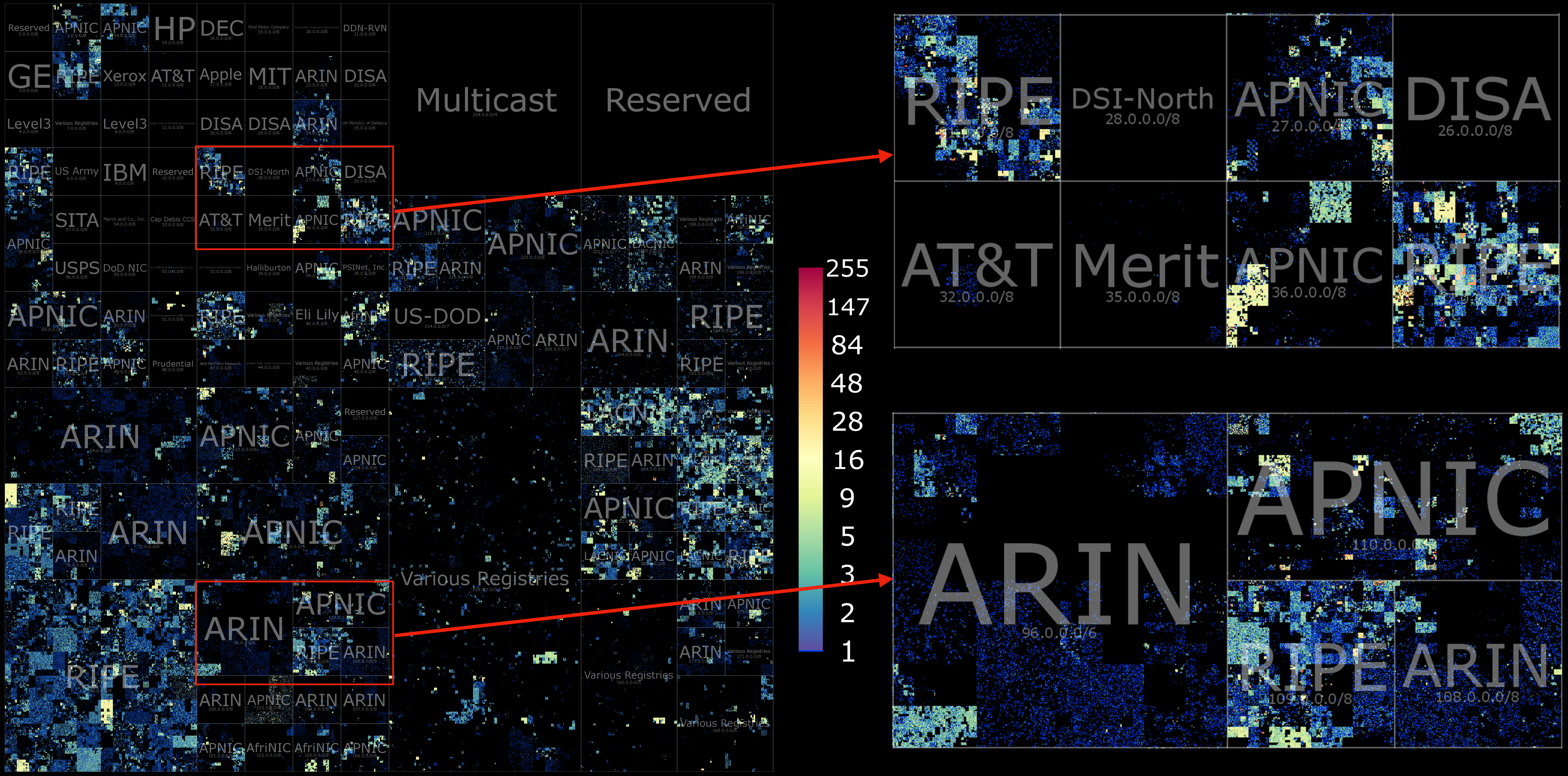


4096 * 4096 bitmap

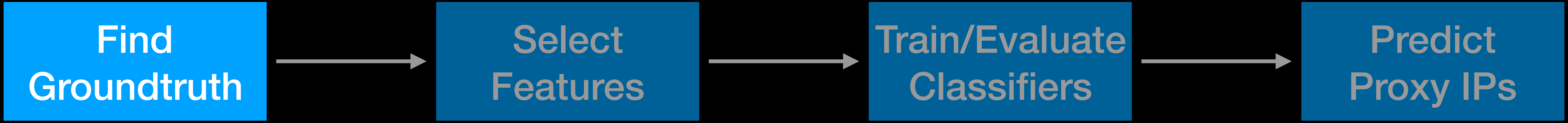
Each /24 IPv4 prefix is mapped to a pixel, using Hilbert curve of order 12

Different pixel colors denote # of proxy IPs for a given /24 prefix

Service Overview: Distribution



Residential or Not

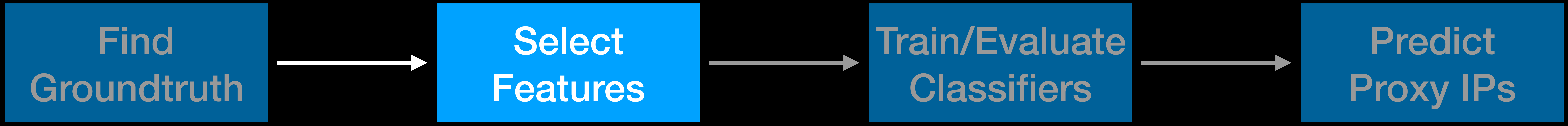


★ GT sources of various noise levels

★ Clean GT for training, noisy for evaluation

Source	Label	# IPs	# /16	# /8	# Training
Manual	resi-clean	79	25	19	79
Device Search Engine	resi-clean	89,345	13,525	195	9,921
Trace My IP	resi-noisy	37,480	11,402	213	0
Filtered IP Whois	resi-noisy	23,264,961	394	31	0
IoT Botnets	resi-noisy	1,699,291	20,112	200	0
Public Clouds	non-resi-clean	53,716,321	968	99	5,000
Alexa Top1M	non-resi-clean	442,989	14,365	213	4,481
Commercial Proxies	non-resi-clean	519	71	44	519
Public Proxies	non-resi-noisy	148,509	14,004	204	0

Residential or Not


Find Groundtruth

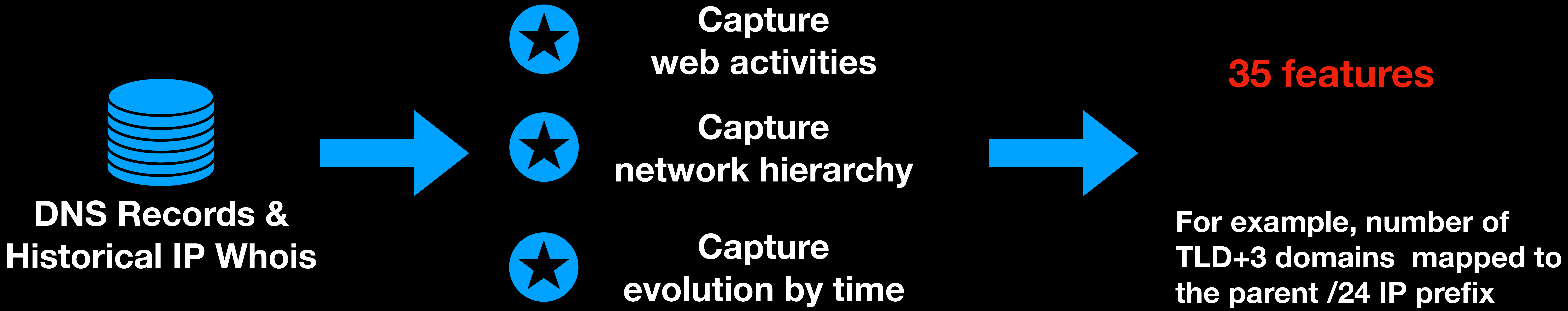
Select Features

Train/Evaluate Classifiers

Predict Proxy IPs

💡 Residential IPs/prefixes are usually web clients instead of servers

💡 Residential IPs/prefixes tend to be directly managed by ISPs


DNS Records & Historical IP Whois

★ Capture web activities

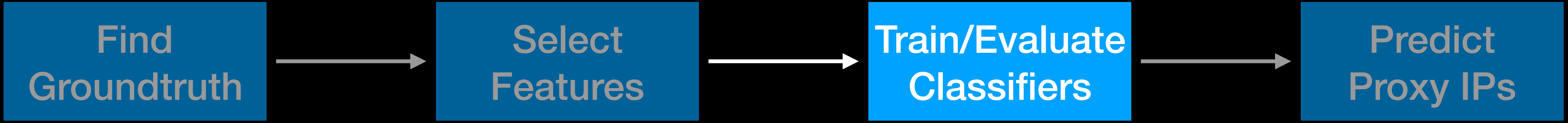
★ Capture network hierarchy

★ Capture evolution by time

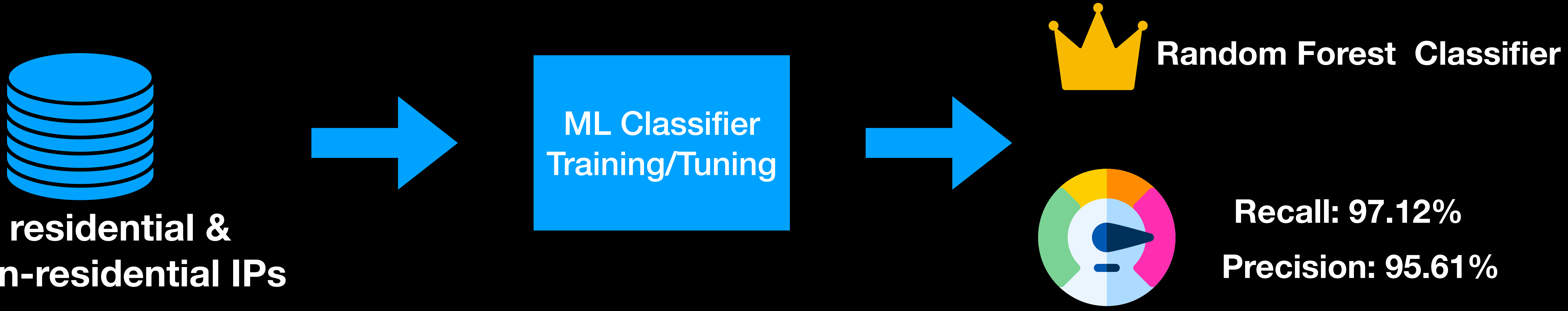
35 features

For example, number of TLD+3 domains mapped to the parent /24 IP prefix

Residential or Not



Find Groundtruth → Select Features → Train/Evaluate Classifiers → Predict Proxy IPs



10K residential &
10K non-residential IPs

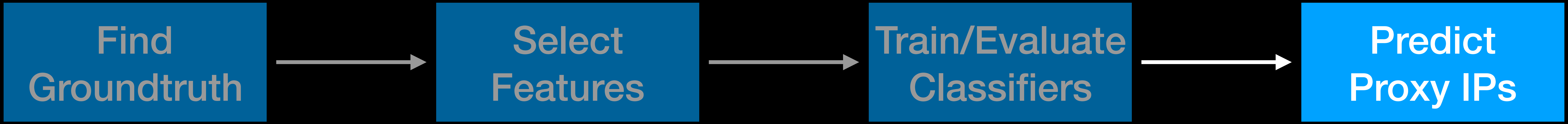
ML Classifier Training/Tuning

Random Forest Classifier

Recall: 97.12%

Precision: 95.61%

Residential or Not



Find Groundtruth → Select Features → Train/Evaluate Classifiers → Predict Proxy IPs

5.9M (95.22%) of 6.2M predicted as residential IPs

Evasiveness

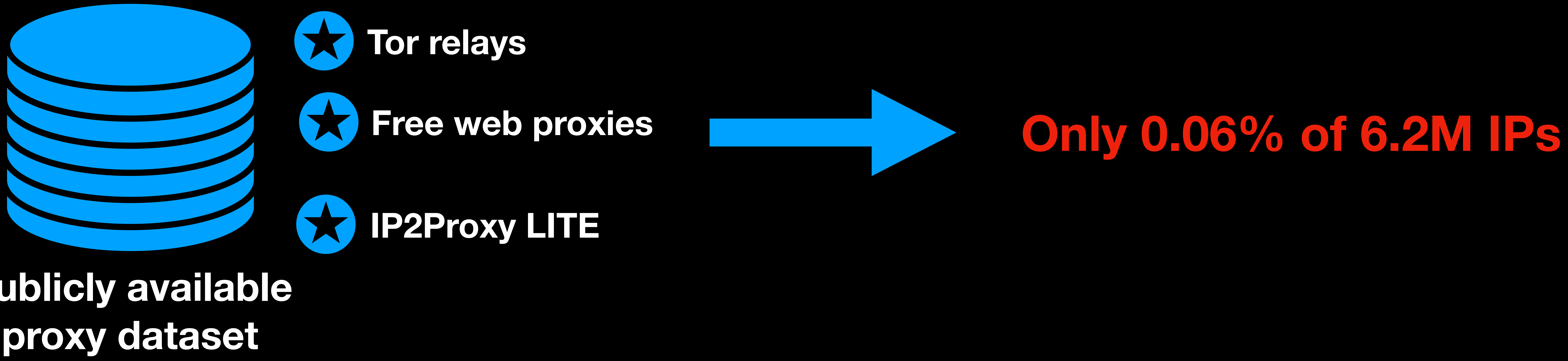
Recognized as proxy?

Identified as malicious?

Evasiveness

Recognized as proxy?

Identified as malicious?



- Tor relays
- Free web proxies
- IP2Proxy LITE

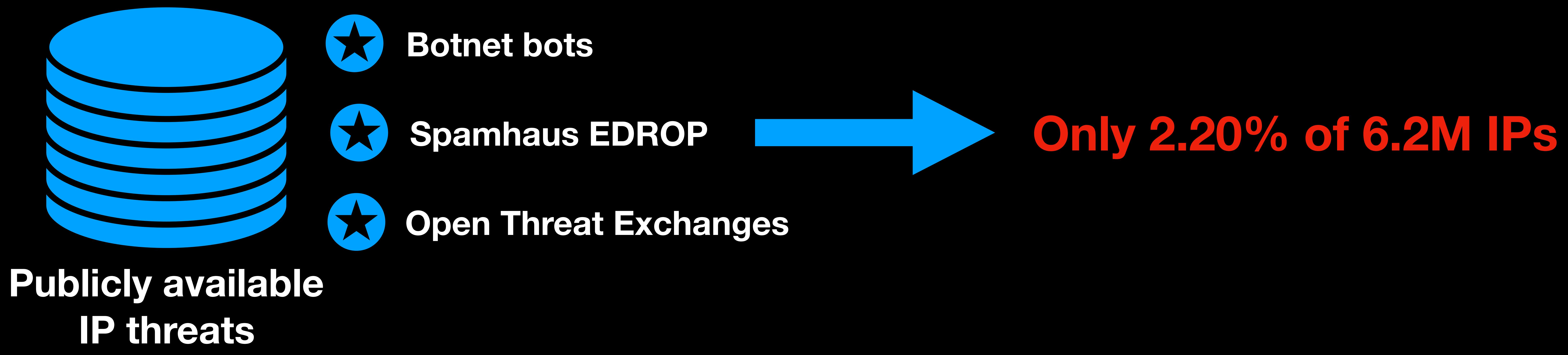
Publicly available proxy dataset

Only 0.06% of 6.2M IPs

Evasiveness

Recognized as proxy?

Identified as malicious?

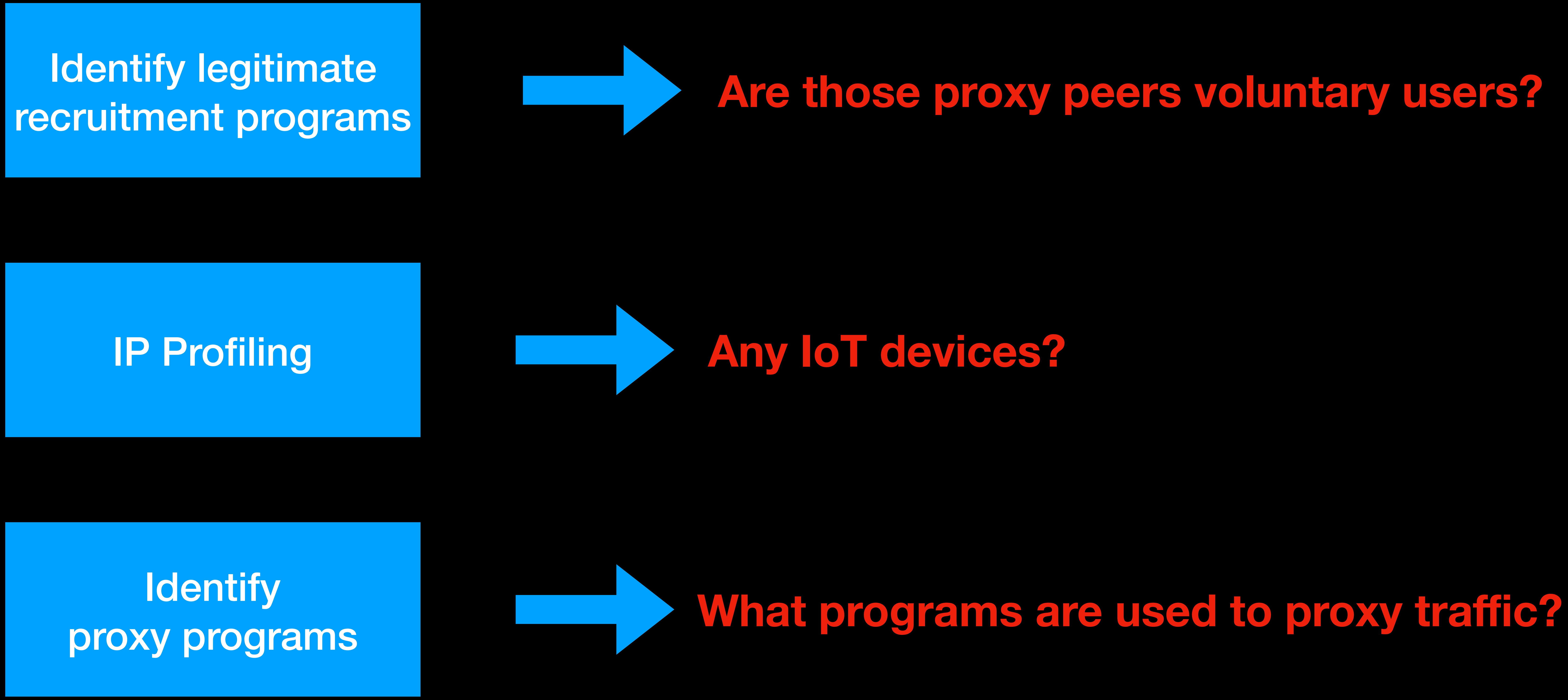


- Botnet bots
- Spamhaus EDROP
- Open Threat Exchanges

Publicly available IP threats

Only 2.20% of 6.2M IPs

Recruitment



Identify legitimate recruitment programs → Are those proxy peers voluntary users?

IP Profiling → Any IoT devices?

Identify proxy programs → What programs are used to proxy traffic?

Recruitment

Identify legitimate recruitment programs

IP Profiling

Identify proxy programs

Only Luminati was found to recruit users through Hola programs

And Hola programs were reported as problematic in previous studies

Recruitment

Identify legitimate recruitment programs

IP Profiling

Identify proxy programs

★ 730K IPs responded to our banner grabbing

★ 550K got device type identified

★ All providers got suspicious IoT devices identified for their proxy IPs, including Luminati

Device Type	Num	(%)
router	114,768	48.42
firewall	25,088	10.58
WAP	24,470	10.32
gateway	22,003	9.28
broadband	17,358	7.32
webcam	13,024	5.49
security-misc	10,608	4.48
DVR	4,249	1.79
media device	2,589	1.09
storage-misc	1,988	0.84

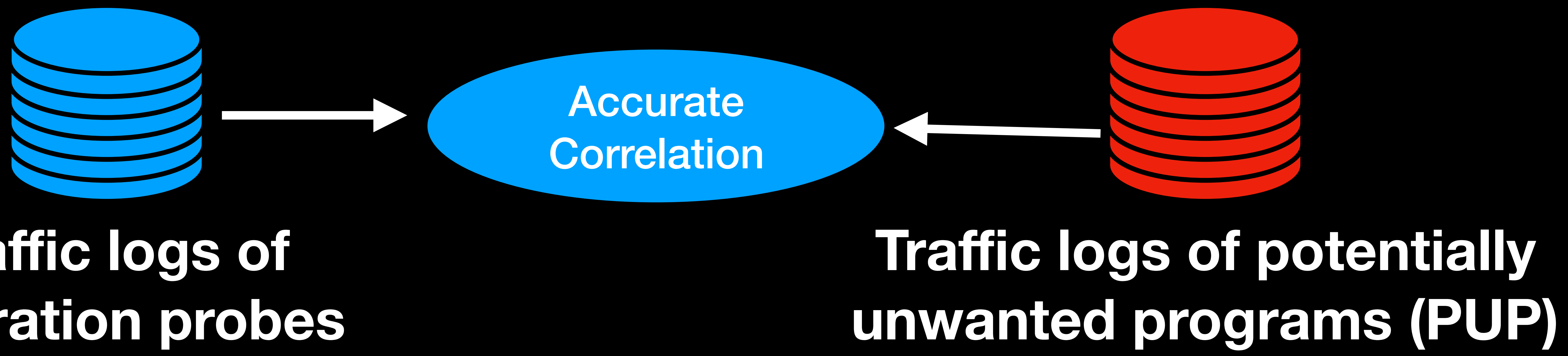
Device Vendor	Num	(%)
MikroTik	86,593	36.53
Huawei	37,545	15.84
BusyBox	18,337	7.74
Technicolor	16,866	7.12
SonicWall	14,122	5.96
Fortinet	9,190	3.88
Dahua	6,258	2.64
ZyXEL	5,601	2.36
AVM	5,272	2.22
Cyberoam	4,558	1.92

Recruitment

Identify legitimate recruitment programs

IP Profiling

Identify proxy programs


Accurate Correlation
Traffic logs of Infiltration probes
Traffic logs of potentially unwanted programs (PUP)

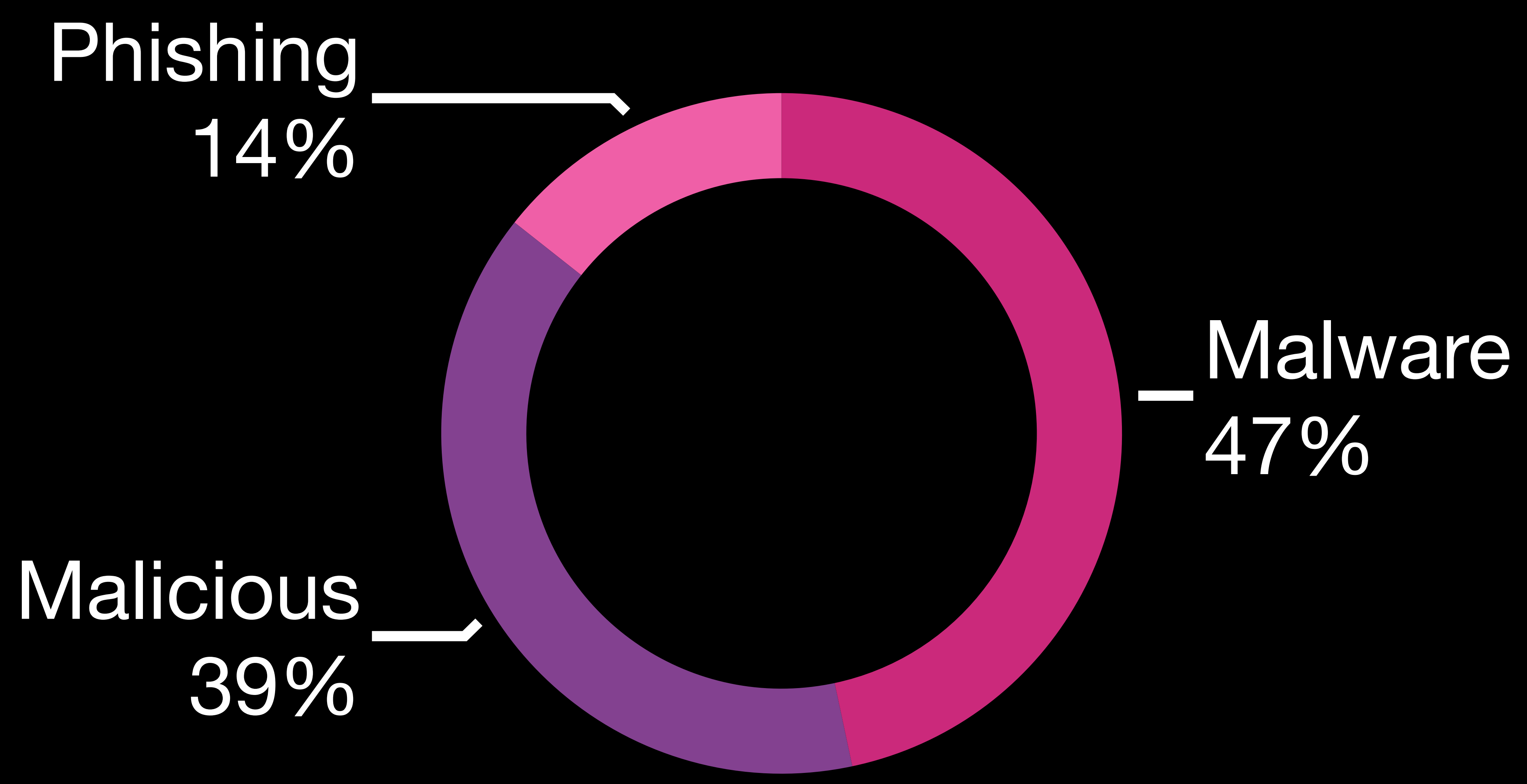
★ 67 PUP samples identified

★ Proxy programs are found for all 5 providers

★ 50 of them were flagged by anti-virus engines

Usage

- ★ For the 67 proxy programs, **5M traffic logs** were sampled to study usage
- ★ 9.36% of the destinations were reported to be malicious by VirusTotal

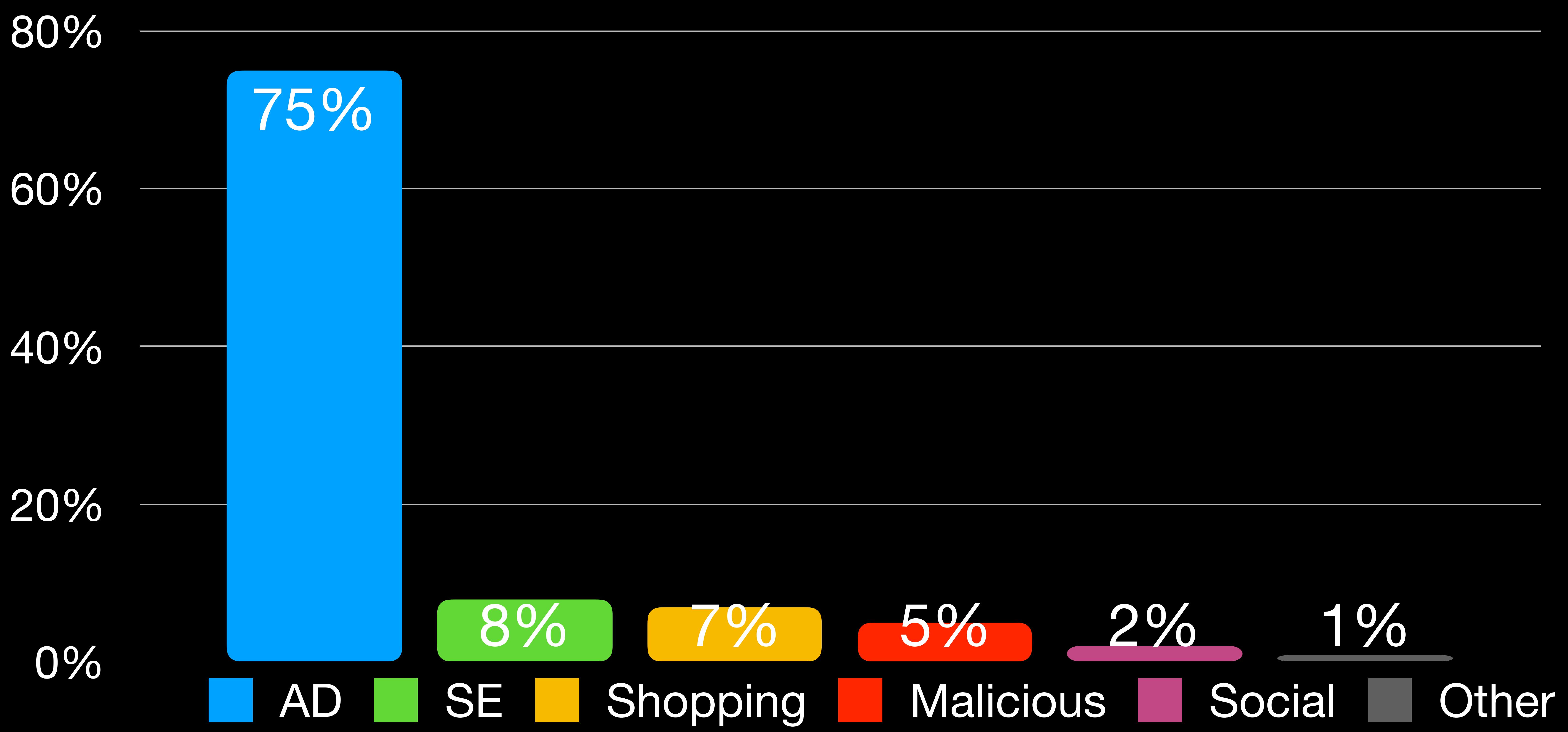


→

**ntkrnlpa.cn,
gwf-bd.com,
fadergolf.com,
www.2345jiasu.com,
www.pf11.com,**

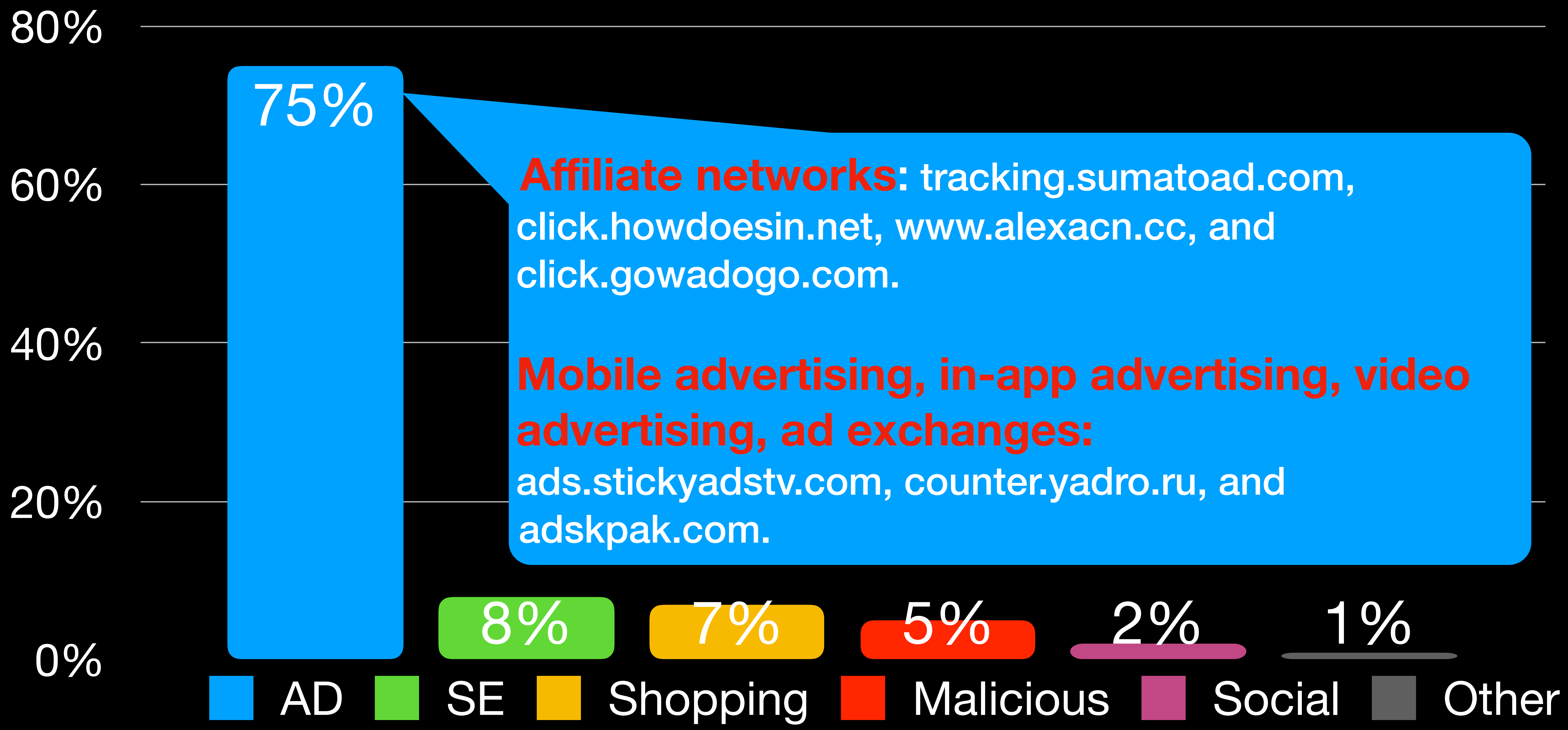
Usage

- For the 67 proxy programs, **5M traffic logs** were sampled to study usage
- 9.36% of the destinations were reported to be malicious by VirusTotal
- Top 1000 traffic destinations were manually studied.



Usage

- For the 67 proxy programs, **5M traffic logs** were sampled to study usage
- 9.36% of the destinations were reported to be malicious by VirusTotal
- Top 1000 traffic destinations were manually studied.

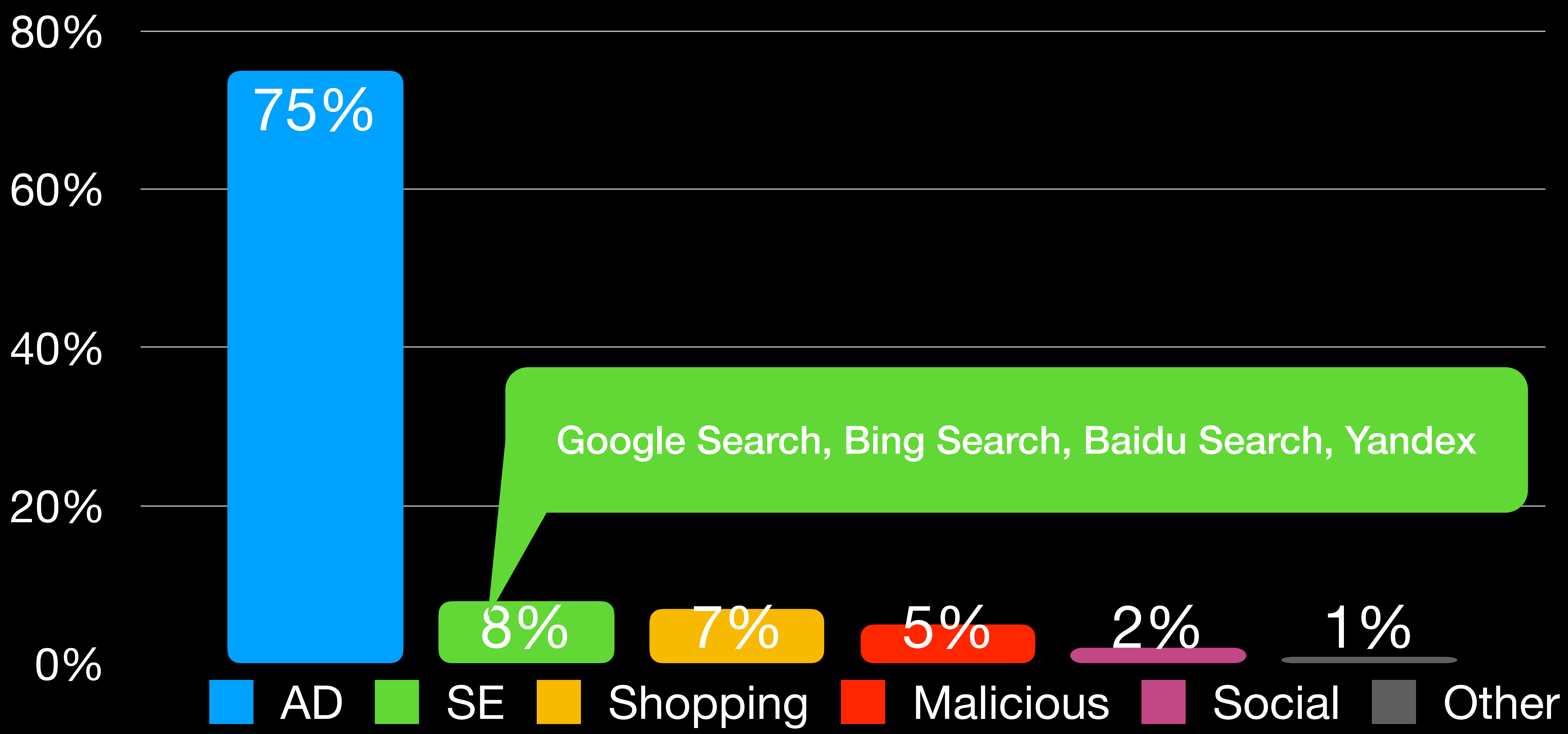


Affiliate networks: tracking.sumatoad.com, click.howdoesin.net, www.alexacn.cc, and click.gowadogo.com.

Mobile advertising, in-app advertising, video advertising, ad exchanges: ads.stickyadstv.com, counter.yadro.ru, and adskpak.com.

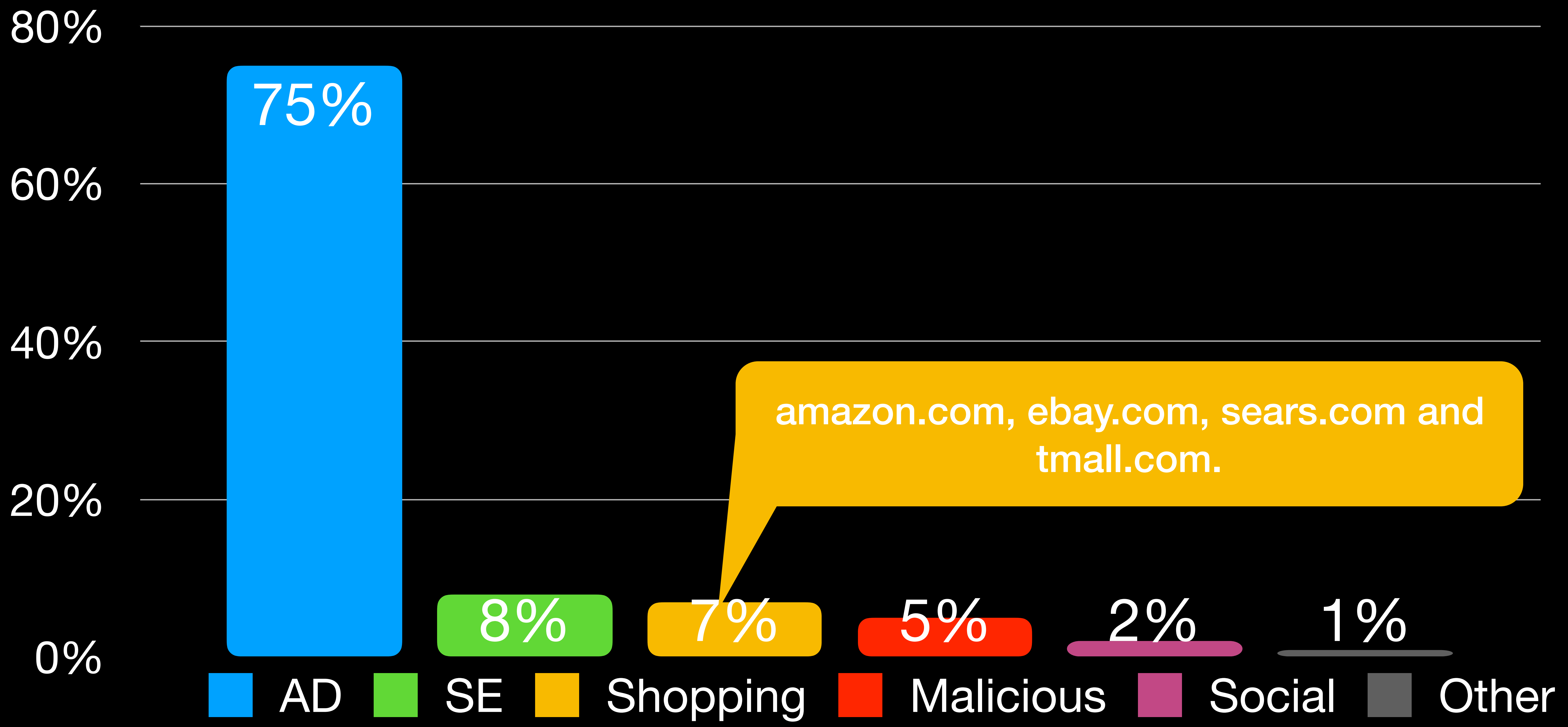
Usage

- For the 67 proxy programs, 5M traffic logs were sampled to study usage
- 9.36% of the destinations were reported to be malicious by VirusTotal
- Top 1000 traffic destinations were manually studied.



Usage

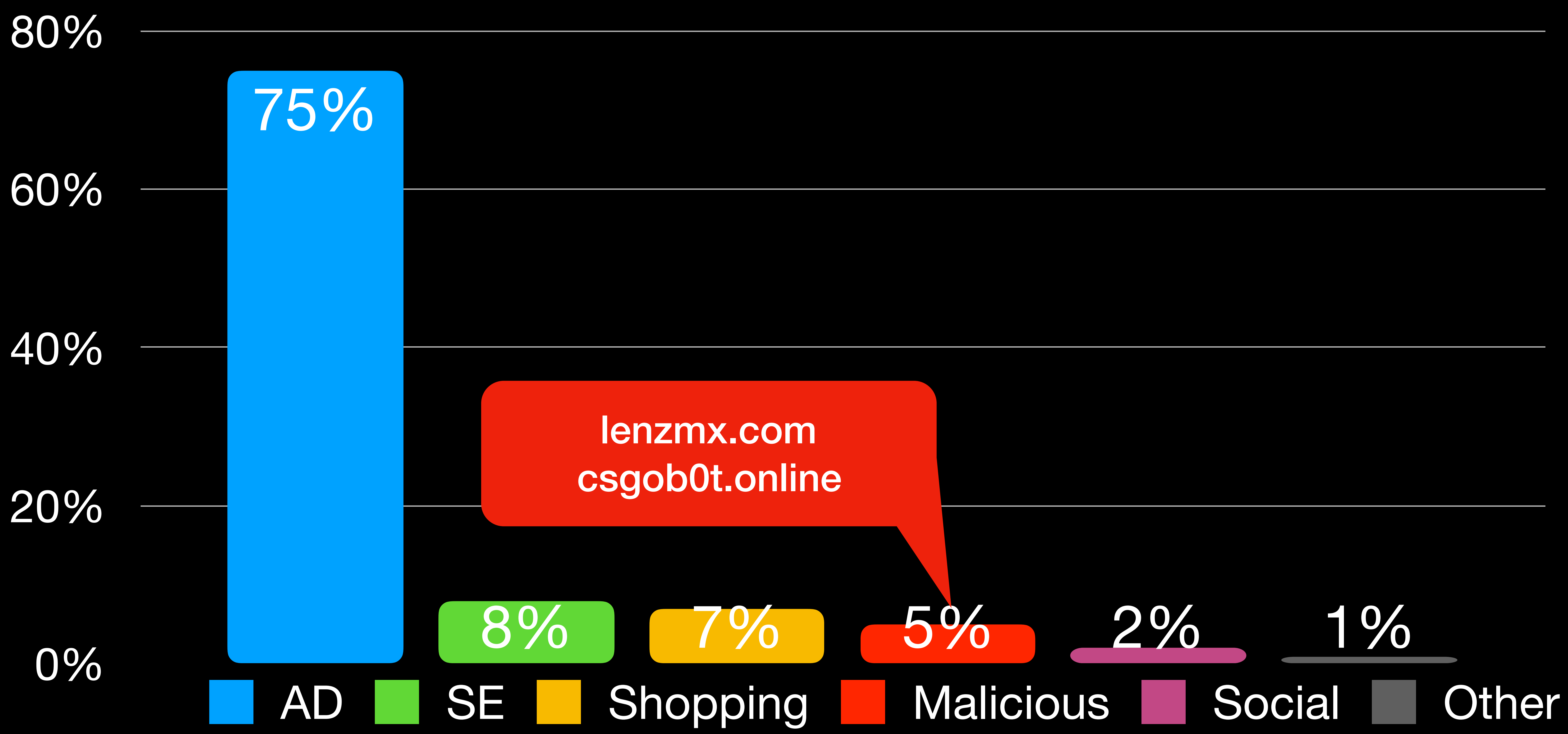
- For the 67 proxy programs, 5M traffic logs were sampled to study usage
- 9.36% of the destinations were reported to be malicious by VirusTotal
- Top 1000 traffic destinations were manually studied.



amazon.com, ebay.com, sears.com and tmall.com.

Usage

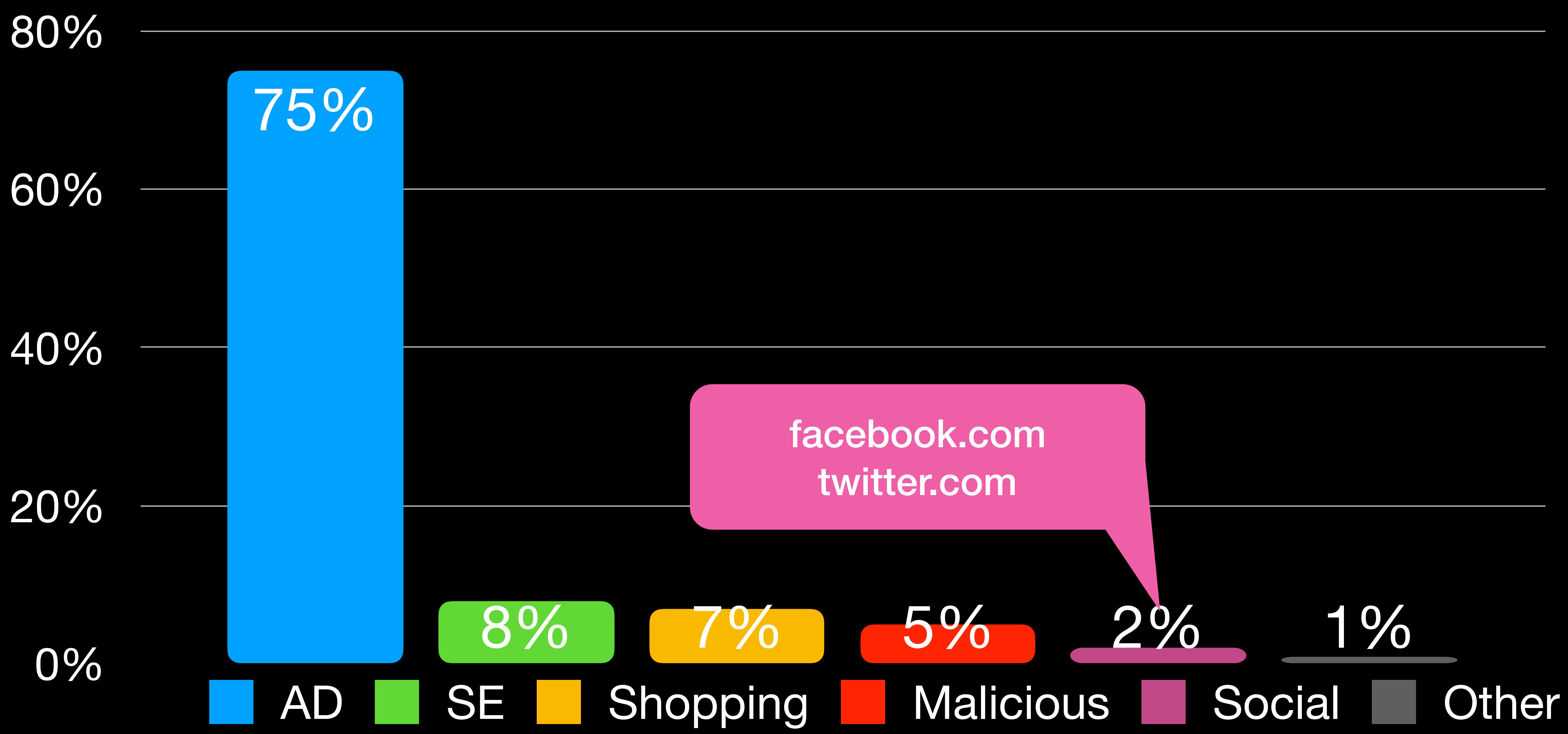
- ★ For the 67 proxy programs, 5M traffic logs were sampled to study usage
- ★ 9.36% of the destinations were reported to be malicious by VirusTotal
- ★ Top 1000 traffic destinations were manually studied.



lenzmx.com
csgob0t.online

Usage

- For the 67 proxy programs, **5M traffic logs** were sampled to study usage
- 9.36% of the destinations were reported to be malicious by VirusTotal
- Top 1000 traffic destinations were manually studied.



Misc. Findings

Connection between proxy providers

Risk to the local network

Long-tailed distribution

	Proxies Online	Geosurf	IAPS Security	Luminati	ProxyRack
Proxies Online		12.5%	0%	0.06%	0.09%
Geosurf	36.3%		0%	0.23%	1.7%
IAPS Security	0%	0%		66%	0.07%
Luminati	0.02%	0.02%	0.07%		0.04%
ProxyRack	0.14%	0.86%	0%	0.2%	

- ★ Proxies Online and Geosurf are the same proxy provider
- ★ IAPS Security is some kind of reseller for Luminati

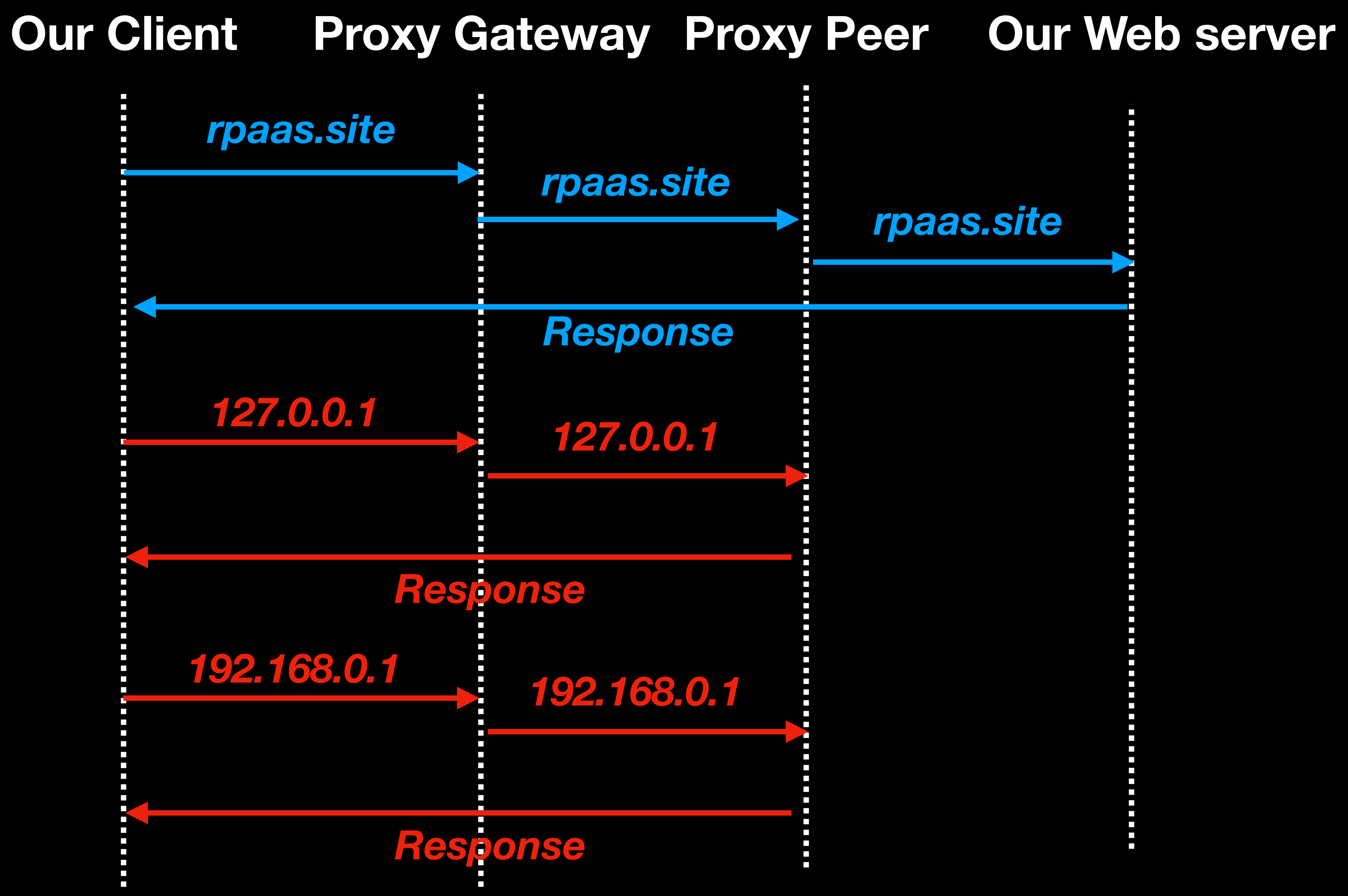
Misc. Findings

Connection between proxy providers

Risk to the local network

Long-tailed distribution

⛔ 3 out of 5 providers allow local traffic



Misc. Findings

Connection between proxy providers

Risk to the local network

Long-tailed distribution

Provider	Top Countries (%)		Top ASNs (%)	
Proxies Online	Indian USA Mexico	32.2 7.8 6.7	9829 8151 24560	8.1 5.4 4.9
Geosurf	India Brazil Mexico	27.9 9.2 9.1	8151 9829 55836	7.2 5.8 4.5
ProxyRack	Russia Indonesia Egypt	8.6 8.1 6.3	1797 8452 45595	5.3 4.7 4.0
Luminati	Turkey Ukraine UK	12.7 7.9 6.1	9121 25019 34984	8.5 1.8 1.8

Summary

Millions of residential IPs with high evasiveness

Potential threats to local network environments

Powerful infrastructure for online abuse activities

A prosperous ecosystem with higher prices and more service providers

Problematic recruitment: a mix of legitimate and suspicious channels

Promising and stealthy monetization channels for compromised devices

> *A lie that is half-truth is the darkest of all lies.*
>
> —Alfred Tennyson

Q&A

xmi@iu.edu

Data & Code: https://rpaas.site