Resident Evil: Understanding Residential IP Proxy as a Dark Service

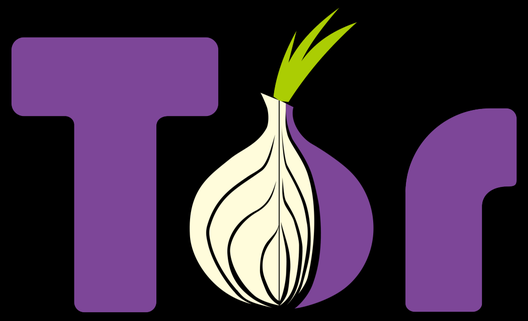
Xianghang Mi, Xuan Feng, Xiaojing Liao
Baojun Liu, XiaoFeng Wang, Feng Qian
Zhou Li, Sumayah Alrwais, Limin Sun , Ying Liu

INDIANA UNIVERSITY

Tsinghua University

King Saud University

Background: Traditional Web Proxies



HTTP/HTTPS
/SOCKS

✗ Exit nodes are constrained

✗ Exit nodes are distinguishable

✗ Exit nodes may be heavily abused

Vulnerable to 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
230,090

EUROPE
345,135

ASIA/OCEANIA
115,045

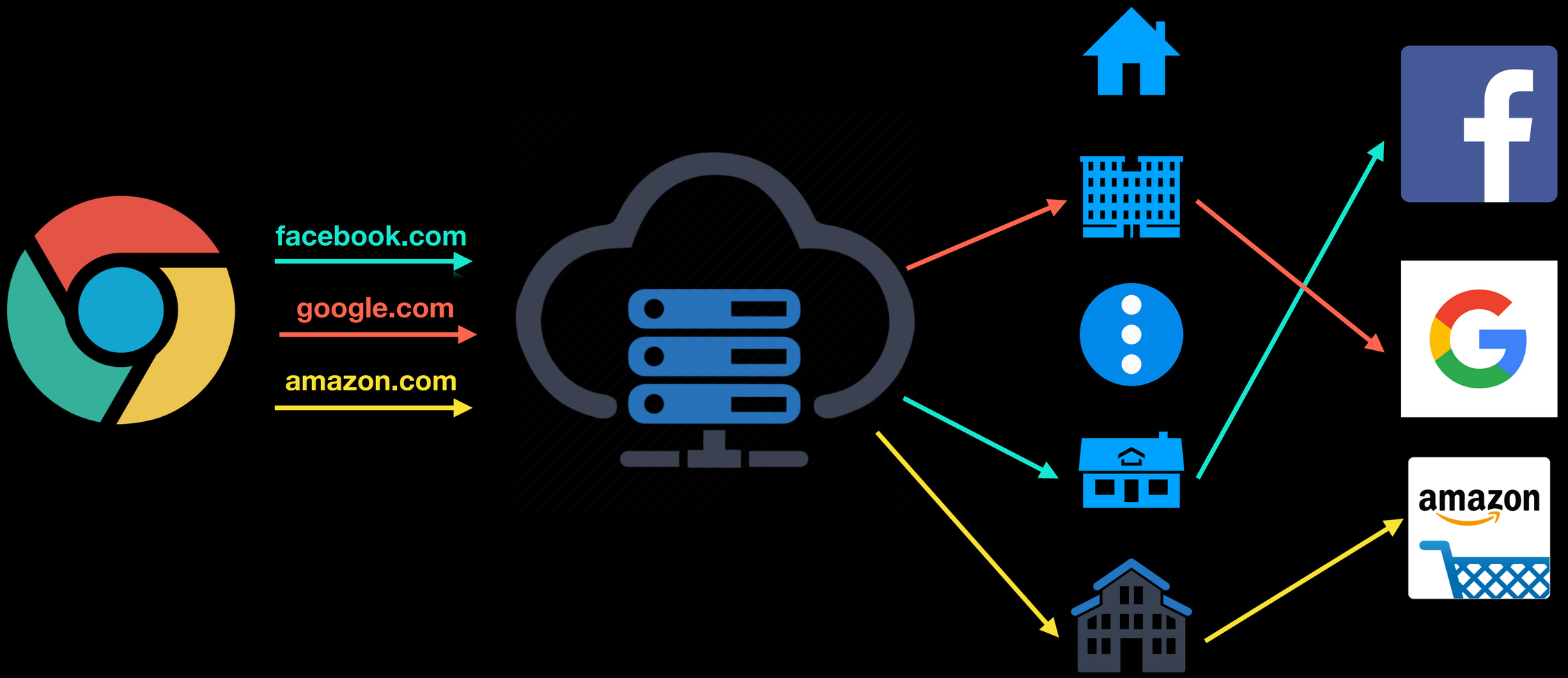
Background: Residential IP Proxy as a Service

- 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

Proxy Peers

Destinations

Service Overview: Scale

Each request is identified by a unique subdomain

Each request/response has payload encrypted and signed

Provider	Price	Payment	Date(s)
Proxies Online	$25/Gb	Paypal	07/06-11/24
Geosurf	$300/month	Paypal	09/17-10/22
ProxyRack	$40/month	Bitcoin	09/18-11/24
Luminati	$500/month	Paypal	09/25-11/01
IAPS Security	$500/month	Bitcoin	09/23-11/01

Service Overview: Scale

Each request is identified by a unique subdomain

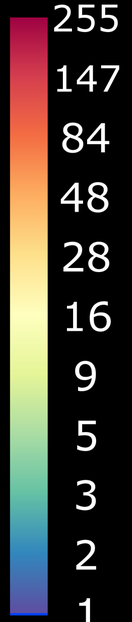
Each request/response has payload encrypted and signed

60+ millions of successful probes

6.2 millions of unique IPv4 addresses

238 countries/regions, 52K+ ISPs.

Service Overview: Distribution

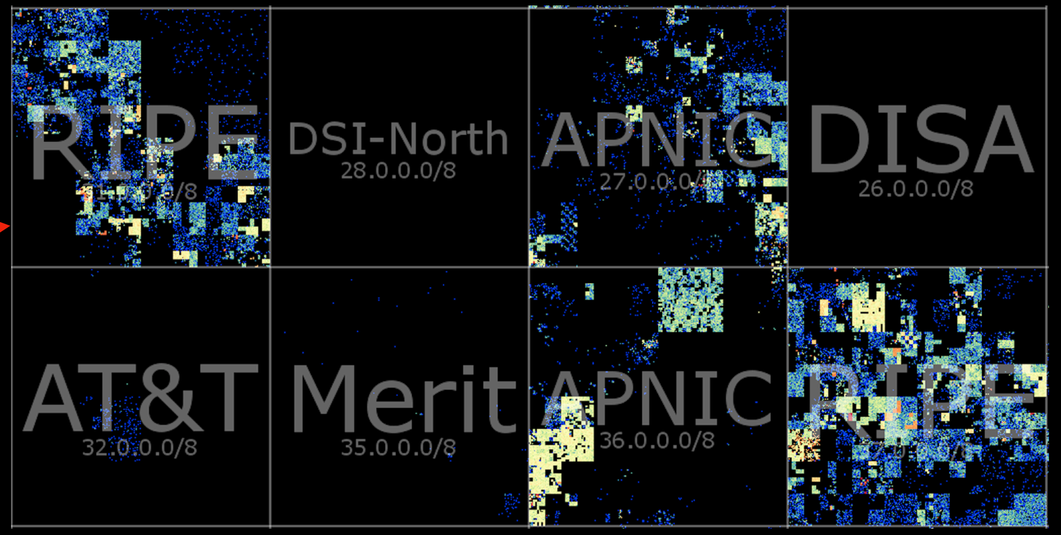


4096 * 4096 bitmap

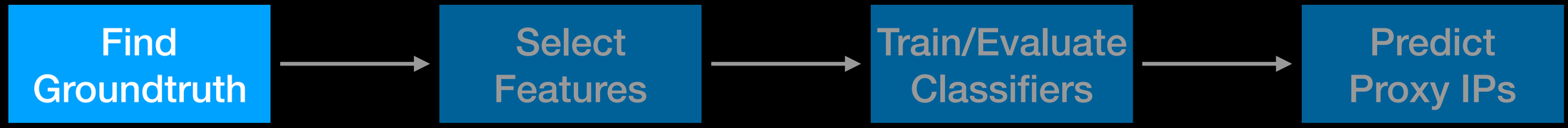
Each /24 IPv4 prefix is mapped to a pixel, using Hilbert curve of order 12

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

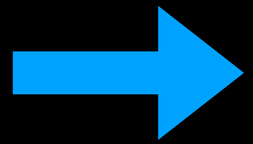
Residential or Not

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

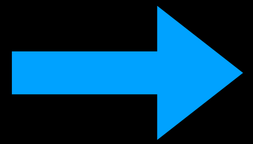
DNS Records & Historical IP Whois



Capture web activities

Capture network hierarchy

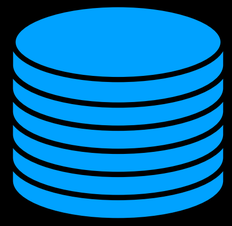
Capture evolution by time



35 features stand out for next step

Residential or Not

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



10K residential &
10K non-residential IPs

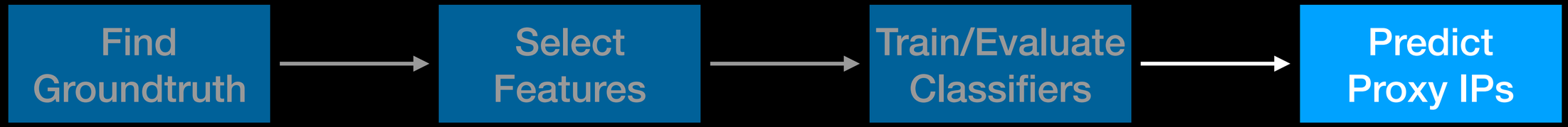
ML Classifier Training/Tuning

Random Forest Classifier

Recall: 97.12%

Precision: 95.61%

Residential or Not



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?

Publicly available proxy dataset

- Tor relays
- Free web proxies
- IP2Proxy LITE

Only 0.06% of 6.2M IPs

Evasiveness

Recognized as proxy?

Identified as malicious?

Publicly available IP threats

- Botnet bots
- Spamhaus EDROP
- Open Threat Exchanges

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	(%)	Device Vendor	Num	(%)
router	114,768	48.42	MikroTik	86,593	36.53
firewall	25,088	10.58	Huawei	37,545	15.84
WAP	24,470	10.32	BusyBox	18,337	7.74
gateway	22,003	9.28	Technicolor	16,866	7.12
broadband router	17,358	7.32	SonicWALL	14,122	5.96
webcam	13,024	5.49	Fortinet	9,190	3.88
security-misc	10,608	4.48	Dahua	6,258	2.64
DVR	4,249	1.79	ZyXEL	5,601	2.36
media device	2,589	1.09	AVM	5,272	2.22
storage-misc	1,988	0.84	Cyberoam	4,558	1.92

Recruitment

Identify legitimate recruitment programs

IP Profiling

Identify proxy programs

Traffic logs of Infiltration probes

Traffic logs of suspicious programs

- 67 different program 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
- 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

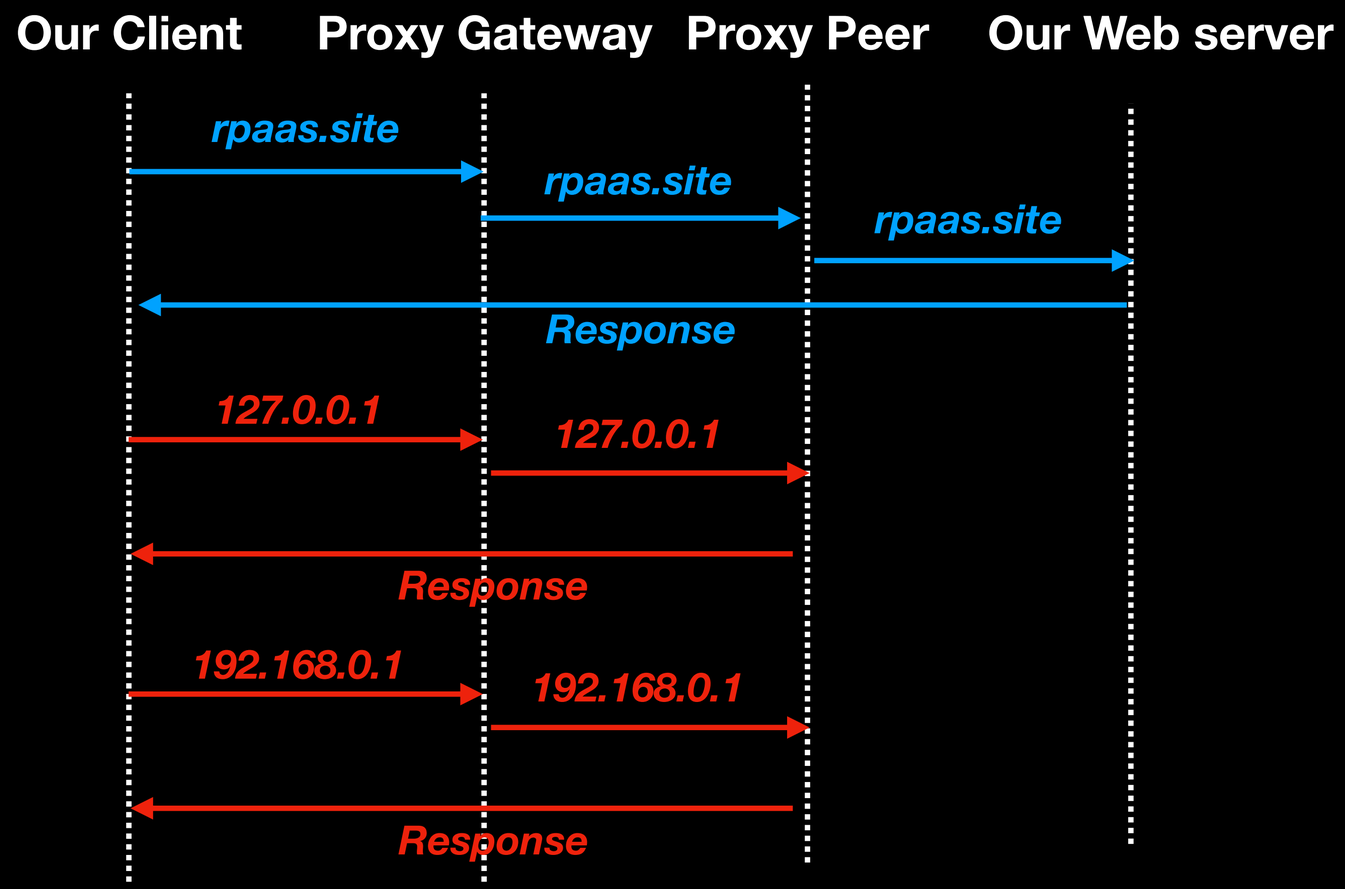
Misc. Findings

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 ISPs	%	Top ASNs	%
Proxies Online	India	32.2	BSNL	6.5	9829	8.1
	USA	7.8	Uninet S.A. de C.V.	5.2	8151	5.4
	Mexico	6.7	Deutsche Telekom AG	2.8	24560	4.9
Geosurf	India	27.9	Uninet S.A. de C.V.	6.9	8151	7.2
	Brazil	9.2	BSNL	4.7	9829	5.8
	Mexico	9.1	Deutsche Telekom AG	2.8	55836	4.5
ProxyRack	Russia	8.6	PT Telkom Indonesia	5.4	17974	5.3
	Indonesia	8.1	Pakistan Telecom	3.7	8452	4.7
	Egypt	6.3	Republican Unitary	3.3	45595	4.0
Luminati	Turkey	12.7	Turk Telekom	8.5	9121	8.5
	Ukraine	7.9	JSC Ukrtelecom	1.7	25019	1.8
	UK	6.1	BT	1.7	34984	1.8

Summary

Millions of residential IPs with high evasiveness

A prosperous ecosystem with higher prices and more service providers

Potential threats to local network environments

Problematic recruitment: a mix of legitimate and suspicious channels

Powerful infrastructure for online abuse activities

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