Countering Residential IP Proxies

Detection Techniques and Strategic Insights

Elisa Chiapponi
elisa.chiapponi@amadeus.com

31st Computer & Electronics Security Application Rendezvous (C&ESAR 2024)
Rennes, France
20th November 2024

amadeus
Global Security Operations

Who am I

- Security Researcher in the **Global Security Operations** of Amadeus
  - Protection of web domains linked to the travel industry
- Expertise in **Network** and **Application** Security
- Work based on current and Ph.D. research and collaborations

EURECOM Sophia Antipolis

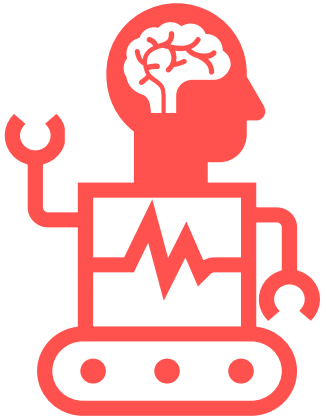
amadeus Global Security Operations

جامعة الملك عبدالله للعلوم والتقنية
King Abdullah University of Science and Technology

Resilient Computing and Cybersecurity Center

SORBONNE UNIVERSITÉ

UNIVERSITY OF TWENTE.



Sophisticated bots

Random user interactions generation

Mimic of **human interaction** to reach their goals

CAPTCHA solving

Sophisticated Bot Attacks – Functional Abuse

Web scraping

- Competitor Monitoring
- Content Reselling
- Illicit Aggregators

Requests | Purchases | Revenues | Slow connections

Denial Of Inventory

- Disrupt supply and demand
- Holding cheapest fares
- Impact revenues and operations
- Defeat the competition

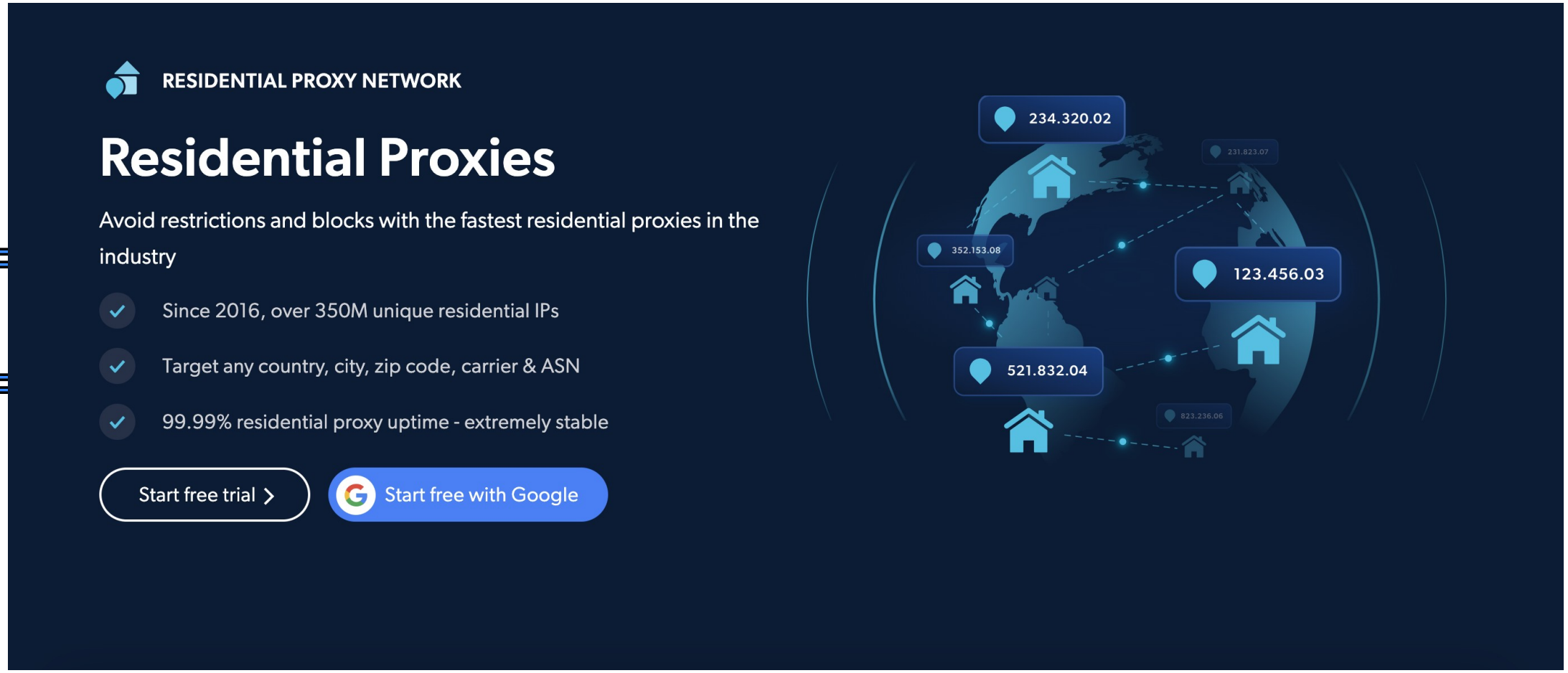
Requests | Purchases | Revenues | Reputation

SMS Pumping

- Impact revenues
- Generate revenue through network operators

Revenues | Reputation

Arms race



Battle Plan

1. Intel Gathering
   Know your adversary

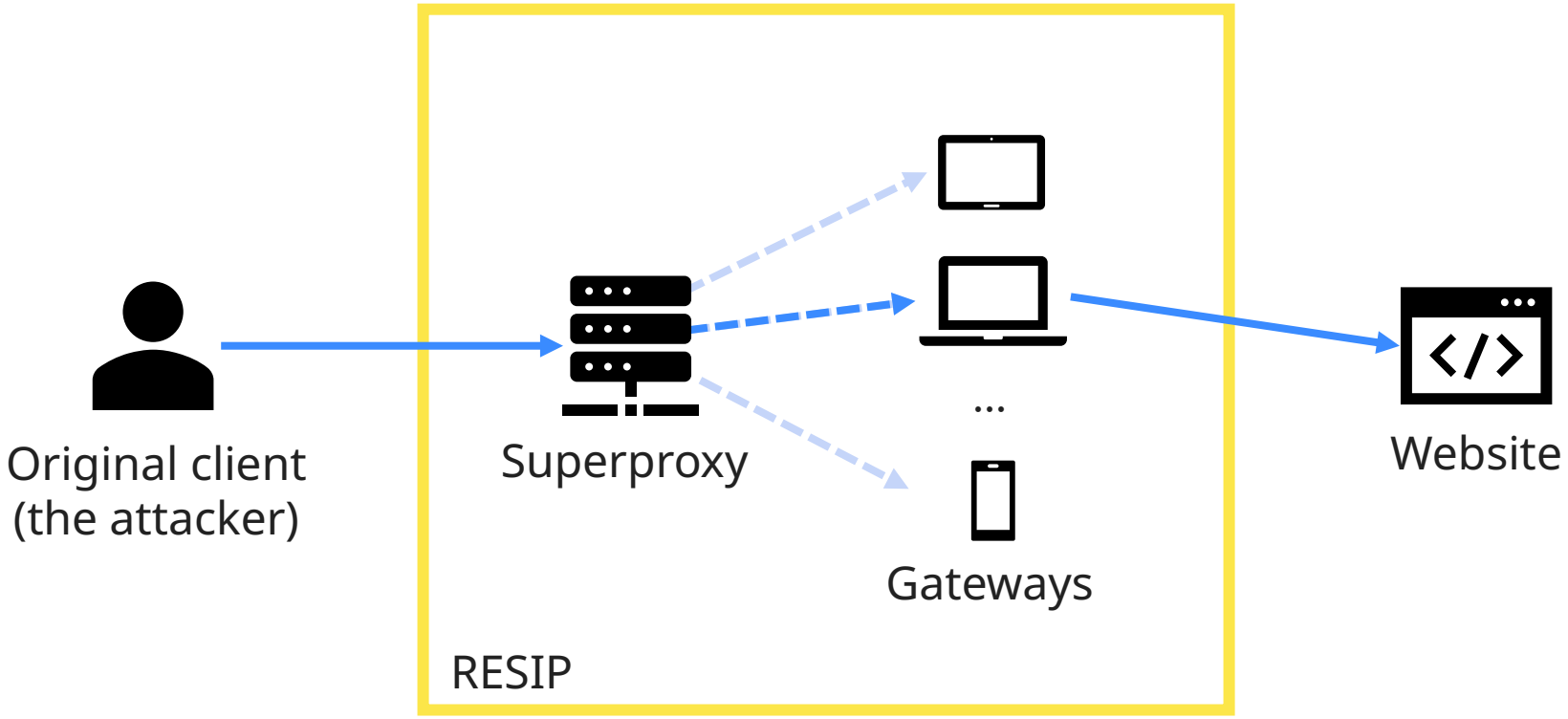
2. Defence Strategy & Combat Phase
   Identify vulnerabilities and exploit them

3. Debrief
   Lessons Learnt

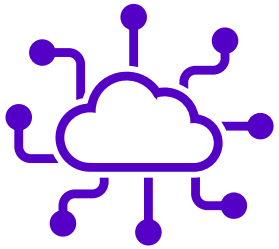
Residential IP Proxies (RESIPs)

- Large networks of **residential devices** (smartphones, laptops, tablets,…)
- Devices **owned** by genuine users who **share** their usage
- No application layer information about being proxied
  - **Indistinguishable** from the requests sent directly by the residential devices at this layer
  - **High probability of false positives** for the traditional server-side bot detection techniques
- Advanced bot traffic **heavily rely** on RESIPs

RESIP infrastructure



Advantages for the attacker



Tens of millions of residential IPs

No private distributed infrastructure

Automated services

Good reputation IPs

No direct traceability

Recruitment process

Free services (e.g. VPN)

Bandwidth payment

Mobile SDKs included by app developers

Infected devices (IoT)

External references:
- M. Frappier et al., Illegitimate residential proxy services: the case of 911.re and its IOCs, 2022.
- X. Mi et al., “Your Phone is My Proxy: Detecting and Understanding Mobile Proxy Networks,” in NDSS 2021.
- A. Vastel. “Ever wonder how proxy providers & BaaS providers obtain residential proxies?”, 2022.

Legitimate but…

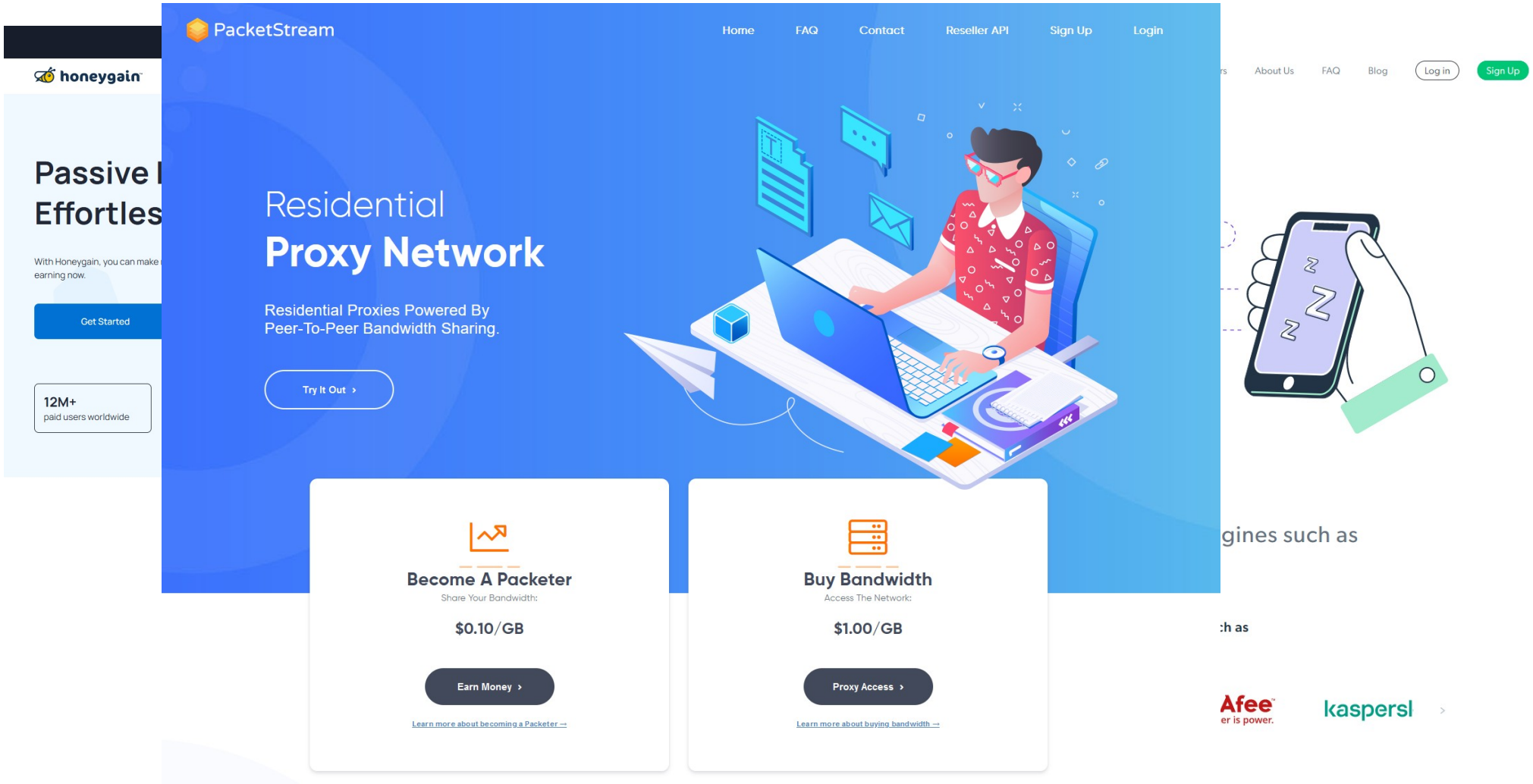
- **Shady** Device Recruitment

- IP addresses reputation and probing **showed**:
  - Credential and stuffing attacks
  - Social Media Spam
  - Fast Flux Proxies
  - Cryptojacking
  - …

- **More expensive** than common VPNs

External references:
- M. Frappier et al., Illegitimate residential proxy services: the case of 911.re and its IOCs, 2022.
- B. Krebs, The Rise of “Bulletproof” Residential Networks, 2019.
- X. Mi et al. Resident Evil: Understanding Residential IP Proxy as a Dark Service, IEEE S&P 2019.
- M. Yang et al., An Extensive Study of Residential Proxies in China. ACM SIGSAC CCS 2022.

Acting as a RESIP gateway



[1] E. Khan et al. A First Look at User-Installed Residential Proxies From a Network Operator's Perspective, CNSM 2024

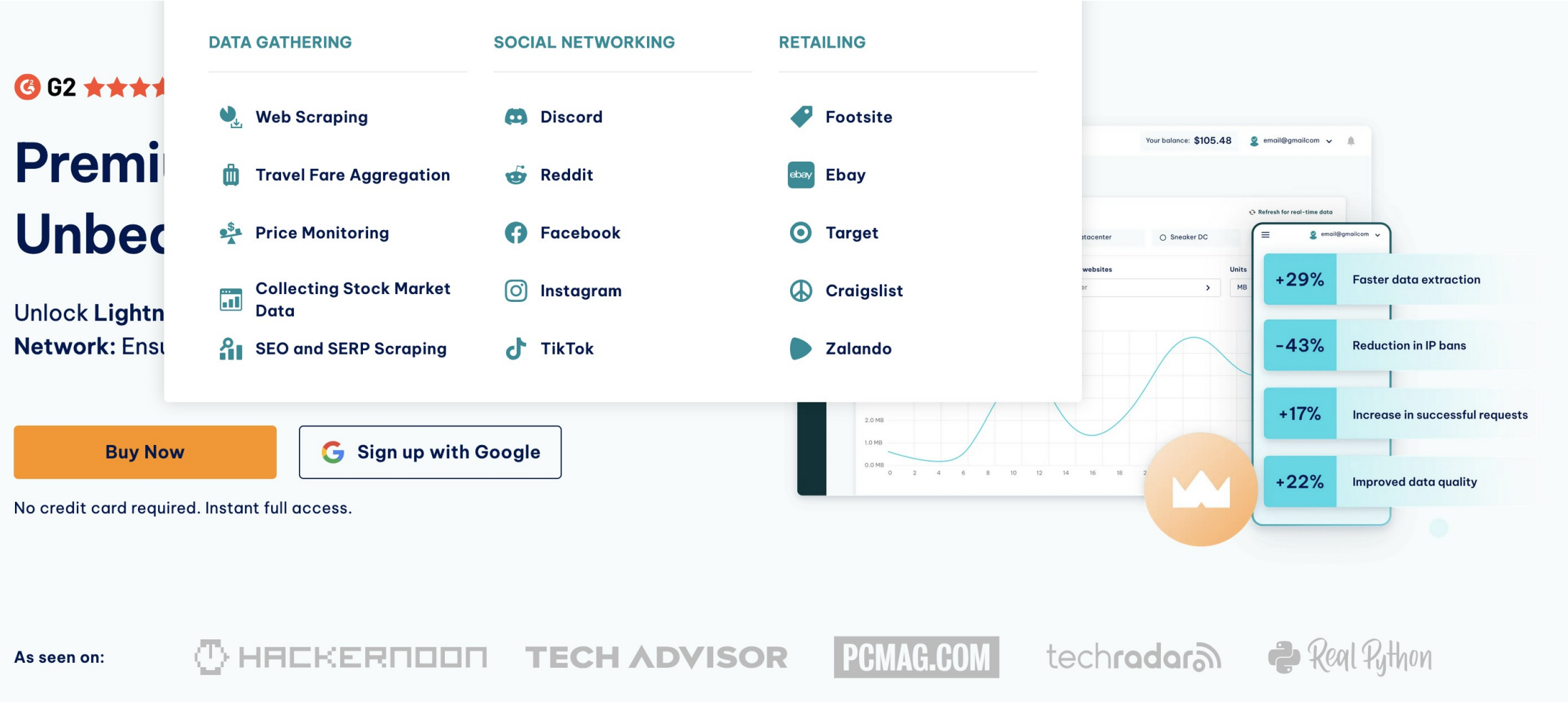
Setup

- .exe files to mimic an **average end-user**
- Examined **8 bandwidth providers**
- **Testbed** (University of Twente):
  - **Windows 11** VMs
  - **Dedicated IP** for each VM to prevent cross contamination
  - IP range classified as **residential**
  - **368GB** - 7.5 months
  - **PCAP** files collection
- Tranalyzer (network traffic analyzer) to **aggregate** PCAP files

External references:
- S. Burschka et al., Tranalyzer: Versatile high performance network traffic analyser, in IEEE SSCI 2016

[1] E. Khan et al. A First Look at User-Installed Residential Proxies From a Network Operator's Perspective, CNSM 2024

Encrypted traffic – First Look



[1] E. Khan et al. A First Look at User-Installed Residential Proxies From a Network Operator's Perspective, CNSM 2024

Online dating apps

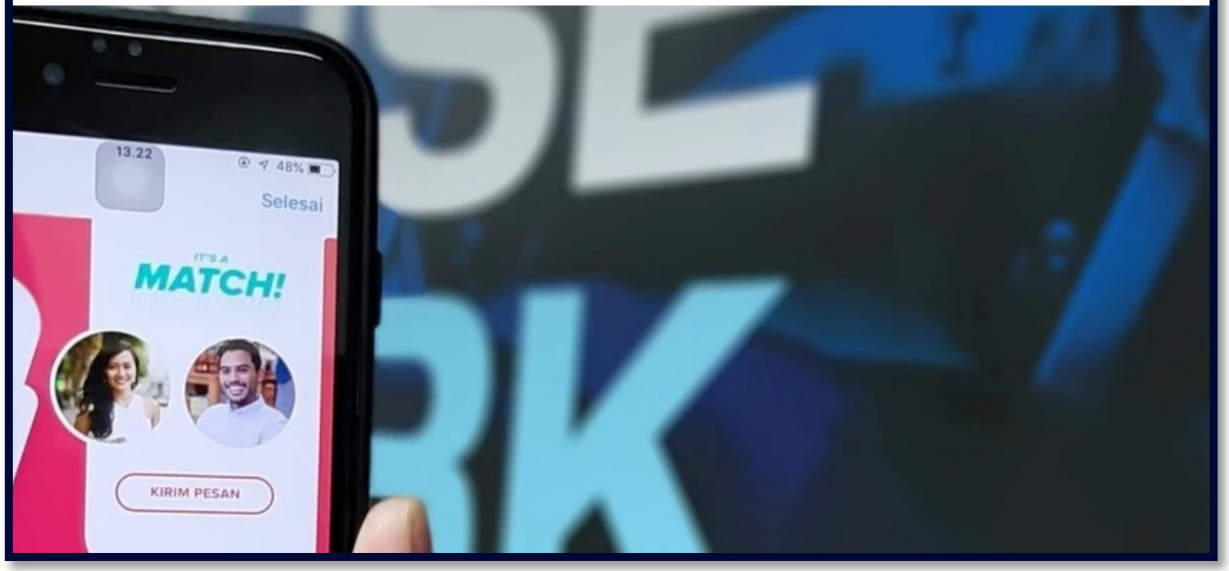
How to Catch a Cheater on Tinder: 8 Signs that Your Match may be a Tinder Bot

WORLD ECONOMIC FORUM

CYBERSECURITY

Tinder Swindler: How 'romance fraud' became a multi-billion dollar cybercrime

May 24, 2024

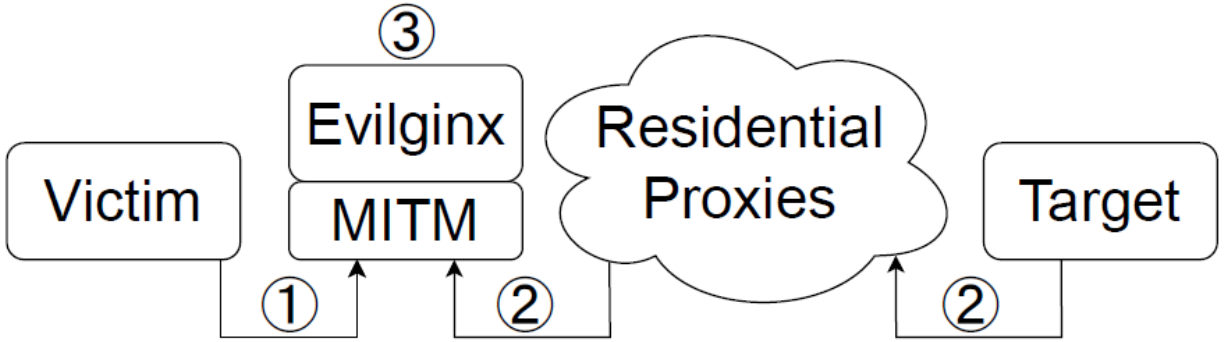


External references:
- https://medium.com/@jennifer.pearson83jp/hot-or-bot-8-signs-your-match-is-a-tinder-bot-b32641a8ff2d
- https://www.weforum.org/stories/2024/05/tinder-swindler-romance-fraud-cybercrime-radio-davos/

[1] E. Khan et al. A First Look at User-Installed Residential Proxies From a Network Operator's Perspective, CNSM 2024

Evilginx

- Man-in-the-middle reverse-proxy attack framework used for **phishing** account credentials along with session cookies
- Evilginx **JA4+ Network fingerprint** found in the collected flows
- **Server Name Indication** of the flow targets: account.booking.com, paypal.com, and m → **Websites not commonly scraped**

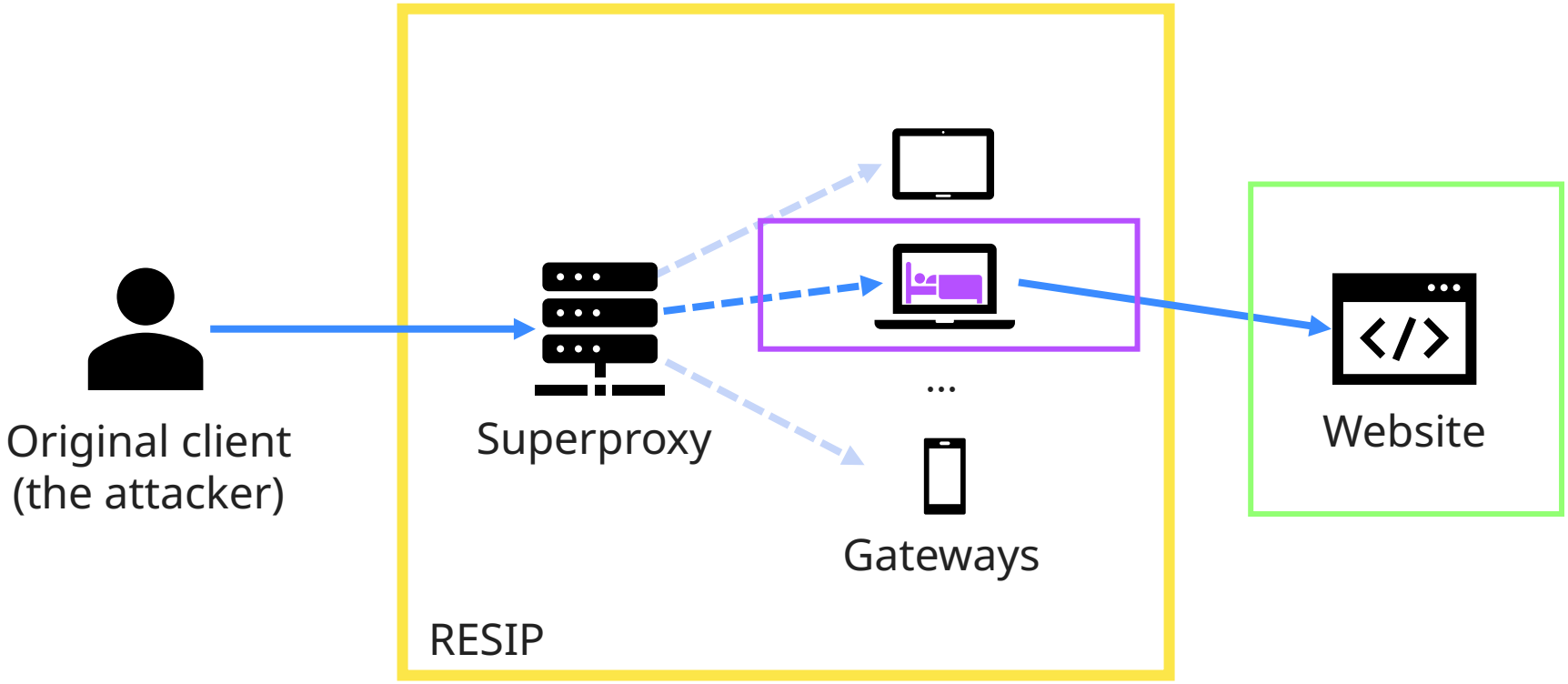


- **Strong indication** of RESIP usage for phishing

External references:
- https://github.com/FoxIO-LLC/ja4

[1] E. Khan et al. A First Look at User-Installed Residential Proxies From a Network Operator's Perspective, CNSM 2024

The other end of the tunnel



[1] E. Khan et al. A First Look at User-Installed Residential Proxies From a Network Operator's Perspective, CNSM 2024

The other end of the tunnel

- **Confirmation** of sophisticated bot attacks for **web scraping** performed through RESIPs
- **Only one** gateway per provider but…
- Analysis of Denial of Inventory IPs in **Spur** (RESIP IP reputation DB)
  - IPs from **7 out 8 providers*** involved in these attacks
- RESIP are **not used only for web scraping** campaigns

*One provider was not identified by Spur at the time of the analysis

External references:
- https://spur.us/

[1] E. Khan et al. A First Look at User-Installed Residential Proxies From a Network Operator's Perspective, CNSM 2024

Battle Plan

1. Intel Gathering

   Know your adversary

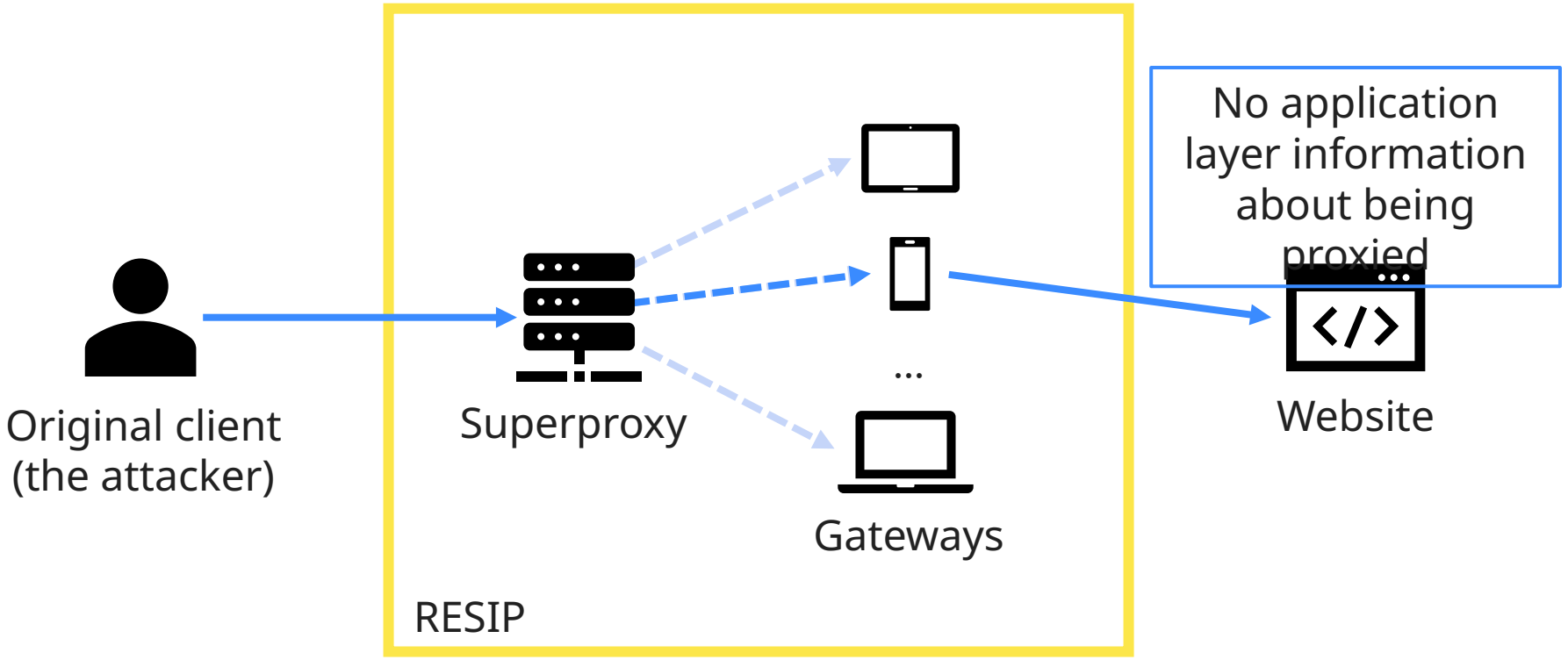
2. Defence Strategy & Combat Phase

   Identify vulnerabilities and exploit them

3. Debrief

   Lessons Learnt

RESIP infrastructure



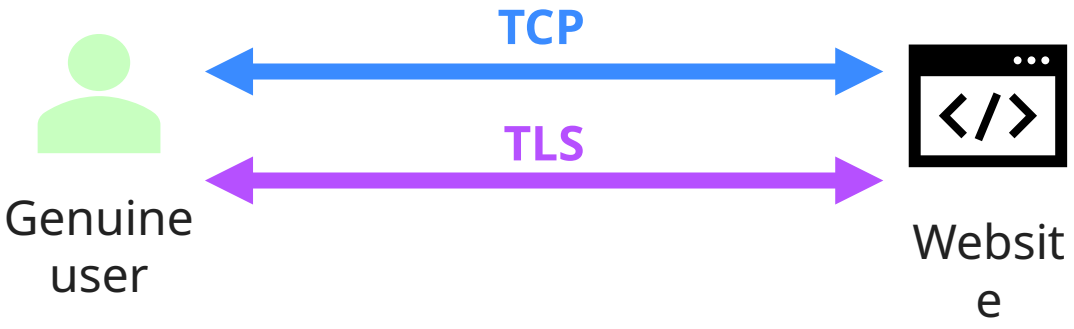
Both direct and RESIP connections are indistinguishable at the application layer **but** are there differences at the **transport layer?**

- **Round Trip Times at the TCP and TLS layers**
- **Retransmission Protocol**

Both direct and RESIP connections are indistinguishable at the application layer **but** are there differences at the **transport layer?**

- **Round Trip Times at the TCP and TLS layers**
- **Retransmission Protocol**

Direct connections

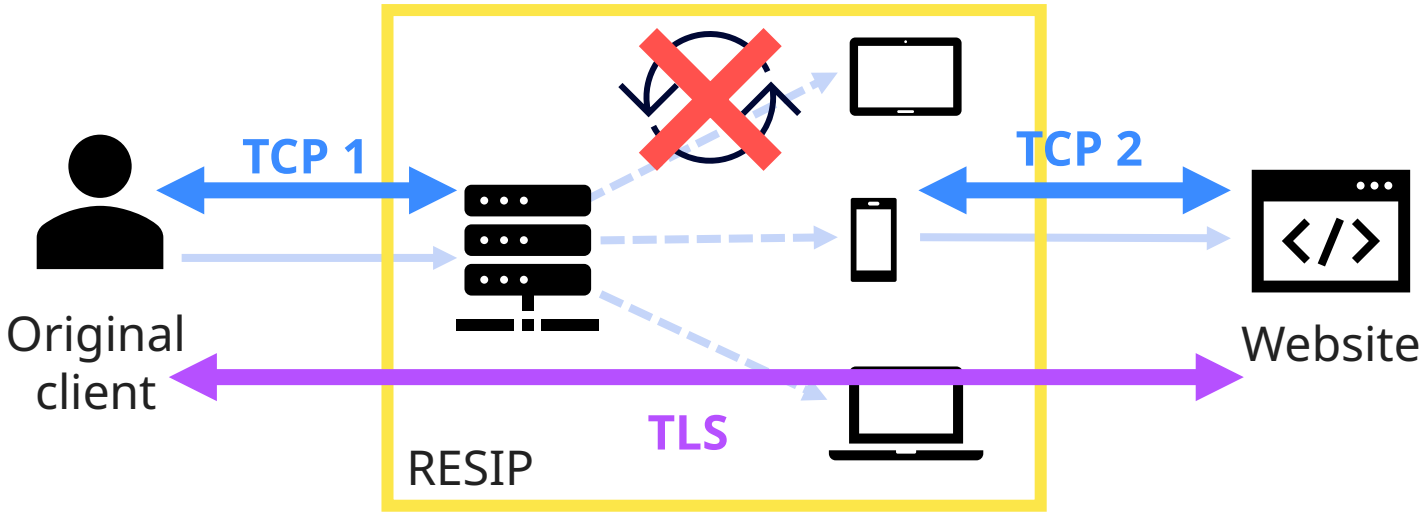


TCP: Transmission Control Protocol

TLS: Transport Layer Security

[2] E. Chiapponi et al. (2022). BADPASS: Bots taking ADvantage of Proxy AS a Service. In ISPEC 2022.

RESIP connection



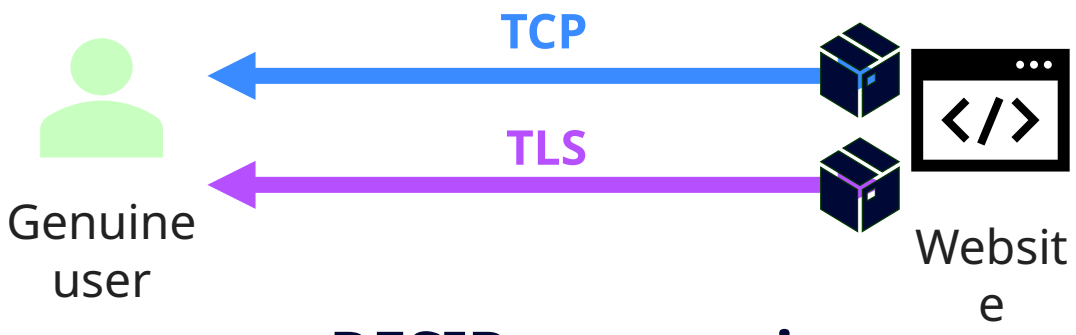
TCP: Transmission Control Protocol

TLS: Transport Layer Security

[2] E. Chiapponi et al. (2022). BADPASS: Bots taking ADvantage of Proxy AS a Service. In ISPEC 2022.

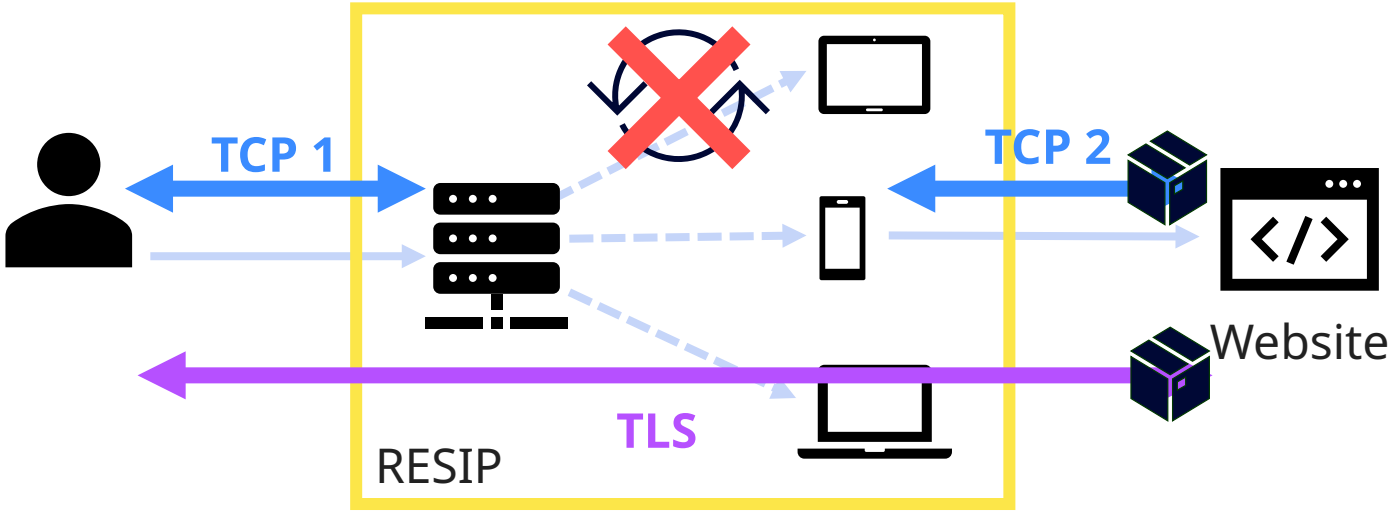
How can we check it at the server side?

Direct connection



$RTT_{TLS} \sim RTT_{TCP}$
for direct connections

RESIP connection



$RTT_{TLS} >> RTT_{TCP}$
for RESIP connections

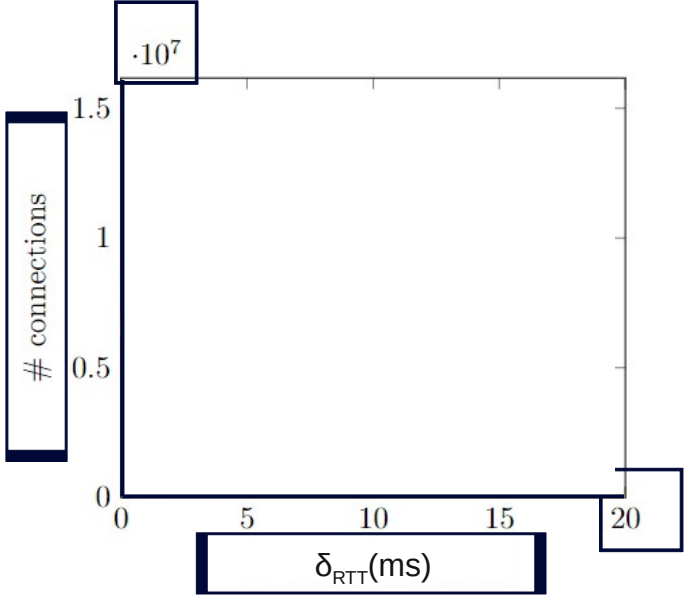
[2] E. Chiapponi et al. (2022). BADPASS: Bots taking ADvantage of Proxy AS a Service. In ISPEC 2022.

From theory to practice

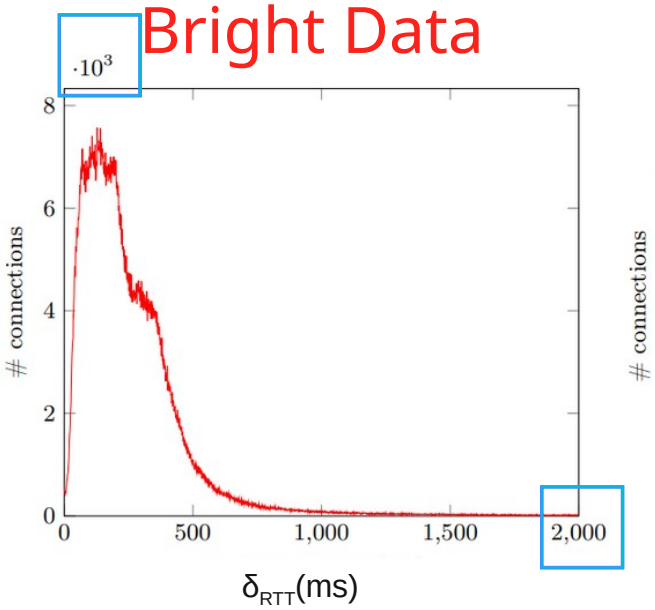
- **2** client/server machines in **11** locations all over the world
- **4** RESIP providers
- **4 months** experiment
- **92M+** connections

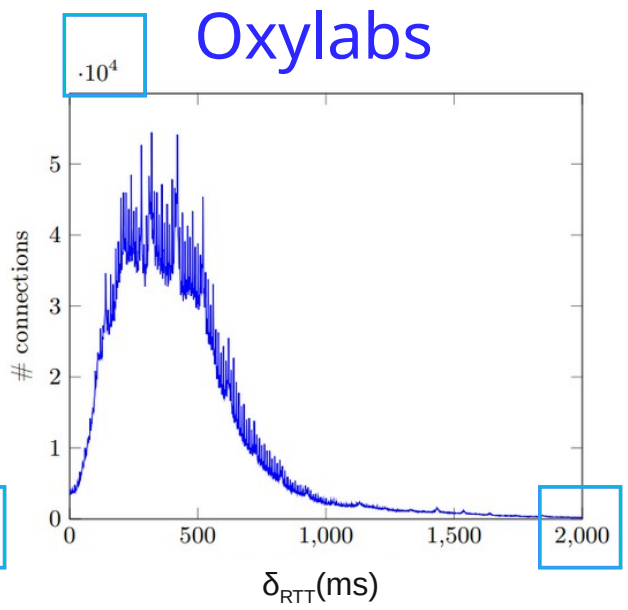
[2] E. Chiapponi et al. (2022). BADPASS: Bots taking ADvantage of Proxy AS a Service. In ISPEC 2022.

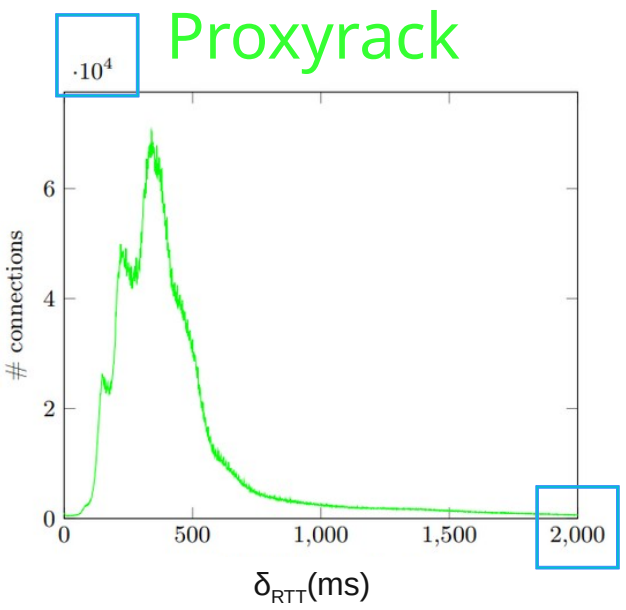
Direct Connections

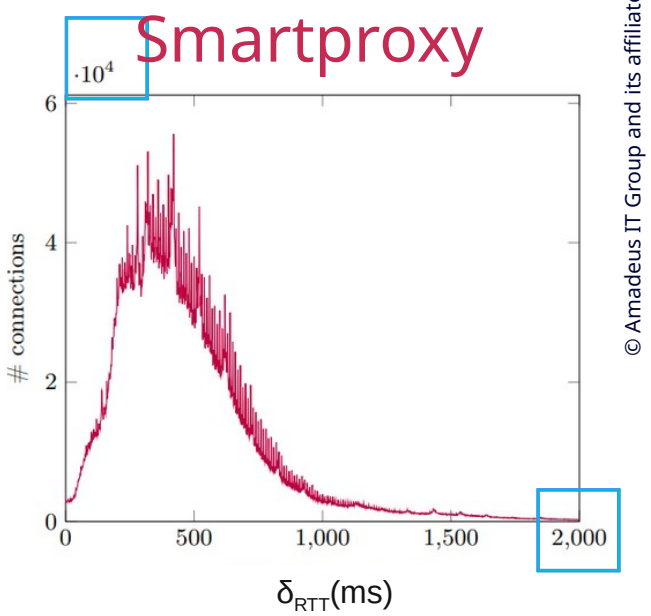


$$\delta_{RTT} = RTT_{TLS} - RTT_{TCP}$$









RESIP Connections

[2] E. Chiapponi et al. (2022). BADPASS: Bots taking ADvantage of Proxy AS a Service. In ISPEC 2022.

RTT Detection

- **δ_{RTT} > 50ms** → RESIP Connection

- **Possible** impacts on the detection technique:
  - Packet speed
  - TLS version

  } **No impact**

  - Client processing time → Browsers and hotspot increase the difference but **below threshold** for direct connections
  - Network delays
  - Geographic location of parties

  } **Small increase** in false negatives

[3] E. Chiapponi et al. (2023). Towards Detecting and Geolocalizing Web Scrapers with Round Trip Time Measurements. In TMA 2023.

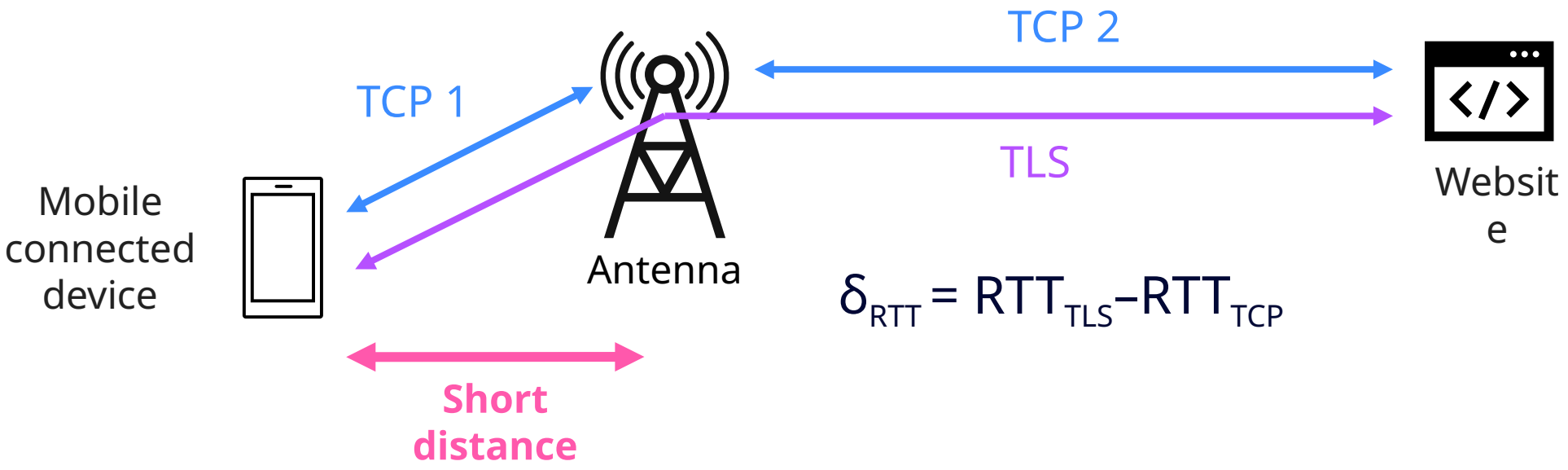
Implementation in real-world

- **Different** from current anti-bot techniques (JS insertion, parameters clustering)
- Amadeus convinced an anti-bot **third party** company to implement the technique
- Analysts currently **using the feature** to detect RESIP campaigns in combination with other parameters

[4] E. Chiapponi et al. (2023). Poster: The Impact of the Client Environment on Residential IP Proxies Detection. In IMC 2023

Mobile connections false positives

- Mobile **TCP Terminating** Proxies



$$\delta_{RTT} = RTT_{TLS} - RTT_{TCP}$$

- δ_{RTT} is **smaller** than RESIP one
  - Confirmation from semi-controlled and real-world data collections

[4] E. Chiapponi et al. (2023). Poster: The Impact of the Client Environment on Residential IP Proxies Detection. In IMC 2023

Detection evasion

- **Downgrading** to HTTP
  - Downgrading **not allowed** + possible generalization
- **Breaking TLS** at the RESIP
  - Technically feasible **BUT**
    - Clients need to accept root certificate from the gateway
    - Gateways devices have access to the content
    - Increased workload for gateways
- **Delaying TCP** packets at the gateway
  - **Unfeasible** since RESIP do not control directly the gateways

[2] E. Chiapponi et al. (2022). BADPASS: Bots taking ADvantage of Proxy AS a Service. In ISPEC 2022.

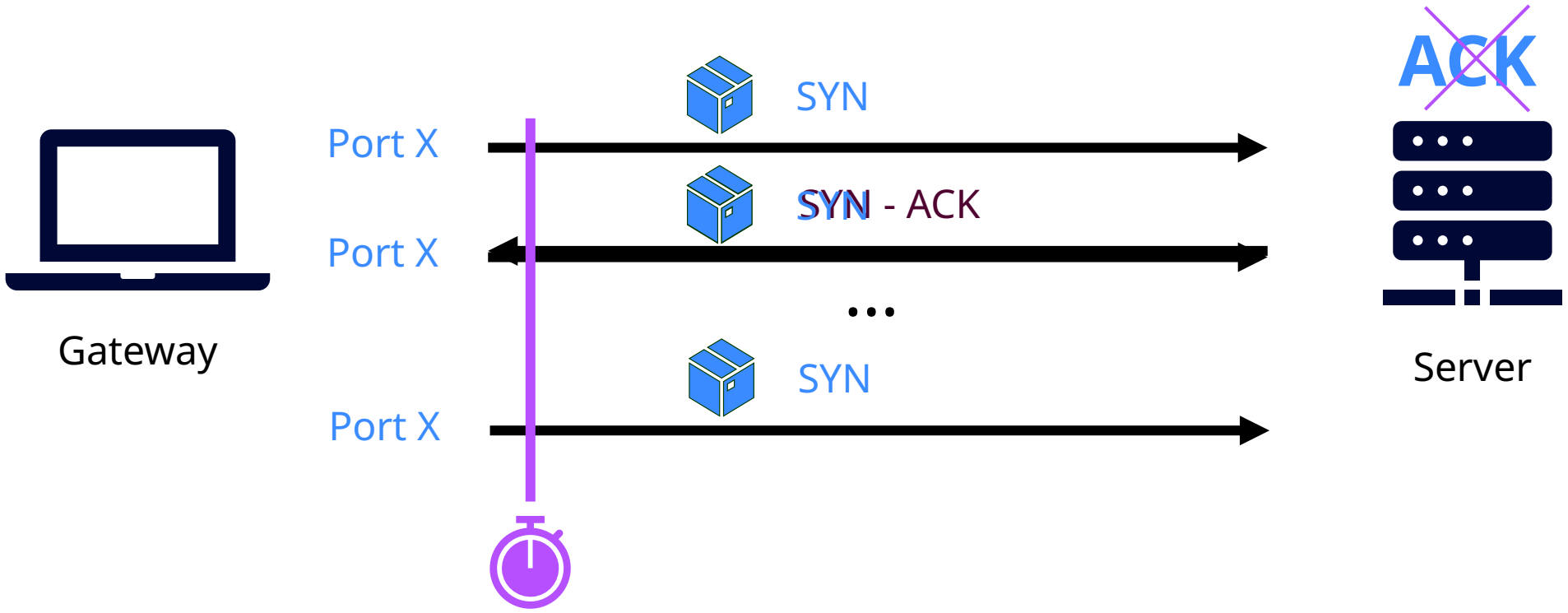
Both direct and RESIP connections are indistinguishable at the application layer **but** are there differences at the **transport layer?**

- **Round Trip Times at the TCP and TLS layers**
- **Retransmission Protocol**

Both direct and RESIP connections are indistinguishable at the application layer **but** are there differences at the **transport layer?**

- **Round Trip Times at the TCP and TLS layers**
- **Retransmission Protocol**

Normal retransmission



[5] E. Chiapponi et al. (2023). Inside Residential IP Proxies: Lessons Learned from Large Measurement Campaigns. In WTMC 2023.

RESIP retransmission (specific providers)

The application layer (not the transport one) handles the sending of packets and the timeouts

[5] E. Chiapponi et al. (2023). Inside Residential IP Proxies: Lessons Learned from Large Measurement Campaigns. In WTMC 2023.

Retransmission detection

- How?
  - **Delay** SYN-ACK packets
  - Check if multiple packets from **same IP and different ports**

- **Detection**: we identify RESIP connections (of specific providers)

- **Attribution**: we identify specific RESIP provider sending requests

- **Mitigation**: if the connection is detected as RESIP the server does not send any SYN-ACK packet. The RESIP waste resources

[5] E. Chiapponi et al. (2023). Inside Residential IP Proxies: Lessons Learned from Large Measurement Campaigns. In WTMC 2023.

Limitations

- **Only** for specific RESIP providers
- The wait time **can degrade** the user experience (1-1,5s) → Usage in **combination** with other techniques
- **Evasion** changing retransmission protocol → Possible **loss in efficiency** and **costs**
- Corner cases for **false** positives → Possible but **unlikely** to happen on a large scale

[5] E. Chiapponi et al. (2023). Inside Residential IP Proxies: Lessons Learned from Large Measurement Campaigns. In WTMC 2023.

Battle Plan

1. Intel Gathering

   Know your adversary

2. Defence Strategy & Combat Phase

   Identify vulnerabilities and exploit them

3. Debrief

   Lessons Learnt

Lessons learnt

- RESIPs appear to be largely used for **non advertised and malicious activities**
  - Online dating apps **frauds**
  - **Phishing**
  - **Denial of Inventory** attacks
  - … (this was just a **first look** of encrypted traffic of a **single gateway** of each network)

- We can used **transport layer differences** between RESIP and direct connections to **detect RESIP** at the server side
  - **Round Trip Times** at the TCP and TLS layers
  - **Retransmission Protocols**

Thank you for your attention!

More questions?
elisa.chiapponi@amadeus.com
or here in person

Presentation based on:

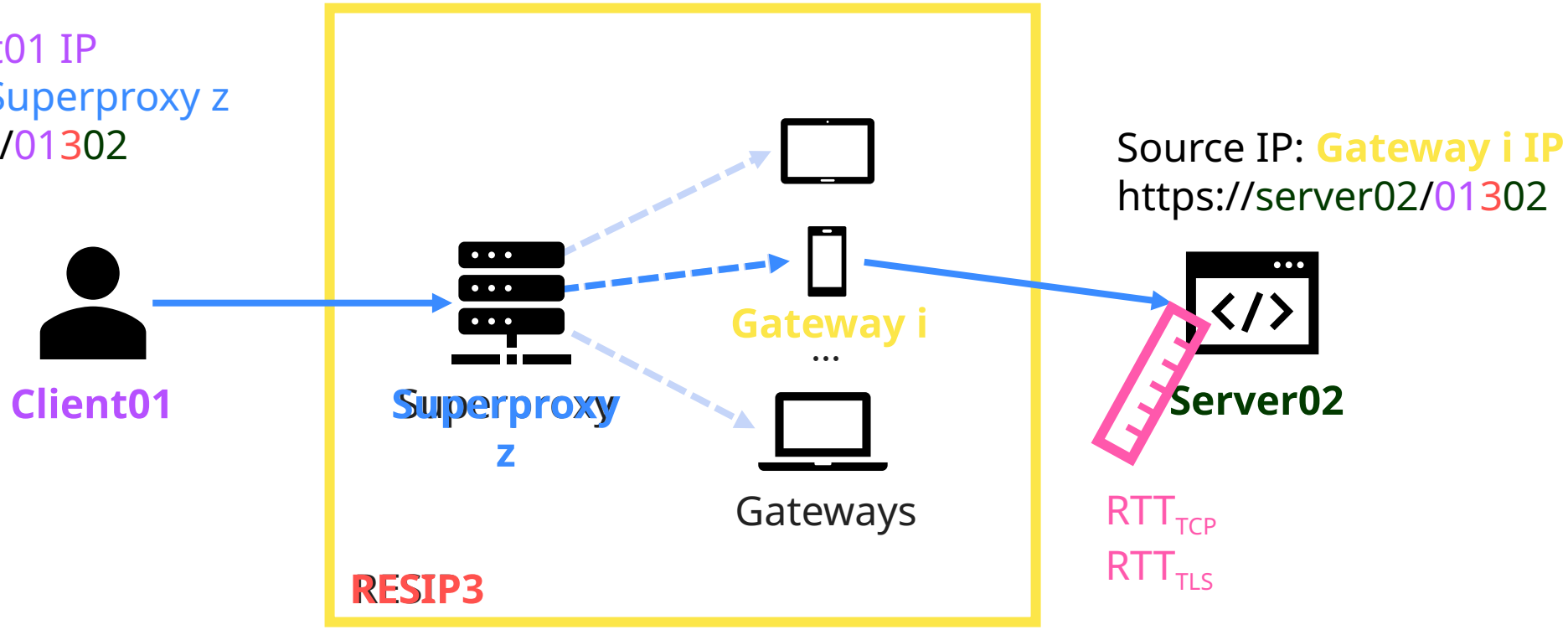
1. E. Khan et al. (2024) A First Look at User-Installed Residential Proxies From a Network Operator's Perspective. In CNSM 2024
2. E. Chiapponi et al. (2022). BADPASS: Bots taking ADvantage of Proxy AS a Service. In ISPEC 2022.
3. E. Chiapponi et al. (2023). Towards Detecting and Geolocalizing Web Scrapers with Round Trip Time Measurements. In TMA 2023.
4. E. Chiapponi et al. (2023). Poster: The Impact of the Client Environment on Residential IP Proxies Detection. In IMC 2023.
5. E. Chiapponi et al. (2023). Inside Residential IP Proxies: Lessons Learned from Large Measurement Campaigns. In WTMC 2023.

Check them here:

Backup slides

Data collection

HTTP Connect
Source IP: Client01 IP
Destination IP: Superproxy z
https://server02/01302



[2] E. Chiapponi et al. (2022). BADPASS: Bots taking ADvantage of Proxy AS a Service. In ISPEC 2022.

Bandwidth Broker earnings

Bandwidth Broker	Start date	Proxied	Flows	Earnings
BrightVPN	2024-03-07	190GB	1.10M	free VPN
earn.fm	2024-04-17	9GB	0.28M	1.69 USD
Honeygain	2024-01-01	48GB	3.90M	20.55 USD
Packetshare	2024-02-27	51GB	2.37M	10.34 USD
PacketStream	2024-04-25	2GB	0.64M	0.21 USD
IP Royal Pawns	2023-11-17	55GB	2.62M	11.48 USD
Proxyrack	2024-01-01	3GB	1.12M	2.07 USD
Repocket	2024-01-01	10GB	1.79M	7.2 USD
Total		368GB	13.82M	53.54 USD