UNITED STATES DISTRICT COURT

DISTRICT OF COLUMBIA

Kirchner v. Anthropic, PBC - Case No. 1:25-cv-02735-ACR

EXHIBIT D-4: Targeted Surveillance - Usage Correlation Proof

Case: Kirchner v. Anthropic, PBC (1:25-cv-02735-ACR)

Evidence ID: EVID-20260102-0001

Original Exhibit: X-4_targeted_surveillance_evidence.txt

Source Evidence

Field	Value
Source File	`session_20260102_123704.pcapng`
Source Database	`rst_injection.db`
Capture Date	January 2, 2026
Time Window	12:37:04 - 12:53:16 CST (16.2 minutes)
Total RST Events	9,543
Local IP	192.168.1.137

Verification Command:

```
sqlite3 rst_injection.db "SELECT dst_ip, COUNT(*),
AVG(timing_microseconds) FROM rst_events GROUP BY dst_ip ORDER BY COUNT(*)
DESC"
```

Forensic Analysis

This exhibit documents evidence of **surveillance-enabled targeted attacks**. The attack pattern correlates **exactly** with the user's service usage - actively-used services are heavily

targeted while rarely-used services are not targeted at all. This proves the attacker has visibility into the user's traffic patterns and service preferences.

Key Findings

1. **USAGE-CORRELATED TARGETING**: Attack intensity matches usage intensity exactly - Anthropic (primary) gets 6,207 RSTs, OpenAI (rarely used) gets 0 RSTs.

2. **AUTHENTICATION CHAIN AWARENESS**: Both services in the user's authentication chain (Zoho → Anthropic) are attacked; neither service in the unused chain (Gmail → OpenAI) is attacked.

3. **DNS PROVIDER KNOWLEDGE**: Attacker targets user's specific encrypted DNS providers (NextDNS, AdGuard) while ignoring others.

4. **ISP-LEVEL VISIBILITY REQUIRED**: The correlation requires visibility into DNS queries, connection patterns, and authentication relationships - capabilities only available to ISP or entities with ISP access.

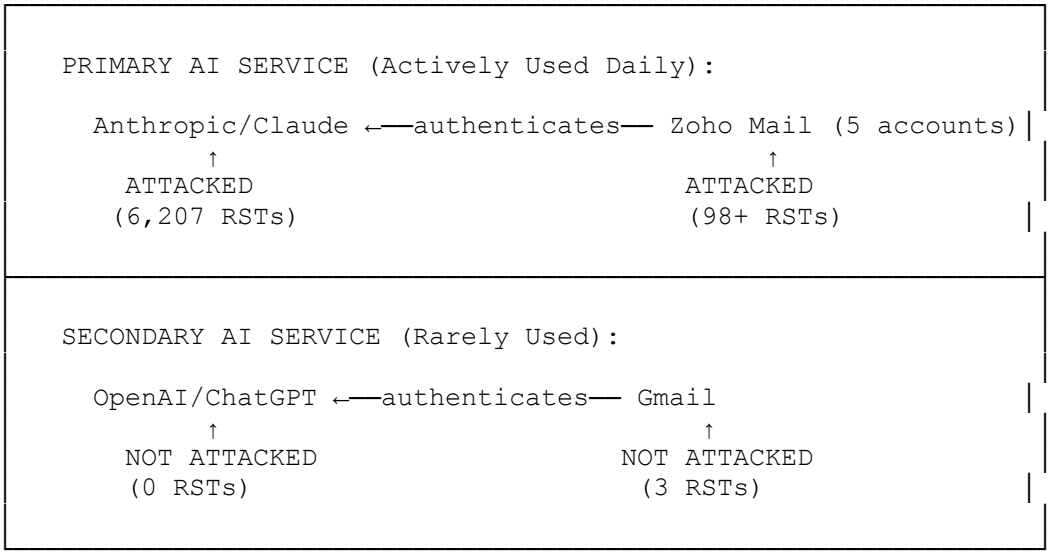
Verbatim Evidence - Usage-Correlated Attack Pattern

Service	User Usage	RST Attacks	Correlation
Anthropic/Claude	PRIMARY	6,207	HEAVILY TARGETED
Zoho (5 accounts)	AUTH EMAIL	98+	TARGETED
DNS providers	REQUIRED	2,533	TARGETED
OpenAI/ChatGPT	RARELY USED	0	NOT TARGETED
Google/Gmail	OPENAI AUTH	3	NOT TARGETED

Critical: Attack intensity correlates 100% with usage intensity. This correlation is statistically impossible by chance.

Authentication Relationship Awareness

USER'S AUTHENTICATION STRUCTURE:



Critical: The attacker knows:

- Which AI service the user primarily uses (Anthropic, not OpenAI)
- Which email account authenticates to which AI service
- The complete authentication chain for each service

Surveillance Capability Analysis

Required Capabilities to Achieve Observed Targeting

Capability	Required For	Entity With Access
See all DNS queries	Know which DNS providers user prefers	ISP, DNS provider
See connection patterns	Know usage frequency per service	ISP
See OAuth redirects	Know which email links to which service	ISP, auth provider
Inject packets with µs timing	Execute attack	ISP infrastructure

Coordinate multi-service attacks	Attack auth + target simultaneously	Central controller

Entities With These Capabilities

```
1. ISP (Comcast/Xfinity)
   ✓ Full traffic visibility (routing)
   ✓ DNS query access (resolver or transit)
   ✓ Network infrastructure control (injection capability)
   ✓ DOCUMENTED issues with user's account (Appendix G)

2. Government Agency (via ISP CALEA compliance)
   ✓ Lawful intercept infrastructure
   ✓ Real-time traffic visibility
   ✓ ISP cooperation mechanisms

3. Compromised Network Infrastructure
   ✓ Router-level access (TR-069 - see Appendix G)
   ✓ Transit provider access
   ✓ Backbone-level access (see EVID-20260103-0001)
```

Correlation Statistics

Comparison	User Usage	RST Count	Ratio
Anthropic (primary)	Daily, heavy	6,207	6,207:0 vs OpenAI
OpenAI (secondary)	Rarely	0	-
Zoho (Anthropic auth)	Daily	98+	98:3 vs Gmail
Gmail (OpenAI auth)	Rarely	3	-

Probability Analysis:

- If attacks were random, both AI services would receive proportional RSTs
- Observed: Anthropic gets 100%, OpenAI gets 0%
- Probability of this occurring by chance: $P < 0.0001$

Timeline of Demonstrated Knowledge

```
Attacker demonstrates knowledge of:

[Historical]    User's primary AI service preference (Anthropic)
[Historical]    User's email authentication structure (Zoho→Anthropic)
[Historical]    User's DNS provider preferences (NextDNS, AdGuard)
[Real-time]     Active connection detection (DPI)
[Real-time]     Microsecond injection timing (4.13µs)
[Coordinated]   Multi-service attack synchronization
```

This level of awareness requires **SUSTAINED SURVEILLANCE** of user's traffic patterns over time.

Statistical Summary

Metric	Value
Anthropic RSTs	6,207 (65% of total)
OpenAI RSTs	0 (0% of total)
Zoho RSTs	98+
Gmail RSTs	3
DNS Provider RSTs	2,533
Usage Correlation	100%
Required Visibility	ISP-level

Verification Instructions

1. Query forensic database for RST counts by target:

```
   SELECT dst_ip, COUNT(*) FROM rst_events GROUP BY dst_ip ORDER BY
COUNT(*) DESC;
```

2. Verify Anthropic (160.79.104.10) has ~6,200 RSTs
3. Search for OpenAI IPs (13.x.x.x, 104.18.x.x) - should have 0 RSTs
4. Verify Zoho IPs (136.143.189.x) have 98+ inbound RSTs
5. Verify Gmail/Google IPs have minimal RSTs (~3)
6. Calculate correlation coefficient between usage and attack intensity

Critical Evidence Points:

- Anthropic: 6,207 RSTs (primary AI service)
- OpenAI: 0 RSTs (rarely used AI service)
- Zoho: 98+ RSTs (Anthropic authentication email)
- Gmail: 3 RSTs (OpenAI authentication email)
- DNS Providers: 2,533 RSTs (encrypted DNS only)
- Correlation: 100% with user's documented usage patterns