

The State of Scams in Digital Communities

Telegram Edition · 2026

Agenda

1. The collapse — communities are the new attack surface
2. Three eras — pre-crypto noise, crypto fraud, the LLM era we're in now
3. How modern attacks defeat classical tools — ten failure modes
4. The AI threat in detail — the GPT-spam economy & the agentic shift
5. The Sybil problem — one actor, many identities
6. Beyond the channel — Telegram is the entry, not the endpoint
7. Ecosystem-specific patterns — where the money is
8. What works — and what doesn't
9. What's coming (*to follow*)
10. What community operators can do (*to follow*)

Parts 9-10 ship in the follow-up deck.

PART 1

The Collapse

Communities are the new attack surface

2025 — by the numbers

\$8.6B

U.S. investment-fraud losses

Largest IC3 loss category. Crypto-investment long cons — victims groomed over weeks or months — account for \$7.2B of it.

FBI IC3 • 2025

43.5M

Telegram groups & channels banned

Post-Durov-arrest platform cleanup (Aug 2024 onward).

CHECK POINT RESEARCH • 2025

+2,000%

surge in malicious Telegram-group scams

Two-month window, late 2024 into early 2025.

SCAM SNIFFER • NOV 2024 → JAN 2025

\$35B+

in stablecoins through Telegram-based illicit marketplaces

Huione & Xinbi "Guarantee" escrow markets — scam tooling, laundering, trafficked-labor brokering. More on these in Part 5.

ELLIPTIC • 2025

Drawn to scale

The same numbers, side by side.

SE-ASIA SCAM-CENTRE REVENUE • PER YEAR

~\$40B



HUIONE & XINBI MARKETPLACE FLOWS • LIFETIME

\$35B+



U.S. INVESTMENT-FRAUD LOSSES • 2025, REPORTED

\$8.6B



DPRK CRYPTO THEFT • 2025

\$2.02B



Different metrics — annual revenue vs. cumulative flows vs. reported losses — but the magnitudes are the point.

Why now

Three forces converging:

01

Telegram is the de facto Web3 social layer. Crypto runs there. So do the scams.

02

LLM-based attack tooling is commodity. Jailbroken AI models sell *on Telegram itself* — from ~\$200/month in 2023 down to \$220 *lifetime* in 2025, advertised openly in adversary channels.

03

Existing moderation was built for the pre-AI threat model. Rule-based bots cannot keep up with conversational, hyper-personalized, multi-turn scams.

Communities are dying — either to spam itself, or to over-moderation deployed to fight it.

PART 2

Three Eras

Pre-crypto, pre-LLM, LLM — and the steepest climb is just beginning

Three eras

ERA I • PRE-CRYPTO

Noise

Scripts. Mass invites. Keyword spam. Drive-by promotion. A nuisance, not predation — little in the room worth stealing.

What classical tools were built for.

ERA II • CRYPTO • PRE-LLM

Fraud

Money enters the room. Spam becomes fraud — pig-butcherer, fake admins, wallet drainers — but still human-run, so it stays labor-bottlenecked and leaves seams.

Where classical tools started losing.

ERA III • LLM

AI-scaled

AI writes the content and runs the personas. The labor bottleneck is gone — one operator runs thousands. We're here now, and the sophistication curve is still climbing.

Where we are now — and accelerating.

Three eras, in sequence — and nothing retires. Defenders face all three at once, and the LLM era is just getting started.

PART 3

Ten Failure Modes

How modern attacks defeat classical tools

01 · Clean-chat admin-impersonation DM

The dominant scam vector of 2025.

The attack:

1. Scammer monitors a public chat. Spots a user asking a question or expressing frustration.
2. Within minutes, DMs that user from a profile copying the admin's display name and avatar.
3. Often with a one-character substitution (`0` for `O` , `l` for `I`).
4. The in-chat message is benign English. The malicious payload — seed-phrase ask, fake KYC link — lives in the DM.

WHY CLASSICAL TOOLS MISS

Public message contains no banned keyword, no link, no media. No profile-similarity check vs. real admins. No notion of "this account joined 3 minutes ago and is already DMing people."

02 · LLM-generated conversational spam

The slow pivot.

The attack:

1. Account joins. Posts on-topic, natural English. Matches channel culture.
2. Engages with other members for several turns. Builds plausibility.
3. Pivots — usually to a link, an offer, or a DM ask.

Powered by FraudGPT / WormGPT / jailbroken commercial models. Sold on Telegram itself — now as cheap as \$220 *lifetime*.

WHY CLASSICAL TOOLS MISS

No banned tokens appear. Messages are unique per user (defeats duplicate-detection). Tone matches the channel. No keyword, no signature, no n-gram statistic catches "the slow pivot."

03 · Edit-after-post payload swap

The simplest bypass — and most bots still don't catch it.

The attack:

1. Spammer posts a clean message — passes all filters.
2. Waits out the join-restriction window (usually a few minutes to an hour).
3. Edits the message to inject a link or scam text.
4. Filters don't re-run on the edit. The edited payload sits in the chat.

WHY CLASSICAL TOOLS MISS

Most moderation bots subscribe to the `message` update only. Telegram's `edited_message` update bypasses the entire detection pipeline. Trivially fixable — most just haven't fixed it.

04 → 06 · The signature problem

04 Homoglyph / zero-width obfuscation

Cyrillic `а` / `е` / `о` and inserted `U+200B` / `U+200C` / `U+200D` split keywords like "airdrop" or "metamask." Blocklists match code-points, not glyph shapes.

05 Image / screenshot scams

Scam URL or wallet address rendered inside a JPEG, sticker, or fake-pinned-message screenshot. No OCR; media-locks block media wholesale, which is unusable in any active chat.

06 Aged-account swarms

Operators buy 3-month-to-5-year-old Telegram accounts with built-up history specifically to defeat anti-spam heuristics. Account-age checks pass cleanly.

07 → 10 · The behavioral problem

07 CAPTCHA bypass + CAPTCHA fatigue

Solver services defeat CAPTCHAs cheaply. Meanwhile, aggressive CAPTCHAs cause real-user dropoff (real members give up and leave). Necessary but insufficient.

08 Lurker-to-DM funnel

Scammer joins quietly, never posts in-channel, scrapes member list, opens DMs. Lurker accounts never trip any in-group rule. Feeds the \$8.6B in 2025 U.S. investment-fraud losses.

09 Hyper-personalized phishing

Messages reference the victim's project, wallet, or on-chain activity. Breached data + LLM = uncanny specificity. No signature exists.

10 Malicious "support" bots

Bots named `@MetaMask_Official_Support` , `@Trezor_Help_Bot` . Bot verification exists since Jan 2025, but adoption is minimal — any bot can still launch under any name.

PART 4

The AI Threat

in detail

The GPT-spam economy

PRICE COLLAPSE

\$220 lifetime

for jailbroken LLM scam tooling — down from
~\$200/mo in 2023

WORMGPT 4 • THE REGISTER • 2025

GROWTH SIGNAL

+1,265%

phishing-volume surge in the 12 months after
ChatGPT launched

SLASHNEXT • Q4 2022 → Q3 2023

DISTRIBUTION CHANNEL

Telegram itself

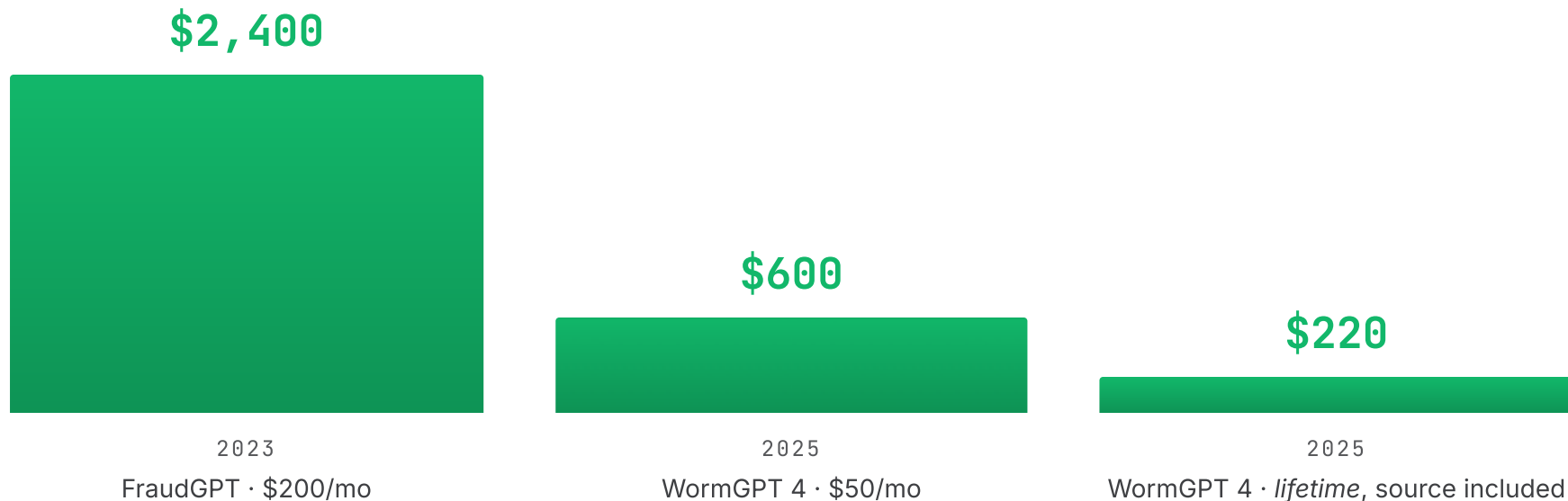
The same platform we're defending is the marketplace for the
attack tools.

CAPABILITY

Multi-turn conversational coherence. Tone-matching
to channel. Hyper-personalization from breached-
data context.

The cost of entry, collapsing

A year of jailbroken-LLM scam tooling, list price.



–91% in two years. The last bar isn't even a subscription.

Why signatures fail

Pre-AI: signatures worked

"Send 1 BTC to address X to receive 2 BTC." Repeated verbatim across thousands of messages. Hash it, block it, done.

Post-AI: signatures don't

Every message is generated fresh. Same scam intent, infinite surface variation. No two payloads look alike.

What remains?

Detection has to move from *what was said* to *what the actor is doing* — behavioral patterns, conversation context, identity signals.

The agentic shift

From AI-generated content to AI-driven campaigns.

Case study: "Patriot Bait" (May 2026)

One Russian-speaking actor — "bandcampro."

17,000-subscriber Telegram channel impersonating an American veteran.

Live since Feb 2021; AI-automated phase Sept 2025 – May 2026.

Python orchestrator ("Quantum Patriot") called Gemini through 73 stolen API keys; Venice.ai-powered "QFS terminal" chatbot for victims.

Jailbreak: convinced Gemini he was an "authorized pentester" — saved to model memory file.

Academic context: the ScamAgents paper (arXiv 2508.06457) formalizes this pattern — autonomous multi-turn agents with memory; current LLM safety guardrails ineffective.

TREND MICRO INVESTIGATION

"What previously required a team of writers, social media managers, IT workers, and malware programmers can now be automated by a single actor using a VPS, a Telegram bot, and API access to frontier models."

The Truman Show

Synthetic communities. Multi-week immersion.

Pattern: OPCOPRO (Oct 2025 onward)

Tightly controlled Telegram + WhatsApp groups built around AI-generated "experts" and synthetic peers.

Multi-week narrative immersion — looks and feels like an institutional trading community.

Companion mobile apps in the App Store + Play Store as WebView shells. Balances, trades, agreements all fabricated server-side.

Then the ask.

WHY THIS DEFEATS EVERYTHING

There is no spam message. The entire context is the attack. Every per-message detector misses. Every lurker → DM heuristic misses. The community itself is fabricated.

PART 5

The Sybil Problem

One actor. Many identities.

One actor, many faces

The economics of multi-account operations.

AGED ACCOUNT PRICE

\$1–\$10

per aged Telegram account on adversary marketplaces.
Three-month-old to five-year-old accounts available
wholesale.

DOCUMENTED CASE

PumpCell ring

Coordinated multi-chain pump-and-dump across dozens of
Telegram channels with synchronized account activity.
Netted ~\$800K in one month.

MEDIUM RING SIZE

50–200

accounts in a single coordinated operation. Larger
professional rings reach into the thousands.

UNIT ECONOMICS (ILLUSTRATIVE)

~\$500 → \$5K–\$5M+

Indicative cost of a 100-account swarm at marketplace prices
vs. payout of a single successful scam. 10× to 10,000× ROI on
the infrastructure.

Why this breaks defense

All classical defenses share one assumption:

Each account is an independent actor.

When accounts are coordinated, the assumption breaks:

Bans don't matter	the actor has 50 spares ready
Account-age checks don't help	they buy aged accounts wholesale
Reputation tracking resets	fresh accounts have no history
Federation banlists fight a hydra	cut off one head, two grow

THE STRUCTURAL GAP

Classical defense has no way to collapse N accounts back to 1 actor. Without that collapse, defense is whack-a-mole forever.

The hydra, at platform scale

The largest illicit marketplaces in history were Telegram channels.

BEFORE • 2021-2025

\$35B+

Huione & Xinbi "Guarantee" —
Telegram-native escrow markets.
Scam tooling, laundering,
trafficked-labor brokering. Paid in
stablecoins.

THE BAN • MAY 13, 2025

**Thousands of
channels removed**

Telegram's largest targeted
takedown, following Elliptic
research and FinCEN's Section 311
designation of Huione Group.

AFTER • WEEKS LATER

~70×

surge in daily volume on successor
Tudou Guarantee — 30% owned
by Huione. Xinbi returned under
the *same account ID*, inflows
+90%.

The takedown worked exactly as designed. The market didn't even slow down.

PART 6

Beyond the Channel

Telegram is the entry. Not the endpoint.

Cross-platform kill chains

The biggest 2025 losses didn't end on Telegram. They started there.

DPRK CRYPTO THEFT • 2025

\$2.02B

Aggregate 2025 total. +51% YoY. \$300M+ attributed specifically to deepfake Zoom & Teams meetings.

VOICE-NOTE DEEPFAKES • 2025

\$5M+

Distress scams using voice clones. 3 seconds of source audio is enough via consumer voice-cloning APIs.

THORCHAIN CO-FOUNDER • 2025

\$1.35M

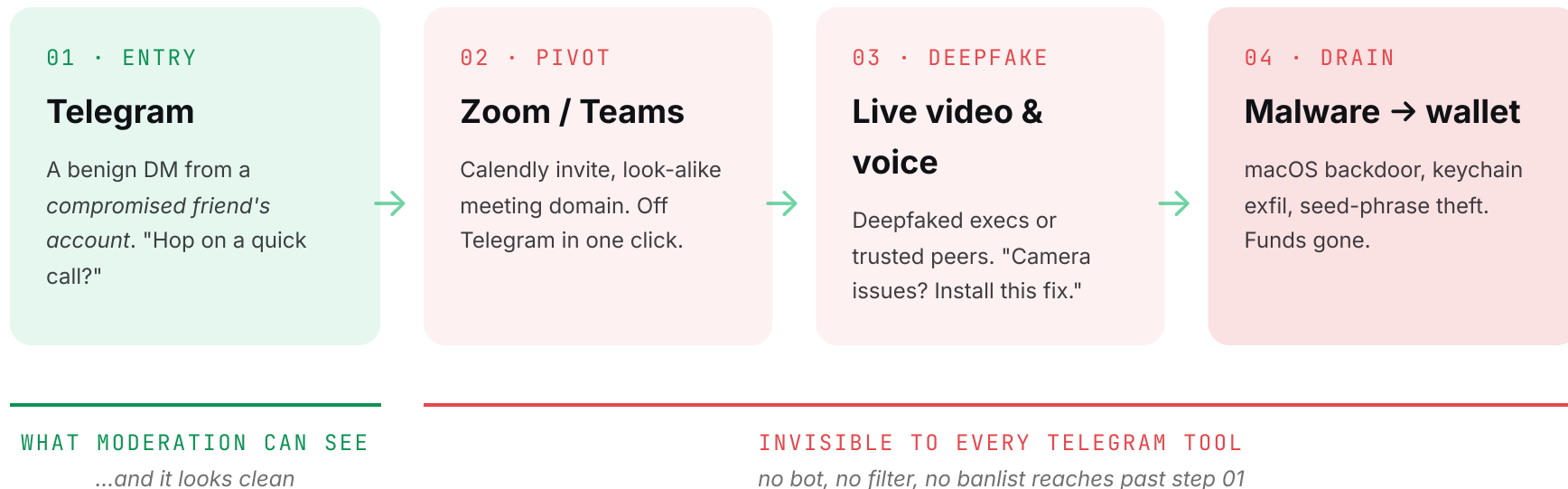
Entry: hacked friend's Telegram account. Pivot: Zoom deepfake call. Endpoint: macOS malware.

THE PATTERN

Telegram entry → cross-platform escalation → deepfake → drain.

Single-platform moderation cannot see the kill chain.

Anatomy of the kill chain



The only step a Telegram defense can touch is the one with nothing to detect.

Who runs the scams

The supply side is industrial — and increasingly, indicted.

SE-ASIA SCAM CENTRES • UNODC

~\$40B/yr

generated by Southeast Asian scam centres. UNODC names Telegram as "convenient and scalable for illicit actors."

PRINCE GROUP INDICTMENT • OCT 2025

\$15B

127,271 BTC seized from chairman Chen Zhi — the largest forfeiture in U.S. history, tied to forced-labor compounds.

FORCED LABOR

~220K

people trafficked into scam compounds in Cambodia (~100K) and Myanmar (~120K). The "scammer" DMing your community is often a victim too.

KK PARK RAID • OCT 2025

2,000+

workers freed from a single Myanmar compound. 30 Starlink terminals seized. One compound, of hundreds.

PART 7

Ecosystem Patterns

Where the money is

Where the money is

Crypto communities bear disproportionate spam load.

TARGETED VALUE

Highest dollar-amount per victim of any consumer category. A successful long-con investment scam nets six to seven figures.

DOCUMENTED CAMPAIGNS

PumpCell: a Telegram-coordinated multi-chain pump-and-dump ring netted ~\$800K in a single month.

SOLIDUS LABS • COINDESK • DEC 2025

ATTACK SURFACE

Crypto-Telegram channels are inherently open (community discovery, free joining). Combined with wallet UX, every member is a wallet to drain.

BRAND IMPERSONATION

Fake "support" bots and channels impersonating major wallets and exchanges remain endemic — verification exists but is barely adopted.

AURA • KASPERSKY • 2024-2025

Ecosystem-specific surface

TON	Telegram-native distribution makes Mini-App phishing the dominant pattern. Wallet drains via in-app webview. (KASPERSKY · 2025)
EVM	Wallet-drainers, fake airdrops, malicious dApp signing prompts. Base + Ethereum take the volume.
SOLANA	Memecoin pump rings, fake token launches, drainer Telegram bots distributing seed-phrase phishing.
CROSS-CHAIN	PumpCell-style coordination: same actors run pump campaigns across multiple chains simultaneously, laundering across bridges.

TON in the crosshairs

The Telegram-native trust UI is the attack surface.

- **Fake Fragment clones**
Copycat username-auction sites that drain wallets through TON Connect when users "claim" or "list."
- **"TON Giveaway" / "Wallet Support" Mini Apps**
Telegram-native phishing UIs that request wallet authorization for fake events or fake support flows.
- **Malicious contract approvals via TON Connect**
Approvals that look like routine sign-in actions but actually grant drain permissions.

THE PROBLEM

Mini Apps run *inside* Telegram. Trust is implicit. No moderation tool scans Mini App manifests, contract approvals, or Mini App reputation.

The drainer economy

Wallet-drainers are SaaS — sold, supported, and updated on Telegram.

INFERNO

Inferno Drainer's 2025 return

"Retired" in 2023, back with encrypted C2 addresses stored *on-chain* and customer-run proxies. ~\$9M in its first half-year back. (CHECK POINT · 2025)

ANGELX

The Angel / AngelX takeover

Inferno handed its codebase to Angel Drainer. AngelX deployed 300+ malicious dApps — targeting TON and Tron specifically.

TRICKMO

TrickMo.C moves C2 onto TON itself

Android banking trojan (new variant, Jan 2026) running C2 over TON's ADNL overlay — no DNS, no takedown surface. (THREATFABRIC · 2026)

WHY IT BELONGS IN A SPAM DECK

The spam *is* the distribution layer — every drainer needs a delivery channel, and that channel is the community attack surface of Parts 3–5.

PART 8

What Works

and what doesn't

What still works

Don't throw it out:

Join-time CAPTCHA + restrict-on-join

~95% spam reduction reported (vendor figures). Highest-ROI single filter.

Federated banlists (CAS-style)

Cross-chat ban-sharing. High precision, free wins.

User-report systems

Cheap human-in-the-loop signal. Don't underestimate it.

Locks (URL / forward / media / invite)

Kills the long tail of low-effort drive-by spam.

Anti-flood (rate limits)

Still defeats the dumb half of swarms.

Account-age checks

Defeats casual attackers (though not professional aged-account farms).

What's plateaued

Rule-based detection against AI-generated content.

No signature exists when every message is unique.

Account-age heuristics against farmed-aged accounts.

Attackers buy aged accounts specifically to defeat this.

Single-platform reputation.

Attackers cross platforms; your bans don't.

Static keyword filters.

Homoglyph and zero-width obfuscation breaks them silently.

The platform's own response

Telegram moved more in 18 months than in the prior decade. It raised costs — not structure.

SEP 2024

Data-sharing policy reversal

IPs and phone numbers now go to law enforcement on valid request. U.S. requests fulfilled: 14 → ~900/yr.

JAN 2025

Third-party verification for accounts & bots

The anti-impersonation primitive defenders asked for — adoption so far is minimal.

2025

Paid messages ("Star Messages")

Per-message fees for non-contacts — economic friction aimed directly at the DM funnel.

2024–26

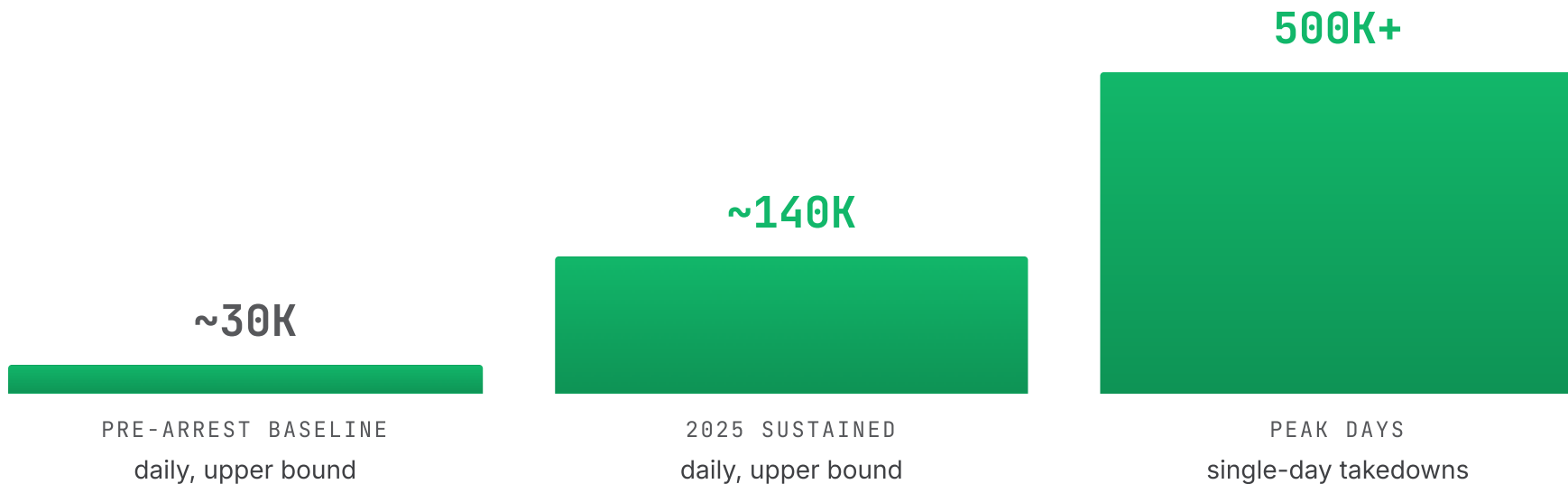
Mass takedowns — 43.5M channels in 2025

Attackers adapted: pre-staged backup channels, "Request to Join" gating that locks out automated moderation. Migration off Telegram barely registers.

Platform enforcement is now real — and the attacks of Parts 3–7 all post-date it. The structural gaps stay community-level.

Enforcement went exponential

Channels & groups taken down per day.



Throughput rose ~17x — and every attack in this deck post-dates it.

What's emerging

Three directions the industry is moving. The "how" is unsolved.

AI-assisted detection.

Necessary against generative spam. Cost discipline is the open problem.

Behavioral signal layers.

Detection that looks beyond message content. Mostly absent from deployed bots.

Verifiable identity.

Wallet-, credential-, or proof-of-human-gated trust for high-stakes contexts. Adoption is the bottleneck.

PARTS 9 – 10

To Follow

Predictions, and the operator playbook — coming in the follow-up deck.

References 1 / 2

[1] FBI IC3 — *2025 Internet Crime Report* (\$8.6B investment-fraud losses). ic3.gov

[3] Scam Sniffer (via Cointelegraph) — *Malicious Telegram group scams up 2,000% since November*. cointelegraph.com

[5] Kaspersky — *Telegram phishing and scams in 2025*. kaspersky.com/blog/phishing-and-scam-in-telegram-2025

[7] Dark Reading — *FraudGPT: malicious chatbot for sale on dark web*. darkreading.com

[9] SlashNext — *2023 State of Phishing* (+1,265% phishing volume, Q4 2022–Q3 2023). slashnext.com

[11] ScamAgents (arXiv 2508.06457) — *How AI Agents Can Simulate Human-Level Scam Calls*. arxiv.org/abs/2508.06457

[13] DarkGram (arXiv 2409.14596, USENIX Security 2025) — *339 cybercriminal Telegram channels, 23.8M followers*. arxiv.org/abs/2409.14596

[15] CoinDesk / Solidus Labs — *Telegram ring ran pump-and-dump network that netted \$800K in a month*, Dec 2025. coindesk.com

[17] Combot Anti-Spam System (CAS). cas.chat

[2] Check Point Research — *Telegram's crackdown in 2026 — and why cyber criminals are still winning* (43.5M bans, takedown rates). checkpoint.com

[4] Aura — *12 Latest Telegram Scams*. aura.com/learn/telegram-app-scams

[6] Coinbase — *How scammers are targeting crypto communities*. coinbase.com/blog

[8] The Register — *WormGPT 4: \$50/mo or \$220 lifetime*, Nov 2025. theregister.com

[10] Trend Micro — *Inside the Influence and Fraud "Patriot Bait" Campaign*, May 2026. trendmicro.com

[12] Check Point — *The "Truman Show" Scam: Trapped in an AI-Generated Reality*, Jan 2026. checkpoint.com

[14] Mobileproxy.space — *Telegram account farming in 2026* (farming-vendor marketing; cited as primary evidence of the gray market). mobileproxy.space

[16] Andrew Kotchnev / Medium — *Top 5 CAPTCHA bots for Telegram in 2026* (vendor roundup; source of the ~95% figure). medium.com/@a.kotchnev

References 2 / 2

[18] Elliptic — *Telegram market takedown: Huione & Xinbi Guarantee* (\$35B+ flows). elliptic.co

[20] CoinDesk — *Illicit crypto trade quickly rebounds despite Telegram's shutdown of \$35B Huione marketplace*, Jul 2025. coindesk.com

[22] U.S. DOJ — *Chairman of Prince Group indicted; 127,271 BTC forfeiture*, Oct 2025. justice.gov

[24] The Hacker News — *BlueNoroff Deepfake Zoom Scam*, June 2025. thehackernews.com

[26] Trend Micro — *AI voice cloning*, Apr 2026 (\$5M+ 2025 losses; 3-second clone threshold). news.trendmicro.com

[28] Infosecurity Magazine — *Telegram exploited in Toncoin scam*. infosecurity-magazine.com

[30] Cryptopolitan — *Inferno Drainer hands over to Angel Drainer (AngelX)*. cryptopolitan.com

[32] TechCrunch — *Telegram rolls out third-party account verification*, Jan 2025. techcrunch.com

[34] Fortune — *Telegram to provide user info to governments on legal request*, Sep 2024. fortune.com

[19] TRM Labs — *How Telegram's ban on Huione and Xinbi reshaped the guarantee-services landscape*. trmlabs.com

[21] UNODC — *Inflection Point*, Apr 2025 (SE-Asia scam centres, ~\$40B/yr, trafficking estimates). unodc.org

[23] Chainalysis — *Crypto hacking & stolen funds, 2025 year-in-review* (DPRK \$2.02B, +51% YoY). chainalysis.com

[25] crypto.news — *THORChain co-founder loses \$1.35M to North Korean Zoom scam*. crypto.news

[27] DailyCoin — *TON ecosystem phishing attacks on the rise*. dailycoin.com

[29] Check Point Research — *Inferno Drainer reloaded: the return of the most sophisticated crypto drainer*, 2025. research.checkpoint.com

[31] BleepingComputer — *TrickMo Android banker adopts TON blockchain for covert comms*, May 2026. bleepingcomputer.com

[33] Telegram — *Star Messages, Gateway 2.0 and more*. telegram.org/blog

Thank you

The State of Scams in Digital Communities

Telegram Edition · 2026

PART 1 OF 2 · MAY 2026

Read the full report: targa.dev/research/state-of-scams-2026

© 2026 Targa-Praetorian · Licensed CC BY 4.0 · cite as: Targa Research (2026). *The State of Scams in Digital Communities — Telegram Edition*.