2025 Report

Cyberfraud in Retail

The Changing Face of Fraud: Malicious Intent Gets Organized

A field guide to fraud typologies for cybersecurity practitioners in retail

NETACEA

Fraud is increasingly acceptable in the TikTok generation

Matthew Gracey-McMinn
VP Threat Intelligence, Netacea

Back in 2022 we produced a field guide on refund fraud typologies that was well received by several major retail brands. What we didn't expect was that the guide would also become seen as a definitive work by the fraud community! During multiple investigations into fraud groups our guidebook has turned up on a number of criminal forums as a window into how the 'good guys' see things.

The **Cyberfraud in Retail report 2025** is an evolution of that first report and is intended to serve as a field guide for anti-fraud and cybersecurity professionals everywhere. And probably a number of fraudsters too, I guess!

This report looks at both digital and offline fraud typologies in retail and considers the factors and trends that set the direction for cyberfraud evolution. Since 2022, one thing's for sure, the phenomenon we discovered of fraud becoming increasingly organized has only accelerated and now the profile of threat actors we track is changing on two fronts.

There's an increasingly professional presence associated with hardened fraudsters, especially those from Russian ransomware gangs, and an influx of more naïve young people for whom fraud is being normalized on social media. Both of these trends are bad news for retail brands and the second points to an increasingly accepting public sentiment towards casual fraud.

We recently ran a survey of over 2,000 consumers in UK and US to see if what our researchers are seeing online is reflected by the public view. Indeed, it is. Only 7% of respondents were not aware of retail fraud tactics, and 81% are familiar with the concept of DNA (Did Not Arrive) fraud.

Almost half of consumers (45%) have been targeted by ads for retail fraud guides or services on social media and almost a quarter (23%) have been tempted into committing retail fraud. Concerningly, a slightly higher number believe that there is an 'acceptable' level of casual fraud before morality gets the better of them, and 34% believe that fraud up to the value of $100 is acceptable.

In fact, 15% would consider more committing serious retail fraud activities 'under the right circumstances' - as in, if they really thought they could get away with it.

But perhaps the most shocking statistic of all is that 16% of consumers surveyed know someone who has participated in an insider fraud scheme.

This does not bode well for retailers facing increasingly sophisticated adversaries as well as a public with a low threshold for opportunistic theft.

The good news is that a combination of technology and intelligence can help you identify and understand your risks both internal and external. Fraud, infamously, is a lagging indicator of a problem, but fraud intelligence can help you detect, and even predict, malicious intent.

Read on to find out how.

The ‘businessification’ of cyberfraud

A growing challenge for retail

The impact of cyberfraud is significant for retail brands. According to Gartner, over $48bn was lost to online fraud in retail in 2023.

But this doesn’t factor in losses from outside of the digital customer journey - essentially confined to the website or app - where post payment fraud and refund fraud are beginning to weigh heavily on the bottom line.

NRF National Retail Federation

In fact, figures from the National Retail Federation suggest that returns fraud could more than double the losses from online fraud to over

$101 billion,

while for every $100 in returned merchandise, retailers will lose $13.70 to return fraud.

Find out more

It gets worse. A dollar stolen is not written off as such. For every $1 lost to retail fraud in the US, it costs an enterprise $3 in recovery or making good on the loss, and then we have the reputational impact, which for any kind of cyber attack on a consumer brand is around 10x more than the financial costs, Gartner says.

All told, the impact of cyberfraud could be a trillion-dollar problem for the retail industry.

Fraud in numbers

Netacea Threat Intel Center tracks listings for over 2,000 digital market places frequented by fraudsters.

These marketplaces mostly sell stolen accounts for digital services including media streaming, gaming, adult entertainment, consumer SaaS products and education services, as well as gift cards and accounts linked to loyalty or bonus points.

Many of these marketplaces are frequented by consumers looking to make a one-off purchase to get access to a service such as Netflix for $10 versus $25 retail, as well as by organized fraud groups trading wholesale or building large-scale operations that require many sock puppet accounts for other more complex fraud typologies.

The number of individual sellers we tracked on these marketplaces increased from

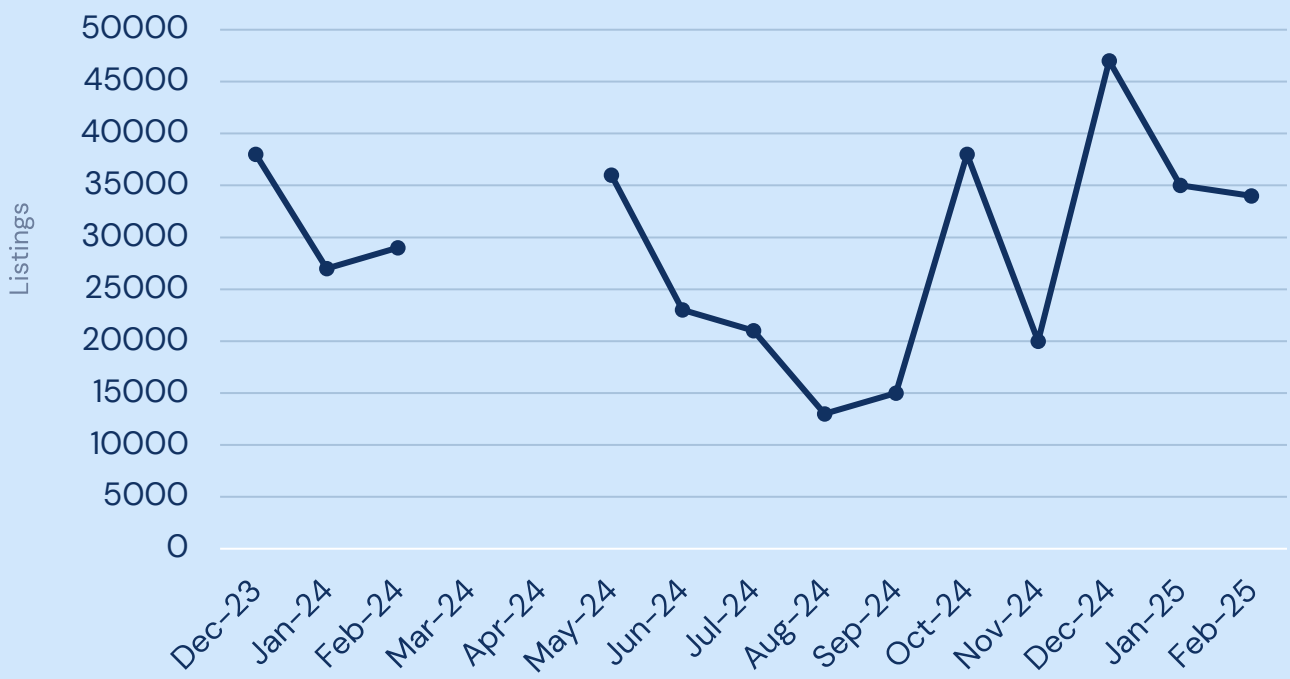
1,718

in January 2024 to reach a 12-month high of

2,785

in January 2025.

Marketplace listings



*missing data is when collection tools were temporarily blocked

Our chart for 2024 shows that the number of listings for available products closely follow seasonal retail trends, picking up momentum in October and November, peaking in December before getting quieter in late January and February.

At their height, the number of individual product listings tops

45,000

per day, averaging around

20,000

per day throughout the year.

While it's possible some of the listings are fake there's little incentive for false postings.

It's true that fraudsters are not exactly trustworthy, but the marketplaces use a reputational reviews system to rate sellers and our Threat Intel Center can confirm that these marketplaces are very much trading in real products.

Selling price of stolen accounts

*missing data is when collection tools were temporarily blocked

How much does a stolen account cost?

We can see from this chart that over the last 12 months, stolen accounts and gift cards have never traded for more than 40% of face value. So, a $100 gift card could be illegitimately acquired for $40, or a premium Netflix subscription would cost $10 versus the $25 monthly face value, for example.

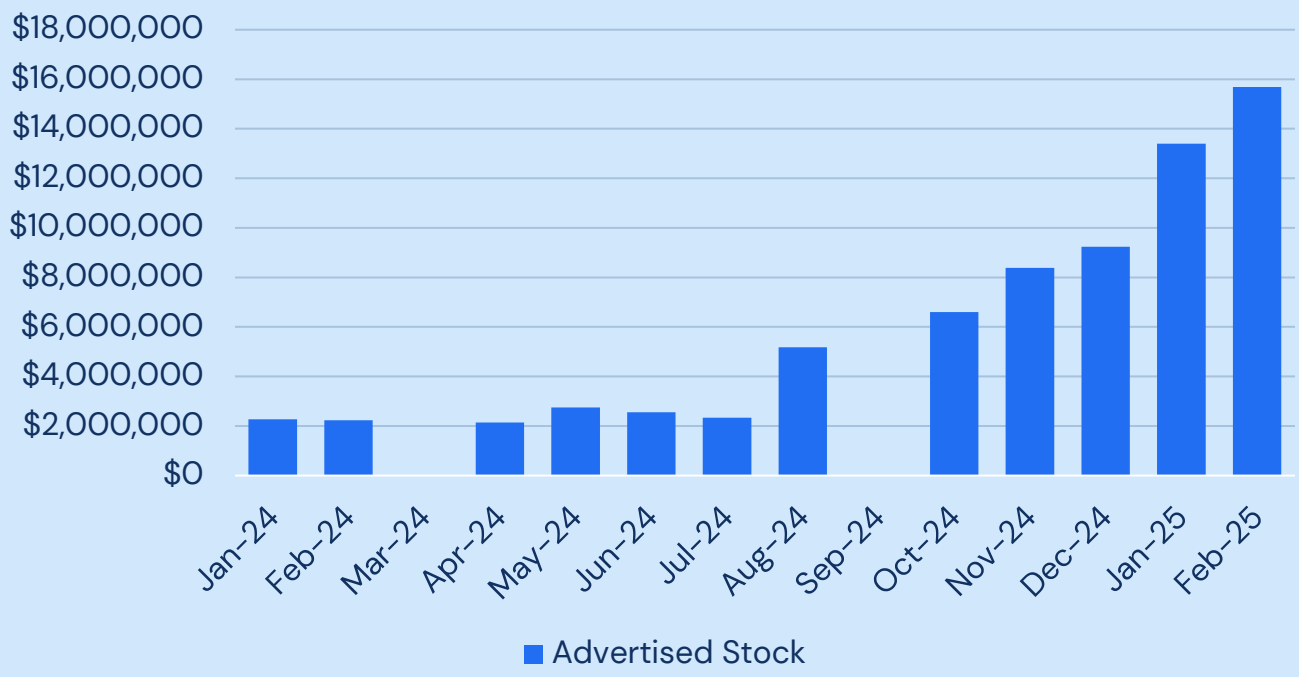
We can see from the data that the discount percentage for gift cards peaked in December at more than 80%, which in line with the spike in stock listings could indicate increased competition in the market or efforts by sellers to shift excess stock.

We can also see that some level of increased discount has persisted in the first two months of 2025, marking at least a 10% reduction in the cost of discount cards year-on-year.

Although we don't show historical data in this report, we can confirm that the data from the last 12 months is typical and an indication of what is an acceptable risk-to-reward ratio for sellers and bulk buyers of stolen accounts.

Generally speaking, a fraud enabler will expect to pocket at least 30% of the value of the fraud being committed, whether that's a stolen product being resold, or a Fraud-as-a-Service (FaaS) offering.

Total value of stolen gift cards available



*missing data is when collection tools were temporarily blocked

What we can see in terms of trends is that the total value of gift cards available on underground marketplaces rose considerably through 2024.

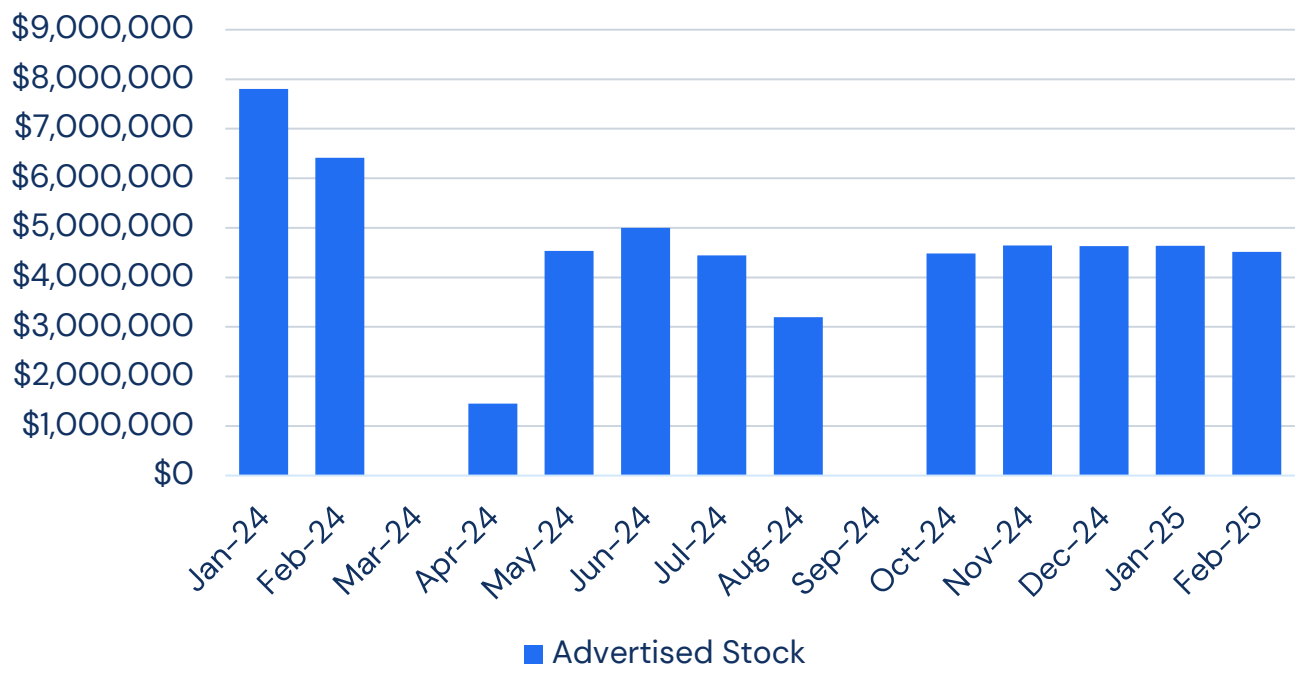
From around

$2.3m at the start of the year to over **$9.2m** in December

Jumping almost

500% year-on-year to over **$13m** in January 2025

Total value of stolen accounts with monetary balance or loyalty points



*missing data is when collection tools were temporarily blocked

Conversely, the value of account balances – stolen accounts that have a payment card, monetary balance, or loyalty points associated – dropped year on year, despite remaining constant throughout 2024.

We can see total balances available hover around the $4.5m to $5m mark monthly throughout 2024. But there was a significant drop on total balance value from $7.8m in January 2024 to $4.6m in January 2025 and a similar decline in February year-on-year. This seems contrary to market trends where US shoppers were expected to spend almost 50% of their holiday gift budget on gift cards in 2024, up nearly 10% on 2023.

While we can only speculate on the reasons for this decline, the most likely reason is that gift card holders were forced to spend their balances to survive the holidays. Research suggests that around 40% of US consumers did not spend their gift card balance on a present for themselves but instead on essentials such as groceries and gas. This is perhaps a reflection of the economic hardship that many people are facing and points to increasing fear of inflation and rising prices.

Money back: Refund fraud on the rise

Netacea's Threat Intel Center also monitors and infiltrates organized fraud gangs directly. We maintain over 700 sock puppet identities through which we access well over 3,000 closed forums and chat groups on the open, deep and dark web, as well as Discord, Telegram and Signal.

This chart shows the popularity of different refund fraud methods in Telegram group chats and is representative of what we see across all channels. Each mention is part of a discussion about fraud typologies related to a specific brand, so it does not include general or non-specific chat.

As you can see, Did Not Arrive (DNA) is the technique of choice for fraudsters, appearing in nearly 40% of all conversations. Its popularity is likely due to the fact it is the easiest method to attempt, with a worthwhile success rate against high volume retailers.

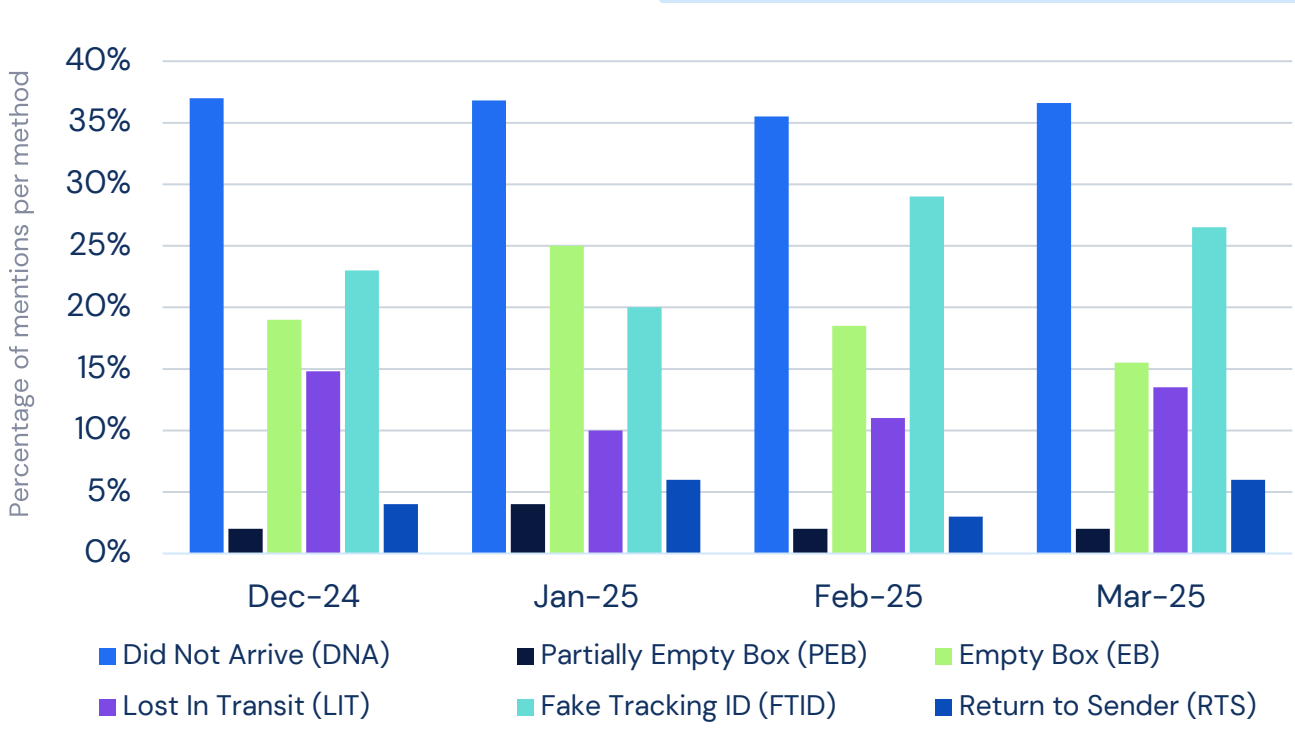
But DNA is less effective against lower volume retailers, certain tracking and fulfilment methods – such as signed for delivery - and high value items.

This is where Fake Tracking ID (FTID) emerges as the second most popular method and where professional fraudsters, known as 'boxers', come into the picture.

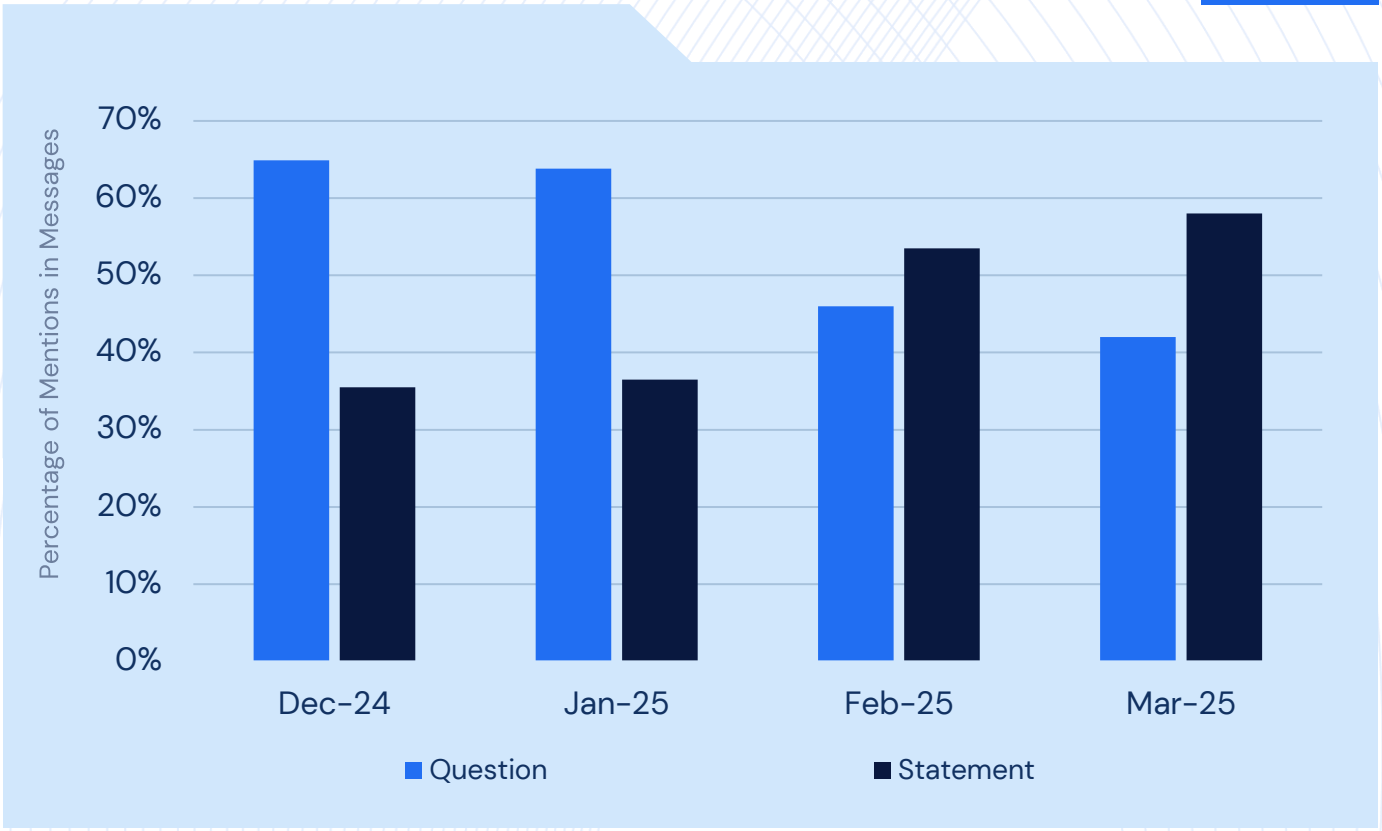
The FTID method of fraud is where the return postage label is altered and used to mail an empty or junk filled package instead of the item for which the refund has been requested. These methods are somewhat more involved and require a level of know -how on the part of the fraudster, which is why we see FTID as more of a service offering from career criminals.

We go into more detail on refund fraud typologies a little later in this field guide.

Postpayment fraud methods



Q&A: Transfer of fraud knowledge in forums



This chart shows comparative trends for people asking questions in Telegram chats, such as 'which fraud techniques work for retailer XYZ?' versus those making a statement, such as 'I successfully performed a DNA claim against retailer XYZ'.

We can see that questions peak in line with retail seasonality and fraud activity increases during the holidays, but questions begin to taper off in February and March leaving the more tenacious fraudsters to stick around and share information.

While some fraud groups are more exclusive than others, it's not hard to stumble across the fringes of the ecosystem through simple search or Reddit posts, which then direct you to more gated communities such as Telegram or Discord chats.

Premium membership services from boxer groups provide access to many of the techniques that professional boxers use to minimize risk of failure for their clients. These services have brand-specific insights for popular retail and reseller marketplaces and include instructive guides tailored to different regions and local couriers.

Some services even provide 24/7 support from trusted members of the community, as well as tiered upgrades for even more niche stores and methods along with their tested conditions.

At the lower end of the market, fraudsters sell packages of 'how to' guides with instructions on attempting fraud yourself. Our search across the 2,000 fraud marketplaces we monitor turned up 175 product listings for variations of refund methods ranging from $40 for a basic instruction guide to several hundred dollars for an 'everything package' claiming to provide details for multiple brands and storefronts.

Impact of law enforcement activity

Law enforcement activity can have a significant impact on data collection.

The data and insights collected and created by the Netacea Threat Intel Center rely on our access to useful data sources.

As well as the thousands of marketplaces we monitor, we also have access to over

3,000

closed threat actor forums on the web and channels like Telegram and Discord.

Some notable recent takedowns include:

Operation Talent

Which took place at the end of January 2025 and resulted in the seizure of markets and forums Sellix, Cracked and Nulled, affecting more than

10 million users in total

These sites were some of the most prolific one-stop-shops for hackers and fraudsters looking for illegal goods, tools and cybercrime services.

We anticipate an impact on the data we collect as new pretenders emerge to fill the void and stocks are relisted on other fraud marketplaces.

Breach Forums, also known as Breached, was taken down in May 2024 after reappearing following a previous takedown in 2023. The site appeared yet again just weeks later at the end of May 2024. While Breached was another of world's most notorious hacker forums with over

250,000

active members, the frequent appearing / disappearing activity over the last 12 months has created a lot of suspicion in the criminal ecosystem that the forum is compromised and could even be a honeypot. As a result, data collected from this site is no longer considered reliable.

Cyberfraud is getting more organized

Retail fraud has long been seen as ‘a victimless crime’.

Social media today is awash with so called ‘influencers’ pumping out promotional ads for refund fraud schemes in an effort to normalize the practice. After all, if you claim you didn’t receive the item you ordered and ask for a refund, no one gets hurt. On the one hand it’s just written off as a loss by the faceless corporation and they probably have insurance. On the other it’s an opportunity to ‘stick it to the man’.

Aside from being a criminal offense, the truth is that these losses are eventually passed on to the customer in terms of higher prices. There’s also the direct impact on consumers from fraudsters misappropriating digital customer accounts, causing personal pain and destroying sentiment between customer and brand.

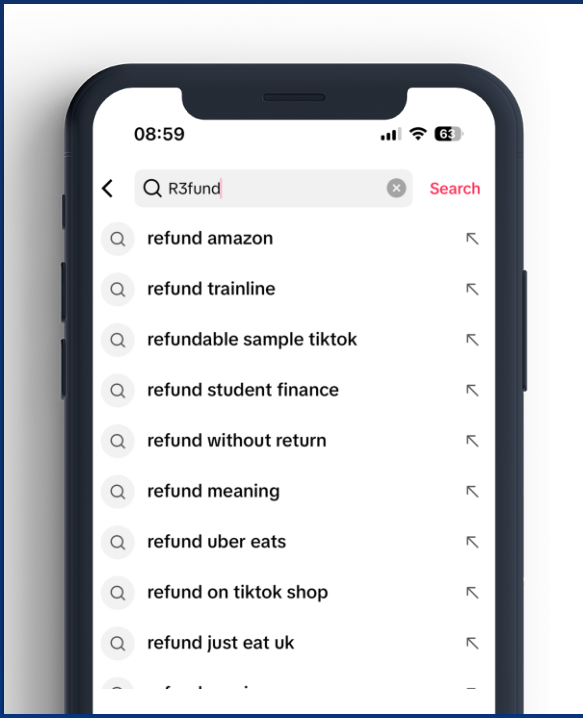
There’s also a more sinister reason refund fraud practitioners are beginning to hit mainstream social media-recruitment. Although professional fraud groups have long advertised fraud-as-a-service (FaaS) in exchange for a cut of the profits, these ads remained in the murky domain of the underground ecosystem-hacker forums on the deep and dark web and invite only Telegram channels.

Their migration to mainstream social media is not only a sign that organized fraud operations are expanding, but the target audience on these channels is also a virtually unlimited source of ‘clean’ user accounts.

Such accounts haven’t yet been flagged for suspicious activity on industry fraud and risk tools. This tips the balance for criminals trying to circumvent security and anti-fraud controls by reducing the number of potential threat signals.

By normalizing fraud to the TikTok community and getting Millennials to “sign up for easy cash”, organized criminals can effectively recruit humans and their marketplace accounts as digital mules for malicious activity. Many of these young people probably don’t understand that what they’re doing could land them with a criminal record or at the very least damage their credit rating.

And there is clearly an appetite within younger, tech-savvy demographics to participate-multiple surveys of American Millennials and Gen Zers over the last few years reveal that around half of respondents have admitted to committing refund fraud.



The point is, cyberfraud is becoming more accessible and the groups behind it are getting more organized. We’re not only seeing more of a presence from the Russian ransomware gangs considered highly dangerous due to their links with money launderers and organized crime syndicates, but we’re also seeing fraud groups adopt the operational practices of these organizations to optimize activities.

This is something we refer to as the ‘businessification’ of fraud because in many ways, these organized crime gangs operate very much like legitimate businesses. They have functional departments and implement rigorous application processes for new recruits, often requiring proof of previous fraudulent activities or a recommendation from a current member.

They have a management layer that dictates strategy and issues action points. Middle managers assign tasks or activities to fraudsters who work in isolation and remain anonymous to the rest of the gang. These workers are issued a target corporation, a quota to hit, and methods by which to defraud the target, and their performance is assessed each day. Failure to meet quotas is not tolerated, resulting in swift expulsion from the group.

These groups practice mature SECOPS and OPSEC, have HR departments, onboarding teams, employee progression and incentivization schemes and their marketing operations, including advertising, are focused on recruiting new blood through promises of easy money and a victimless crime.

For fraud teams and security teams in retail enterprises, these trends are very concerning, but at least they are happening outside of your defensible perimeter. So, what happens when you've got an enemy on the inside?

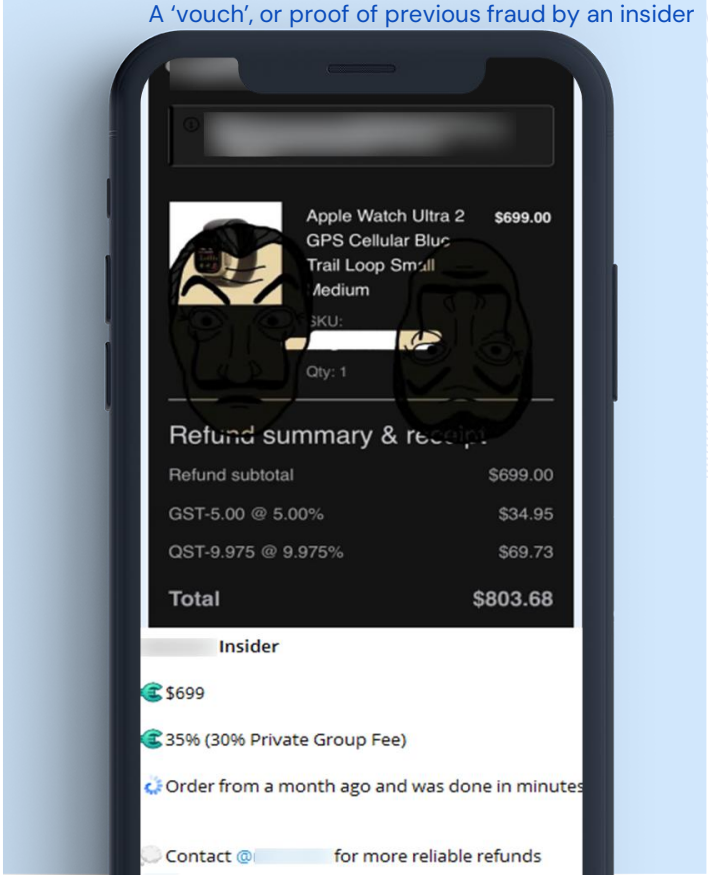
The FaaS and the Furious: Unpacking the insider threat

Corrupt insiders are a significant threat to enterprise brands as they can manually bypass the checks put in place to prevent refund fraud. They are recruited on the promise of significantly increasing their legitimate salaries and it's not unusual to see insider fraudsters taking up to a 35% cut of each refund transaction value. On a $1,000 Apple Watch that's a significant temptation for shouldering a small amount of risk in many cases.

These fraudsters are able to manipulate invoices, shipping labels and other documentation, circumventing controls put in place for the fulfilment process. They are notoriously difficult to detect because they operate in the human layer of business logic where digital tripwires are of little use. Detection is further compounded by the fact that fraud tends to be a lagging indicator of a problem and is often not appropriately identified until several financial quarters later.

Just one corrupt individual can cause significant damage. In one investigation the Netacea Threat Intel Center carried out for one of the biggest retailers in the US, a single insider was found to be facilitating refund fraud across no less than six organized fraud groups.

A 'vouch', or proof of previous fraud by an insider



This is known as Fraud-as-a-Service (FaaS) and is very typical of trends we're seeing across the fraud ecosystem, where every aspect of the fraud lifecycle or kill chain is available for rent from a specialist provider. There is money to be made selling enablers for each step in the process, or for wannabe career criminals to piece together a sophisticated end-to-end operation of their own.

Let's dig into the methods used by fraudulent insiders specifically:

Returns fraud Methods:

Returns Fraud methods are used by refund fraudsters to simulate returning an item, without actually doing so. These methods are used when the refund fraudster is claiming a refund for reasons such as receiving a wrong or damaged item, and the store requires the original item to be returned. Refund fraudsters generally employ third parties to assist them, including Boxing and Scanning services.

Boxing Services

Boxing Services, or Boxers, perform label manipulation and postage for refund fraud services. The refund fraudster provides the original postage label and, if necessary for the delivery carrier, any weight and dimension requirements to the Boxer.

Fake Tracking ID

Boxers are primarily used for the Fake Tracking ID (FTID) refund fraud method. FTID is where the return postage label is altered and used to mail an empty or junk filled package instead of the item for which the refund has been requested. There are two dominant versions of FTID currently in use.

In the first, all information linking the package to the customer or order is removed. This is intended to cause the return center to throw out the junk package and prevent them from tying it to the customer. At the same time, the delivery tracking will show the package as having been delivered, entitling the customer to their refund.

In the second and more widely used method, the delivery address is also modified. Here, the intention is that the package is delivered to an unrelated address and the recipient throws out the junk package, removing evidence of the fraud. The delivery tracking will show that the package was delivered to the return center, entitling the customer to their refund.

Lost in Transit

Refund fraud services may also employ Boxers when using the Lost in Transit (LIT) refund fraud method. Boxers print the postage labels using special disappearing ink, which fades over time. The postage label will be visible when the package is scanned in by the delivery carrier and tracking will record that the package is in transit. After some time, the label will fade and the delivery carrier will no longer be able to deliver the package, causing the tracking to eventually be marked as lost in transit. However, refund fraudsters currently prefer using scanning services for LIT refund fraud.

DNA

The DNA (Did Not Arrive) method is a universal social engineering tactic where individuals falsely claim non delivery of received packages, leveraging carrier systems to target items like electronics or clothing.

A related method, the Empty Box method, involves claiming that a delivered package was empty, deceiving retailers into issuing a refund or replacement.

Double Dip

The "double dip" refund method involves exploiting the process of returning an item for a refund, then claiming a second refund through another method at the same time, allowing the fraudster to keep both the product and the twice the money.

Refund with Reship

A "refund with reship" means claiming a refund for an order while also requesting a replacement, effectively receiving both the refund and the item.

Scanning Services

Scanning Services assist refund fraud services by abusing inside access at delivery carriers to fraudulently manipulate tracking information. This allows packages to be marked as LIT, damaged, or returned to sender (RTS) when they have actually been delivered to the intended recipient.

Pricing for these services ranges from £25 to £150 depending on the scan code requested, the service provider and the postal company.

Scanning Services assist refund fraud services by abusing inside access at delivery carriers to fraudulently manipulate tracking information. This allows packages to be marked as LIT, damaged, or returned to sender (RTS) when they have actually been delivered to the intended recipient.

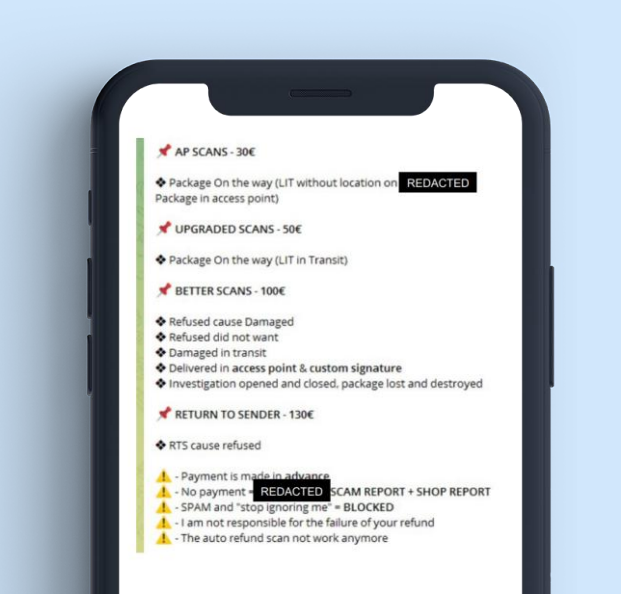
Pricing for these services ranges from $25 to $150 depending on the scan code requested, the service provider and the postal company.

Insider access to delivery carriers via access point accounts or sub-accounts is also traded on underground forums for between $100 and $750. Our research suggests that these accounts may be being taken over through credential stuffing, using tools such as OpenBullet. Once access to an administrator account is obtained, many sellers will create multiple sub-accounts to be resold to expand their opportunity.

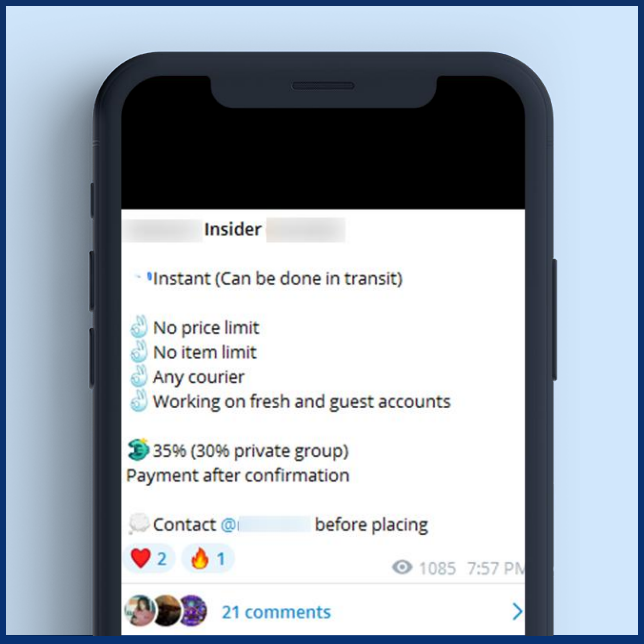
Most refund fraud service providers now prefer to use scanning services over Boxers to lend weight to their social engineering attempts.

For Example,

A scanning service can make it look like a package was refused and returned to the store it was purchased from by combining a RTS or damaged scan with a delivery scan, using the store's address and a fake signature. This strengthens the refund fraudster's case for a refund to be provided for their customer.



Scanning services advert on Telegram



Advert for insider access to multiple delivery carriers on Telegram channel

Fulfilment fraud

Many of the fraud typologies outlined above take place within the fulfilment process, specifically at courier companies used by the retailer.

Additional issues we have surfaced on behalf of large retail brands include:

01.

Fraudulent document acquisition

Fraudsters without valid identification documents or driver's licenses are using Telegram groups to obtain counterfeit or stolen documents. We're also seeing AI deepfakes used by early adopters.

02.

Order manipulation using bots

After registering as a courier, fraudsters employ bots to secure the highest-paying and most desirable delivery orders to get access to more valuable shipments.

03.

Account deactivation and reactivation

Drivers who frequently cancel less-profitable orders risk having their accounts deactivated due to low performance. However, there are Telegram groups that provide services to quickly reactivate deactivated accounts, bypassing standard penalties.

04.

Bypassing verification systems

When accounts are being created this generally requires facial verification checks. These groups are able to bypass this with video recordings and images which could be assisted by Machine Learning.

As you can see, there is a highly sophisticated social element to today's most prevalent fraud strategies, which already gives some indication of the answer to the question:

Why is fraud so difficult to manage?

Online Fraud Prevention (OFP) strategies have typically focused on preventing fraud at online checkout. They are constrained to the digital user journey on the retailer's website or app and tend to focus on answering the question,

"Is this a legitimate customer using my website right now?".

In this respect they are focused primarily on user behavior and fail to consider the ultimate intent of that visitor. But as we have seen, adversaries are now using business logic attacks to exploit opportunities at any point in the business processes – digital or otherwise - to deliver a favorable outcome for the fraudster.

So, the threat now goes beyond the traditional digital concerns of Account Take Over (ATO) and credit card fraud, to encompass broader challenges like brand impersonation, counterfeiting, fake websites, and refund or post payment fraud.

Purely from an operational perspective this puts retailers on the back foot because fraudsters don't operate in silos. They look at the entire landscape for opportunities and risk is largely dispersed among several enablers.

Enterprises, however, do operate in silos and fraud and cyber teams have not historically shared information, tools and knowledge. This means losses that should be attributed to fraud can become a lagging indicator of a problem, as the issue may be misunderstood as a 'cyber threat' and dealt with by a different team entirely or not appear on the balance sheet until months later.

To address this shifting trend, practitioners of online fraud prevention must adapt and incorporate new capabilities to create a holistic layered defense. A five-stage plan has been outlined by Gartner and while it is centered on the retail sector, it applies just as well to many adjacent industries such as hospitality, healthcare, gambling and marketplaces.

Five elements of online retail fraud prevention success

- Incorporate fraud intelligence
  Threat Intelligence
- Mitigate account takeover
  Bot Management
- Protect against automated business logic abuse
  Bot Management
- Prevent marketplace collusion
  Threat Intelligence
- Identify/block post payment fraud
  Threat Intelligence

The five stages of online retail fraud prevention

01

Establish Fraud Intelligence

Security and fraud teams need to look beyond the typical confines of the digital customer journey and address the expanding attack surface for fraud.

Threat signals gathered from outside the OFD (Online Fraud Detection) platform can add important contextual insight on historical, current and future threats, help eliminate intelligence blind spots and enable automated preventive measures.

Example

Detecting shipping labels manipulated prior to item return, intended to ensure tracker confirmation without sending the items which are then sold on.

02

Mitigate Automated Account Takeover (ATO)

This leads on to mitigating ATO, which remains one of the top vectors for perpetrating online fraud. Although many retail organizations are reluctant to introduce any significant friction to the user journey, resulting in limited use of multifactor authentication (MFA) and one-time passcodes (OTPs), it's possible to layer passive controls that introduce little to no friction.

Solutions include behavioral biometrics, behavioral and user intent analytics and device security posture assessments.

Example

Detecting malicious activity once a user is authenticated to a compromised account.

03

Protect Against Business Logic Abuse

Adversaries are increasingly taking the time and effort to study the business logic used by their target retailers with the intention of exploiting loopholes for their own gain. Within the confines of the digital user journey, common concerns include automated abuse of promotions or incentives, loyalty programs, and inventory depletion.

Beyond digital interaction, business logic abuse extends to things like refund fraud and postpayment fraud.

Business logic abuse is considered a significant threat because it is very difficult to detect. Essentially, the application or process is behaving as expected, even though it is being manipulated to give the attacker a favorable outcome.

Example

Detecting automated ordering of products accompanied by a free gift of some value when the product is returned for a refund, but the gift is kept for resale.

Or, automatically creating large numbers of free digital accounts (Fake Account Creation) in order to perpetually take advantage of sign up offers and services.

04

Mitigate Marketplace Collusion

Marketplace fraud impacts a much broader segment of the ecommerce vertical, including retail, hospitality, food delivery, gambling, ridesharing and other gig-based services. Marketplace vendors are mostly independent small businesses that live and die by their reputation. Meanwhile, consumer confidence in the marketplace as a whole can be quickly eroded by prevalent bad actors.

Security practitioners dealing with online marketplace features should look for solutions such as graph networks and behavioral and intent analytics tools that can help detect and identify marketplace collusion.

Example

Spotting fake customer reviews of a product or business to manipulate its reputation.

05

Protect Postpayment Processes

As stated above, most anti-fraud systems focus on prevention of fraud at the point of purchase or operate exclusively within the digital customer journey. Online fraud prevention initiatives for retail and ecommerce use cases need to extend support for events that occur after the purchase, such as returns and concessions.

The National Retail Foundation estimates that between 10% and 14% of online retail returns are fraudulent, accounting for tens of billions in losses.

Example

Detecting shipping labels manipulated prior to item return, intended to ensure tracker confirmation without sending the items which are then sold on.

The fusion of cyber and fraud teams

To better deal with this cross-functional threat it is expected that cybersecurity and fraud prevention budgets will eventually be combined as stakeholders and teams using these tools become increasingly aligned and integrated.

In fact, by 2028, analyst Gartner expects that 20% of large enterprises will shift to cyber-fraud fusion teams to combat internal and external adversaries targeting the organization, up from less than 5% in 2023.

Cyberfraud fusion centers are the future of fraud fighting and will be part of the strategic direction of retail and ecommerce brands over the next decade, combining elements of traditional cybersecurity, cyberthreat intelligence and fraud prevention, breaking down operational silos to create a more comprehensive approach to defense and response.

Using Fraud Intelligence to fuel fraud decisioning engines

Threat Intelligence Services (TIS), also known as Cyber Threat Intelligence (CTI), are well established in the cybersecurity domain for keeping practitioners informed about historical, current and future threats.

They include a range of Open Source and proprietary information regarding threat actor Tactics, Techniques and Procedures (TTPs) and Indicators of Compromise (IoC). Some TIS offerings will provide malware analysis and insights, and those specializing in online fraud will do the same for automated business logic attacks and bots.

Some cyber analysts and their counterparts in the anti-fraud function will be used to getting specialist information from Digital Risk Protection Services (DRPS). These services include monitoring and takedown for fake or counterfeit brand websites, detection of Account Takeover (ATO) attempts, and monitoring of chatter on adversary forums and groups on the deep and dark web and social platforms like Discord and Telegram.

Some of the more comprehensive offerings will include phishing mitigation, social media monitoring, and even specialist protection and training options for high-profile stakeholders like VIPs and executives.

We're expecting to see an increasing convergence of these two intelligence sectors in line with the fusion of enterprise cyber and fraud teams. The resulting output of both intelligence disciplines will create what Gartner calls Fraud Intelligence, and signals created from this analysis should be fed into the enterprise's fraud decisioning engine to help determine, and even predict, if activity is malicious.

Understanding the cyberfraud kill chain

Alongside intelligence services, analysts and practitioners of ecommerce fraud prevention should leverage specialist attack and defense frameworks to build rules, policies and probabilistic models to assist in cyberfraud prevention and mitigation.

The open-source OWASP Business Logic Attack Definition (BLADE) Framework Project, originally created by Netacea researchers in 2022, has become the industry-standard attack framework for combatting cyberfraud.

The OWASP BLADE Framework is an attack framework in the vein of MITRE ATT&CK and the Lockheed Martin Cyber Kill Chain. But where these frameworks focused on traditional network intrusion type attacks, the OWASP BLADE Framework filled a vacuum around enterprise business logic attacks.

Originally, the OWASP BLADE Framework focused on helping security analysts identify the six stages of complex, tailored and targeted attacks within the digital user journey – how malicious visitors to enterprise websites, mobile apps and APIs can manipulate the business logic to behave as intended but provide a favorable outcome for the attacker. But the framework can also be used to identify and understand attacks across the set of rules and decision-making processes that comprise a company's business logic. This covers everything from manufacturing to distribution and fulfillment – including the logic that governs how orders are processed, inventory is managed, and goods are distributed.

With the rise in fraud typologies involving company insiders and supposedly trusted third parties like fulfilment specialists, a framework like OWASP BLADE can be invaluable for identifying vulnerabilities and malicious activity in the 'soft' or human layer of business logic.

BLADE, Business Logic Attack Definition Framework.

Find out more about the OWASP BLADE Framework Project

What are kill chains, phases, tactics, and techniques?

Although business logic attacks typically have an overall objective, they are made up of multiple stages or phases, with each prior action designed to set up a subsequent step.

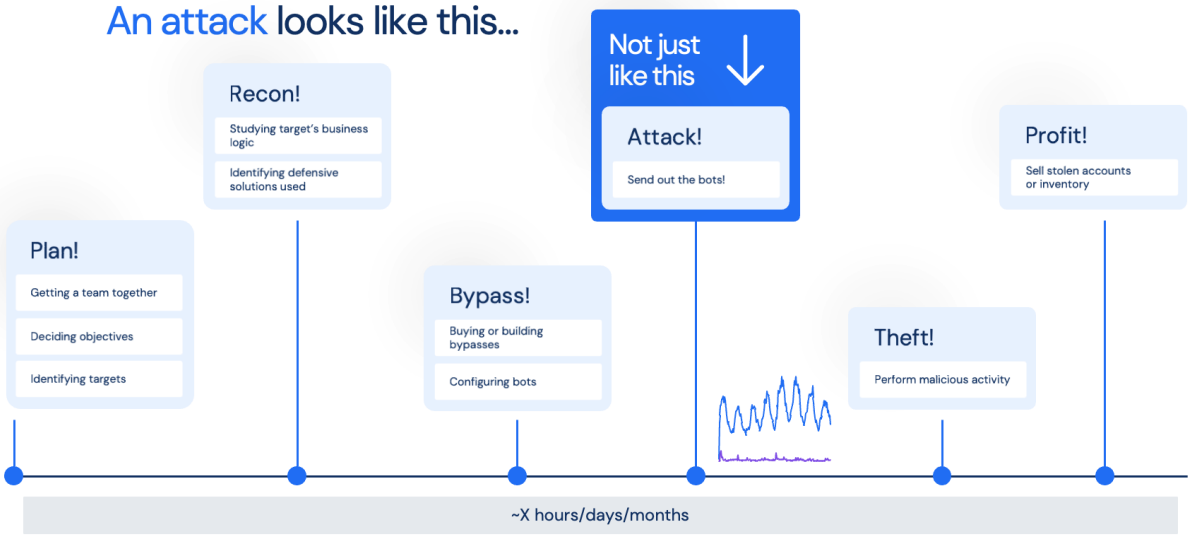
For example, to achieve the general goal of a scalping attack, the attacker may first need to scrape product pages continually to pinpoint the exact second items go on sale (a technique), rotate their IP addresses and create multiple fake accounts to bypass purchases-per-customer limitations (tactics), and automate 'Add to cart' and purchase activity – this sequence is what we call the 'kill chain', representing the overall lifecycle of an attack objective.

Outside of the digital user journey, Scanning Services are an example of a business logic attack designed to assist refund fraudsters by abusing inside access at delivery carriers to fraudulently manipulate tracking information. This allows packages to be marked as lost in transit (LIT), damaged, or returned to sender (RTS) when in reality they have been fraudulently delivered to the intended recipient.

The OWASP BLADE Framework is a useful mechanism for understanding the scope of business logic abuse and identifying opportunities for disruption beyond what a large swathe of the industry considers an 'attack' - something that is only happening when thousands of bots are active on their website.

The reality is that this is only the attack execution phase and is just one stage of the overall cyberfraud kill chain. There are five other phases preceding and succeeding this phase that provide more opportunities for disruption and the value of a framework like OWASP BLADE is in understanding this.

Visibility of fraud beyond the confines of the digital user journey.



Effective Threat Intelligence and Digital Risk Protection Services will be able to extract valuable insights and actionable information from the phases preceding the attack execution, including who the organized criminal group or adversary is, what activities they plan to carry out, and how they plan to do that, including what they know about their targets defenses and how they intend to bypass them. In terms of insight post attack, threat intelligence can help identify what was stolen or affected and where it was sold or profited from.

Intelligence services should be combined with some kind of real-time digital defensive solution that is capable of detecting and mitigating sophisticated abuse of your web estate's business logic. A dynamic anti-bot solution that can respond to automated attacks changing their TTPs without adding undue friction to the user journey and affecting legitimate visitors.

Armed with these capabilities, an enterprise cyberfraud team should be able to:

01 Forsee potential threats

02 Have the right defenses and countermeasures in place

03 Predict, prevent, or disrupt an attack before it is executed

04 Respond effectively to an attack that is executed

Turning cyberfraud from lagging to leading risk indicator

Andy Ash
CISO, Netacea

When I recently spoke at an event, I asked the audience of CISOs, "who in this room maintains a risk register?". As you might expect, most hands went up. I followed up with, "now who has bot management on that risk register?". Only a few hands left.

For most of those businesses, especially those in retail, the effect of fraud is a real and often realized risk. Automated fraud poses just as much financial risk as traditional attack vectors - just look at the numbers at the start of this report - so why are automated attacks (bots) not reflected on risk registers in the same way ransomware is?

As a CISO who talks to a lot of other CISOs, both at Netacea customers and in the industry at large, I believe the sophistication of a company and the maturity of its risk functions plays a large part in recognizing such risks.

At the start of this report, we investigate how fraud groups are highly organized and very detail driven. They understand the business logic and processes of their targets even better than employees and they seek opportunities for exploitation across the entire attack surface available to them.

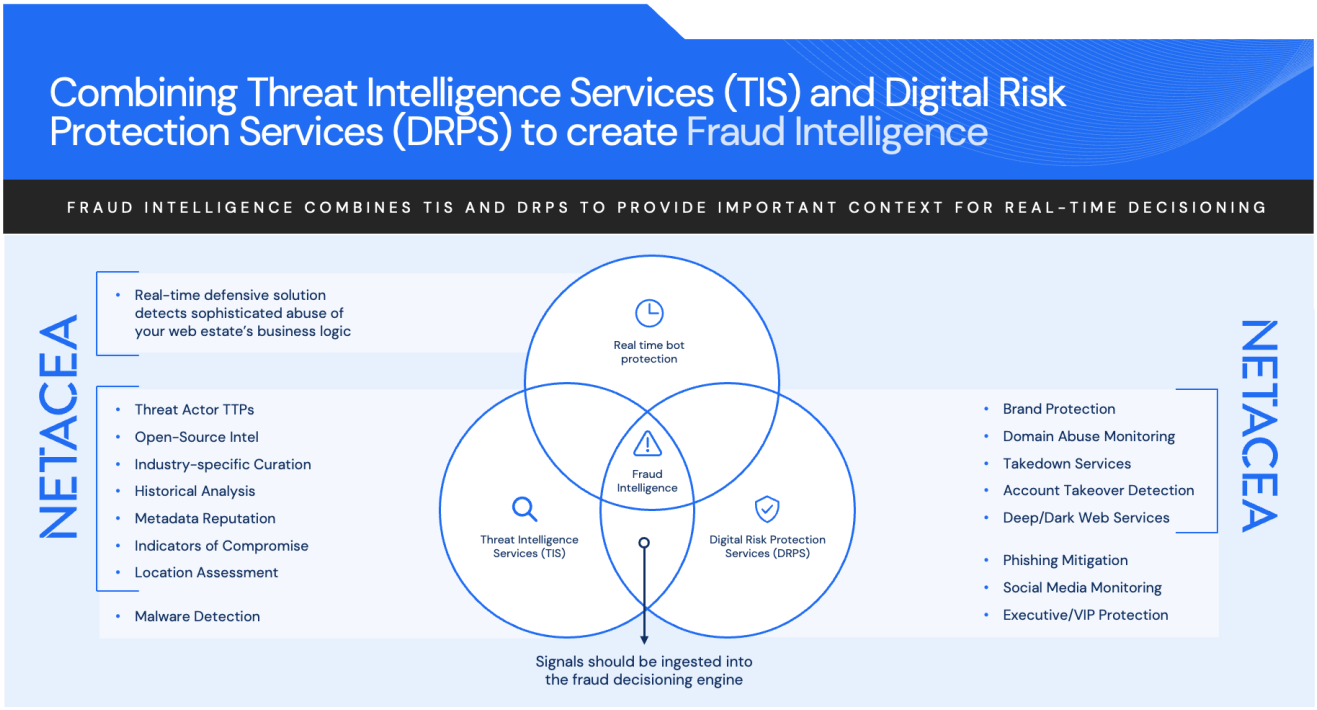
Many enterprises, however, have siloed functions that look at parts of a problem but struggle to see the bigger picture. It's these gaps between silos where fraudsters can slip through undetected. If your web infrastructure team picks up higher volumes of automated traffic crawling critical paths on your site but doesn't relay this info to the fraud or loss prevention team, you might have missed a strong suggestion of an imminent attack. And with fraud being a lagging indicator, by the time the issue is confirmed it's often too late. Change is starting to happen, however, and some of the largest retailers we work with are actively pulling their cyber and fraud teams closer together to address this very issue.

These companies with a more sophisticated attitude to risk also recognize fraud as having much broader implications than just monetary loss. One of the top US retailers we work with halted spiralling infrastructure costs from serving unwanted web traffic – we're talking billions of requests here – by looking at the bot problem. Not all this traffic was malicious in the sense that bad guys were trying to break in, but their inventory systems were heavily monitored by scalpers waiting for the next drop of high-value goods, and culling this traffic saved money and possibly their brand.

This is another reason cyberfraud should be seen as a wider challenge – it erodes your brand sentiment with consumers. Anyone who's had their loyalty points stolen or had to pay over the odds for a ticket, or the latest console or sneakers knows what I'm talking about, and I'm certain brand sentiment is on a good number of risk registers out there.

There could be a new trend that pushes bot management up the ladder however, and forces companies to pay attention to their non-human visitors. With AI embedded in everything and AI agents popping up everywhere, we're going to see even more non-human traffic on the web, a lot of it completely machine-to-machine, even agent-to-agent. What's interesting is that we've been thinking about this phenomenon for a long time in the building of our Intent Analytics engine.

Given the changing nature of traffic, identifying whether an identity is human or not seems redundant. It's really about exposing that entity's intent.



Understanding and disrupting malicious intent at the human level and the machine level

Netacea is a specialist in cyberfraud prevention and expert in detecting and disrupting automated business logic attacks. The humans running our Threat Intel Center are also the founders of the OWASP BLADE Framework, which since launch has become an industry standard in understanding business logic attacks and is used by some of the biggest names in the tech industry as well as our competitors.

Using Machine Learning models, AI automation, and human expertise, Netacea analyzes tens of billions of malicious requests daily across some of the world's busiest and most criminal - targeted web estates.

It's worth remembering there is a human adversary behind every attack. Bots do not act without human direction.

The BLADE Framework enables security practitioners to understand where adversaries are in the lifecycle of their attack plans and helps piece together the kill chains of their tactics, techniques, and procedures. But that information needs to be acted on to be useful and to create an effective defense.

Consider that malicious intent is decided at the human level – by the adversary – at the moment they form their gang or decide on their objective. Will they be stealing customer accounts or data or scalping high profile consumer goods, for example? Deciding on this objective will confine the attackers to a standard set of attack techniques and a predictable kill chain combining these techniques with various tactics. This is useful information for a defender to have.

The fraudster's intent also establishes the boundaries for target businesses they can go after and the next phase in their planning lifecycle is to analyze the business logic used on those websites, mobile apps and in some cases APIs to find opportunities for manipulation or hijacking. This is also the point at which adversaries will seek to detect any bot management solution in place on the website.

Bypass methodologies and configuration files are easily found on the open internet for almost all bot management solutions, from workarounds for the trivial rules-based approaches used by some WAF vendors, to reverse-engineering of the client-side agents deployed by most specialist bot mitigation solutions.

Once these tools are in the hands of the adversary, intent has been set at the machine level and the bot behavior is programmed in. Tens of thousands of bots now have their mission and once released they will endeavor to complete that objective relentlessly, with the more sophisticated automations changing their behavior in response to dynamic defensive solutions encountered.

It is at this point that the benefit of threat intelligence cannot be underplayed. Insight from within the adversary community can give retail brands a heads up on exactly when and how organized criminals are going to attack. In some cases, threat intelligence teams can infiltrate these fraud gangs, gain the trust of the leadership, and socially engineer a disruptive outcome.

In one example, the Netacea Threat Intel team learned of a project to develop anti-bot bypasses for specific solutions within the botting community ahead of a high-profile ticketing event that was seen as a significant opportunity by scalpers.

Our operatives infiltrated the community and socially engineered themselves into a trusted position by demonstrating a comprehensive knowledge of the anti-bot ecosystem. We maintained this position through the development of the bypass modules and even through the sale of those modules to other members of the criminal ecosystem.

It was only on the eve of the opportunistic attack that we took action to help ticketing vendors close the security gaps, so when adversaries launched their automated attacks with high hopes, they were met with failure and the fraud forums immediately lit up with angry chatter.

This approach not only thwarted the attack itself but burned the reputation of the bypass creators and put their customers out of pocket in terms of money they spent on the bypass modules and resources invested in 'the big heist' that didn't happen.

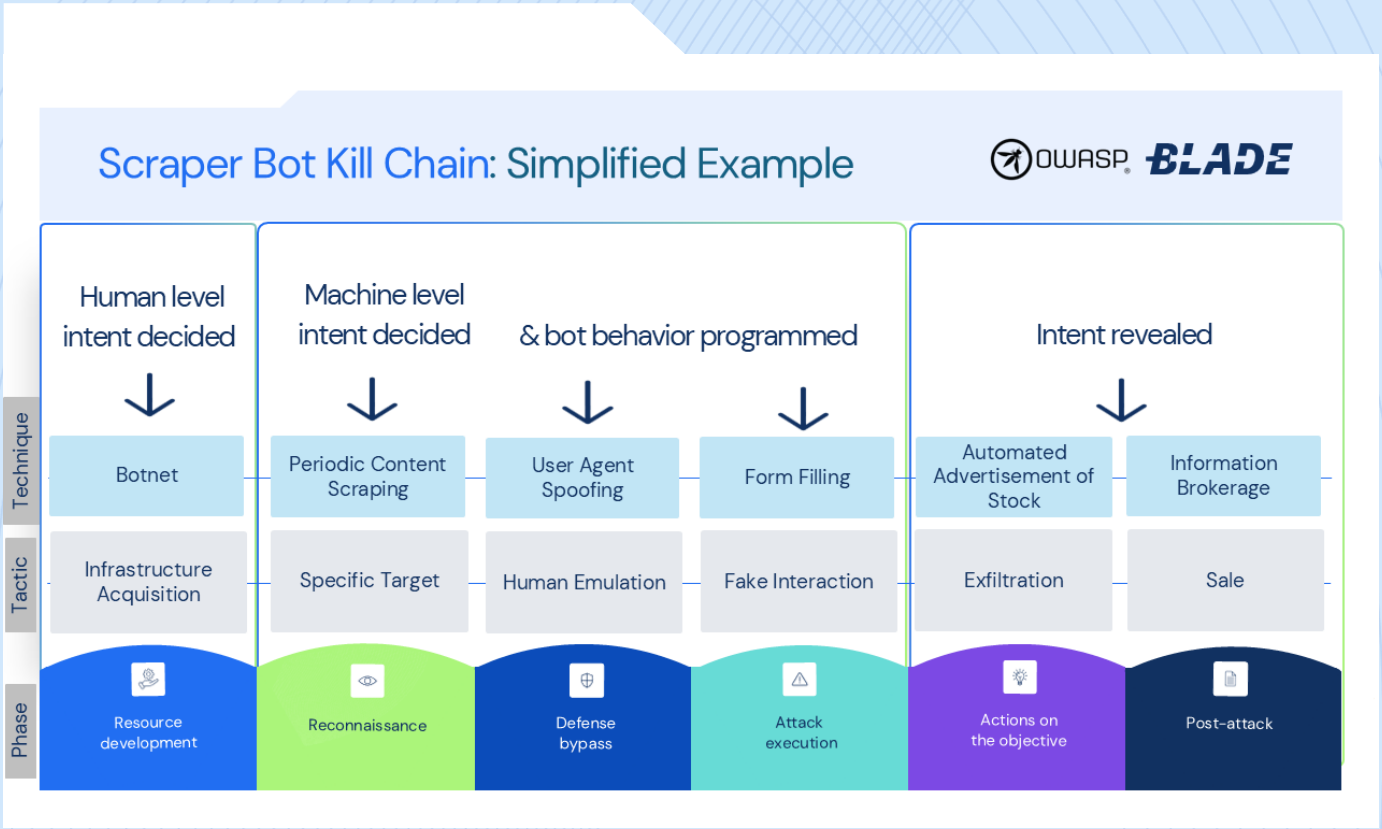
The point is that defensive approaches focusing solely on the attack execution phase do not benefit from this insight and give retailers only a single opportunity to stop a fraud operation from being successful. Without an intelligence-first approach you're leaving actionable insight on the table.

That's not to say the attack execution phase isn't important. If your defenses don't perform here then you suffer the consequences, and no amount of foresight or hindsight can retrospectively change that. But making effective use of that foresight can give you multiple opportunities to stop an attack being successful and limit potentially negative outcomes 'on the day'.

You can even extract insight from the post attack phases when adversaries are taking actions on their objective by stealing accounts, data, or inventory, and then profiting from its sale on dark and gray markets. Forensic analysis at this point can reveal past breaches and highlight existing exploits for a retailer's business logic assisting in prediction of future attacks.

None of this undermines the importance of having an effective real-time anti-bot solution in place, however. Armed with all the foresight in the world, some of today's highly sophisticated automated attackers can change tactics and techniques dynamically or exhibit benign behavior until they make it past rudimentary defenses.

So, even if you know this kind of complex attack is coming, you need a flexible and adaptive detection and mitigation solution to stop it being successful.



The evolution of bot defense and attack technology

It's impossible to ignore that AI has changed the game. We may have crested the peak of inflated expectations and be on the downhill slope to the trough of disillusionment with regards to what AI is capable of today, but there's no escaping the fact it is and will continue to have an impact on offensive and defensive technologies.

The most prevalent threat in today's landscape is what AI has delivered in terms of scalability. Effective automation means a single human adversary can be responsible for tens, hundreds, even thousands of attacks. On the other side, defensive AI is the only way blue teamers can handle the flood of alerts they are drowning in and automate standard responses so they can focus on the real issues.

And as the technology matures, the ongoing game of cat and mouse between red and blue teams is going to become more reliant on autonomous learning and response.

What exactly is a sophisticated bot attack? Read the explainer.

Find out more

The OWASP BLADE Pyramid of Pain

The original Pyramid of Pain was created in 2013 by David J Bianco, a cyber threat evangelist and SANS Institute instructor now working for Splunk, to illustrate the level of pain inflicted on an adversary when the indicators or techniques of an attack are countered by defenses.

For example, if the IP address being used in the attack is blocked it is easy to move to another IP address, but if the tool itself is identified and blocked it is more challenging to continue the attack.

The Pyramid of Pain originally focused on network intrusion and TTPs sat at the apex, with tools and Indicators of Compromise beneath.

At this level you were operating directly on adversary behaviors not just their tools. Forcing adversaries to learn new behaviours was the most painful outcome and the ideal from a defense effectiveness standpoint.

Read the explainer for the original Pyramid of Pain.

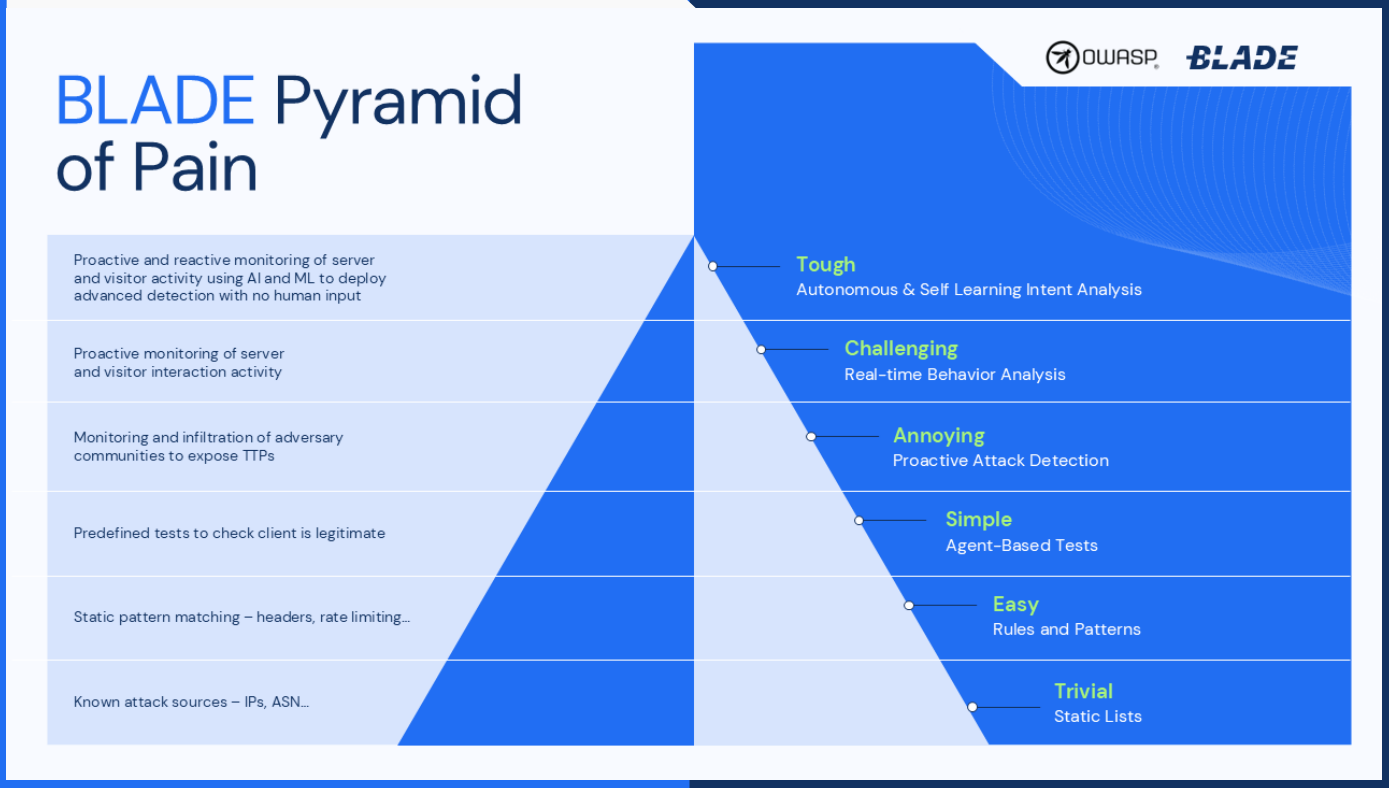
Find out more

But in the Pyramid of Pain reimagined within the OWASP BLADE Framework, to make it relevant to business logic attacks, TTPs sit just above the mid-point. This creates a new apex highlighting the more dynamic and autonomous nature of innovations in both attack and defense concerning business logic abuse.

Exposure of TTPs is only annoying in nature because automated attacks are sophisticated enough to change their behavior dynamically and in real-time in an attempt to evade defensive solutions.

Note:

The OWASP BLADE Pyramid of Pain considers defense techniques in isolation. Best practice requires layering or the adoption of multiple techniques for comprehensive protection – otherwise known as as Defense in Depth. A threat feed of malicious IPs alone will not protect you against sophisticated attacks but can significantly thin the ranks of a bot army and prevent more expensive defensive resources being spent on trivial techniques.



Innovation at the top of the Pyramid

This version of the Pyramid of Pain shows how attacker innovations in bypassing defense have prompted specialist bot detection solutions like Netacea to look at real-time analysis of visitor behavior and intent.

Lower down the Pyramid, you see that WAF and CDN bot protection modules are simple or easy for attackers to overcome, as they rely on rules and patterns or agent-based tests at the point of first request.

Specialist bot protection solutions really put a flag in the ground for where the industry is today by making use of threat intelligence to build in proactive attack detection, which is annoying for an attacker, while real-time behavioral analysis of website visitors makes the lives of adversaries more challenging.

But the top of the Pyramid is autonomous and self-learning analysis of visitor intent. This is where we get into the realm of predictive defense and the ability to detect and mitigate a malicious visitor even when they are behaving benignly.

This approach, able to detect real-time shifts in behavior and expose malicious intent, is where we make life hardest for even the most committed adversary.

Today, the top tiers of the pyramid are where we find only the most determined and well-resourced criminal adversaries. But in the years to come this is where we will see the longer-term impact of offensive AI, driving autonomous experimentation to respond to evolving defences.

In this way, the BLADE Pyramid of Pain serves as a mechanism to demonstrate that through Netacea's patented innovations in autonomous and self-learning intent analysis, the blue team currently has the edge and will continue to do so even as levels of sophistication increase.

Interested in the OWASP BLADE Framework Project?

Find out more

BLADE Pyramid of Pain

The four generations of bot management

Experts at Netacea have witnessed several generations of evolution in bot management as the offensive and defensive capabilities outlined above respond to each other.

1

The first generations of bot defense were based on static checks for bot-like behavior and on blocking visitors that failed those same checks. While the complexity of the checks continued to increase the underlying premise remained the same – test for pre-determined behavior and flag any visitor that fails the test as malicious.

This approach gave way to the first generation of static lists and rules and is the approach still used at a basic level by WAFs today. You essentially compile a list of known bad actors, either from reputational sources or those that failed the tests outlined above, usually based on IP address, and block visitors on that list.

This approach can be circumvented by simply rotating to a new IP address not on the list, which took anti-bot thinking to the second generation and a re-evaluation of the premise of detection.

2

If the attackers are mimicking legitimate visitor behavior, then the approach should be to analyze the device where the requests are made from. This led to the development of agent-based solutions which deploy within the client, usually by injecting some JavaScript into the browser, to validate the device and how the device was being used.

The challenge here is that machines are very good at pretending to be other machines and the signals ingested are not always indicative of malicious intent.

Client-side approaches can also be reverse engineered – the code is right there, even if it is obfuscated, and adversaries can work out what responses the anti-bot solution is expecting to receive in order to bypass the checks.

Somewhat counterintuitively, the agent-based approach also introduces security risks. Flawed updates can take large numbers of endpoints offline, as we saw with high-profile stories like the CrowdStrike outage of 2024.

3

Generation three moved the logical point of analysis to the server and is where Netacea patented its adaptive approach. By constantly analyzing all visitor activity in real-time based on each visitor's interaction with the server, we were able to identify known attackers and new indications of attack, even going so far as to being able to predict malicious intent and adapt defenses appropriately.

An additional benefit of the server-side approach is that the solution is invisible to attackers and cannot be detected, reverse-engineered or bypassed. In fact, Netacea Bot Protection remains the only solution without a commercially available bypass in the wild.

4

But we didn't stop there. In anticipation of future threats driven by offensive AI where Netacea expects self-learning models to deliver adaptive responses at machine speed, we are building and already deploying defensive AI and ML models that are capable of responding to and even predicting these evolving attacks.

AI automation is turning the tables on attackers

Andy Still
CTO, Netacea

The level of sophistication of automation we're seeing in adversary ecosystems suggests that attackers are generally winning in terms of defense bypass in many sectors.

Looking at web application protection specifically, CAPTCHA can now be easily bypassed through a range of on-demand services including human-powered CAPTCHA farms, and increasingly, with AI.

Similarly, device based anti-bot detection is hitting the limits of usability. Once the front line of anti-bot defence it is now bypassable with readily available tooling or even by request to an on-demand API bypass service. Machines are very good at pretending to be other machines. So today, device data is useful only as an additional but easily compromised data source.

These on-demand fraud services are representative of the rapid growth of an entire industry of products and services to define, control, launch, distribute and manage the spoils from large scale automated attacks. And the barrier to entry for any budding criminal kingpin is very low – limited technical knowledge required.

The platforms behind these services are like ERP systems for fraud management, even including distribution channels for sale of tasks to workers or mules who carry out a high-risk component of the actual fraud activity in relative isolation and sometimes with no real legal understanding of what it is they're doing. As the survey findings at the start of this report revealed, losses from fraud are seen as the cost of doing business.

This trend is evolving in parallel with the migration into cyberfraud as another revenue stream from higher profile criminal industries like ransomware, where it can provide large numbers of lower value and lower risk transactions.

There's clearly opportunity. Even the volatile impact of the US administration on consumer pricing plays into the hands of scalpers who thrive on scarcity of goods and already have the automated tooling in place to target even commodity items. Ultimately, the risk of prosecution is dwarfed by the potential benefits in cyberfraud.

Furthermore, the evolution and adoption of AI can only enhance the ferocity of this battle and therefore defense needs to be taken seriously in reflection of the risk.

Enterprises need to understand that there is no longer any 'easy' solution to automated attacks, like dropping a CAPTCHA or checking a device fingerprint. The only way to detect today's sophisticated attacks on business logic and processes is through end-to-end understanding of the attack typologies and AI assisted – moving towards self-learning - identification of malicious intent.

NETACEA

Malicious Intent Exposed

Netacea, the next-generation, enterprise-class bot detection and response specialist, provides a better way to stop bot attacks at scale.

Netacea is a recognized leader for its innovative use of threat intelligence and machine learning to deliver better detection of bot attacks across websites, apps and APIs.

Netacea's patented server-side integration analyzes all web traffic at the edge, providing comprehensive real-time protection through a single, lightweight integration that is invisible to attackers.

Contributors

Andy Still, CTO

Andy Ash, CISO

Matthew Gracey-McMinn, VP Threat Services

Cyril Noel-Tagoe, Principal Security Researcher

James Middleton, Director Product Marketing