Transmit Security
Research & Insights

JUNE 2026

ONLINE SECURITY WHITE PAPER

How to protect your customers in one of the toughest years of threat.

A practical, lifecycle-based framework for defending banks, fintechs and consumer platforms against the new generation of AI-enabled fraud and scams.

By Ken Palla
Consulting · Online Security & Fraud Prevention

EDITION 1.0

ABOUT THE AUTHOR

Ken Palla

Two decades on the front line of online security for financial institutions.

RECOGNITION

Legends of Fraud Award, FraudCON, Israel — 2019

Since 2005, Ken has been in Online Security. He was a Director at MUFG Union Bank, retiring in early 2019. He helped shape the initial responses to the U.S. 2005 and 2011 FFIEC Regulatory Guidance to improve online security for US Banks.

He is an early adopter and has selected and implemented a number of online security products. Ken was an advisor to the RSA eFraud Global Forum and a Program Committee member for the annual San Francisco RSA Conference. He was on The Knoble Scam Committee for three years.

He has published many white papers—on the need to focus on online customer safety, on online authentication and on how to select a multi-factor authentication solution. Most recently, his white papers and blogs have been on consumer financial scams. These recent white papers and blogs focus on controls to reduce scams and what countries are doing about scam reimbursement.

He also was the editor for the complete list of definitions of financial scams, published by The Knoble in 2022. In 2019, he received the Legends of Fraud Award at the 3rd annual FraudCON conference in Israel. He is currently consulting to banks and to online security vendors.

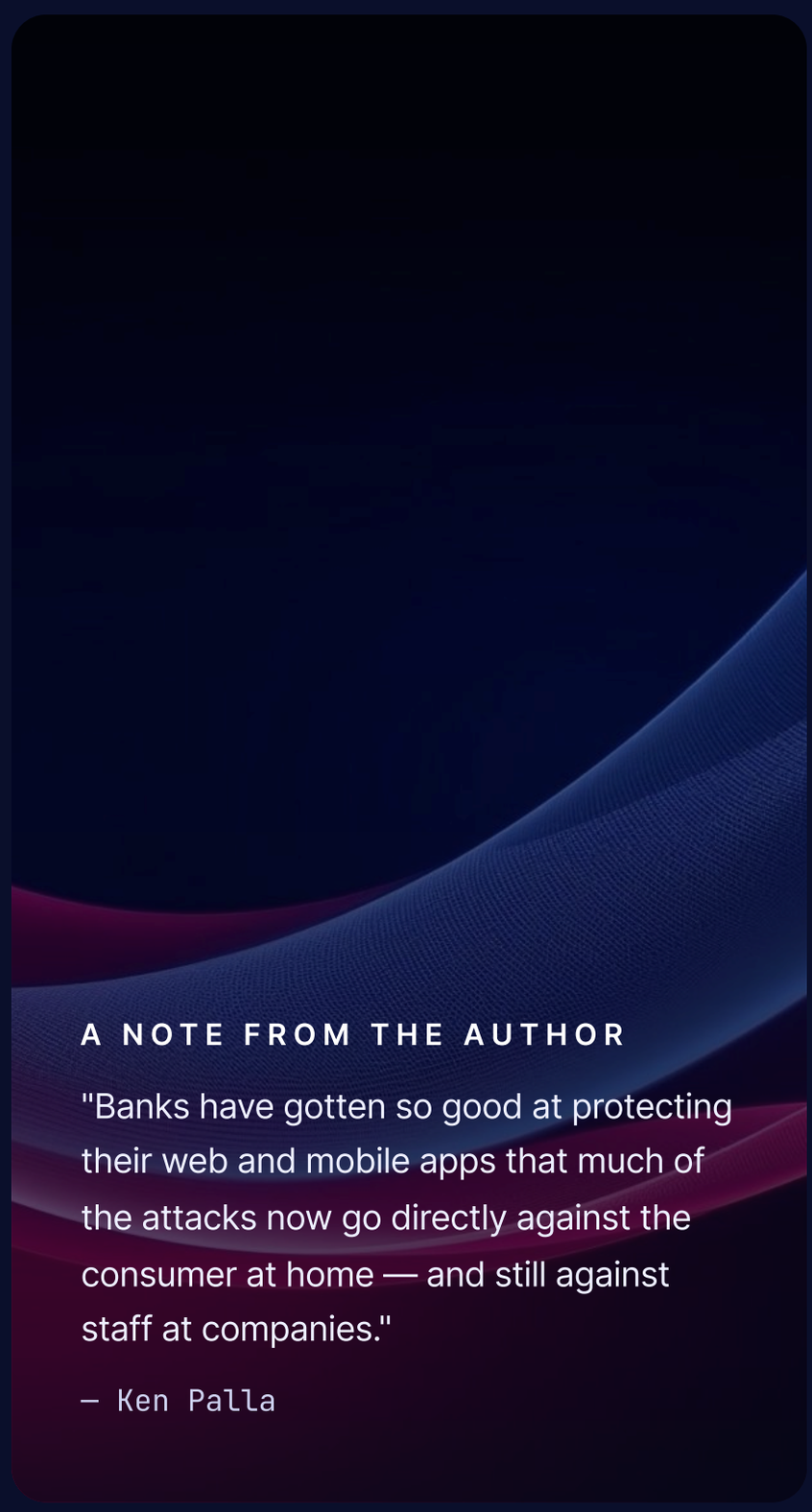
20+
Years in online security
Spanning FFIEC guidance, malware, ATO, scams and GenAI.

RSA
eFraud Global Forum
Advisor and program committee contributions.

2022
The Knoble
Editor of the published list of financial scam definitions.

OUTLINE

What you will find inside this paper.



A NOTE FROM THE AUTHOR

"Banks have gotten so good at protecting their web and mobile apps that much of the attacks now go directly against the consumer at home — and still against staff at companies."

— Ken Palla

EXECUTIVE SUMMARY

2026 has arrived like a winter blizzard.

Generative AI (GenAI), which only really entered the stage in late 2022, has become so powerful for the fraudsters. GenAI and the more recent Agentic AI component has the ability to create and execute fraud scenarios, and at scale, to really change the game for financial institutions, eCommerce, crypto exchanges, gaming and other entities.

Part of this change is the ability of GenAI to break existing security controls. Just look at account opening where the fraudsters can:

01 Efficiently use stolen PII data to at scale input the same applications at multiple banks.

02 Use graphic capability to create supporting government documents.

03 Create videos from a photo that can defeat liveness testing during the application process.

INTERPOL 2026 GLOBAL FINANCIAL FRAUD THREAT REPORT

"AI-enhanced fraud is 4.5× more profitable than traditional methods. 'Agentic AI' systems can autonomously plan and execute complete fraud campaigns, from reconnaissance to ransom demands. "

This requires companies to rethink and retest how well their existing security solutions are really working.

4.5×
More profitable: AI-enhanced fraud vs traditional methods (Interpol, 2026).

$20B
Estimated annual U.S. consumer scam losses today — with $150B+ exposure ahead.

600+
APIs managed by financial services firm on average (BAI); thousands at the largest banks.

This is compounded by the April 2026 announcement by Anthropic of the new Mythos model. This model has the capability to effectively find vulnerabilities in operating systems and application code. As an example, it is reported that Mythos could easily scan all bank customer facing applications and identify vulnerabilities. Anthropic has currently restricted this model to a limited number of entities.

There is also concern that in the next few years (2029-35) quantum computers could defeat existing encryption methods. Developing new encryption methods is no longer a long-term project.

On the positive side, we are starting to see how GenAI can be used to detect and prevent fraud and scams. There are AI agents talking to scammers to collect money mule account information, AI agents that help the fraud analyst interdict with customer/scam victim and plans for Gen to AI help detect fraudulent transactions and help with fraud analyst decisioning.

This white paper uses the lifecycle of the customer as a framework for discussing the online threats and recommended controls. The report also puts a strong emphasis on 'left of boom' controls. 'Left of boom' controls identify where companies can take action that will most prevent a fraudulent transaction from ever occurring.

THE FINANCIAL TIMES CONFIRMS THE RISK SEVERITY OF MYTHOS

"Senior international financial officials have warned the latest AI models from US tech companies could threaten the world banking system by exposing weaknesses in lenders' cyber defence."

Some examples of 'left of boom' controls:

01 **Strong online account opening** to prevent money mule accounts from being opened.

02 **Phishing resistant multi-factor authentication** to prevent account takeovers.

03 **Continual education** with companies to prevent business and consumer email/voice/video compromises.

Companies have limited money to spend on security. So, spending money 'left of boom' can have the most impact to:

1. **Reduce fraud and scam losses**, and
2. **Reduce operational expense** when losses occur (e.g. cost of closing bad accounts, attempts to recover the funds, creating SARS, etc.).

The full life cycle security solution stack needs to include two components that may or may not be part of a company's existing security stack. These two components are:

1. **Consumer scam prevention**
2. **Money mule detection and removal**

Consumer scam losses have grown in the past three years. This is because transnational organized crime (think of the massive scam compounds in Southeast Asia and elsewhere with hundreds of thousands of scammers) has realized banks have tightened their online security, but the consumer sits unprotected.

> In the US, reported scam losses are around $20 billion per year. The Federal Trade Commission estimated annual consumer scam losses could exceed $150 billion per year. The same amounts per capita are reported in other countries around the world.

Money mule accounts are one of the primary reasons (along with crypto exchange accounts and crypto ATMs) that fraud and scam transactions can successfully execute. Report after report by various countries show there are large numbers of money mule accounts available for the fraudsters.

As a result, commercially reasonable security and protection of the banking ecosystem and its customers requires that financial institutions have consumer scam prevention and money mule prevention programs. In some countries, like the UK, it is regulatory policy to have these programs.

This white paper will cover both online security threats and online security mitigation. The primary focus will be on financial institutions, but there will be some additional commentary.

One final thought is on the selection of vendors to help solve the online security problem. Look for vendors who are constantly innovating, offer multiple components of the solution stack and have partners that can help fill the solution stack. And remember, machine learning and AI are tools the vendor deploys, but focus on the specific solution features and how they can be measured.

Every few years, it pays to reframe what online security really means.

This author has been working in online security for financial institutions for over 20 years. Way back in 2005, it was a swirling mess as fraudsters had figured how to access bank customer accounts (on the web only at that time) with relative ease. After all, the primary (and often sole) online security control for banks was User ID and Password. The major newspapers in the US were constantly writing stories on how bank customers were losing money. These were losses due to unauthorized transactions initiated by the fraudsters themselves.

Around 2011, we started to see malware on commercial customer PCs causing losses and hear of business email compromises, where the CEO 'purportedly' tells the Finance team to send a wire to close an important deal. Then the email-type attacks expanded to vendor email compromises (the vendor telling the company 'about a new bank account to send payments for future invoices'.

And again, it expanded to retail real estate closings, where the buyer 'receives an email with a new bank account to send the closing funds'.

Many online security controls were added between 2005 and 2026 to mitigate a number of the most serious fraud attack vectors. Unfortunately, the business/vendor email compromises and the real estate transfer scams have continued unabated, because the attack is against the customer directly and not the bank online systems. To show how bad it is today, I was just at a US title company (part of a US mortgage transaction) and they could not stop warning me about wire fraud.

But now we come to 2026 (and even a few years before), and history is repeating itself where we are seeing a renewed volume of consumer losses. This time it is different for consumers. It is social engineering attacks against consumers causing authorized transaction losses (transactions actually executed by the customer).

A NOTE ON TERMINOLOGY

In this paper, the author has taken the liberty to identify all financial institutions (banks, credit unions, wealth management, fintechs, neobanks, etc.) as simply 'banks'. Much of what will be discussed is also applicable to other entities including eCommerce and crypto exchanges.

WHERE ATTACKERS FOCUS

Fraudsters target the largest banks (more customers) and the weakest banks (often small institutions forced to rely on third-party vendors to add controls).

What makes these scam attacks so effective is that they are often initiated by large transnational organized criminal groups and maybe even by nation state actors (e.g. North Korea). The scam playbooks are very well crafted and proven to effectively work again and again. In the US, it is estimated that almost $150 billion per year is lost by consumer scams. In many other countries around the world, there are commensurate loss figures.

Banks have generally gotten so good protecting their web and mobile apps, that much of the attacks now go directly against the consumer at home and still against staff at companies.

Also, since 2022, we have seen the rise of Generative AI (GenAI). This has been billed by OpenAI, Anthropic, Google and others as a powerful wave of innovation, similar to electricity and the telephone. As this wave builds momentum, we are seeing a tragic by-product of Gen AI being effectively used by fraudsters in 2025/6 to create quite serious cybersecurity attacks, including attacks against online banking.

And as always, fraudsters attack the largest banks (the most customers to go after) and the weakest banks (often small banks that are forced to rely on third party vendors to add controls on their own timeline to help fight fraud and scams).

It is with this knowledge this white paper was written in 2026 to help banks defend against fraud and scam attacks being crafted with the very best of new tools, often requiring much less skill on the part of the fraudster.

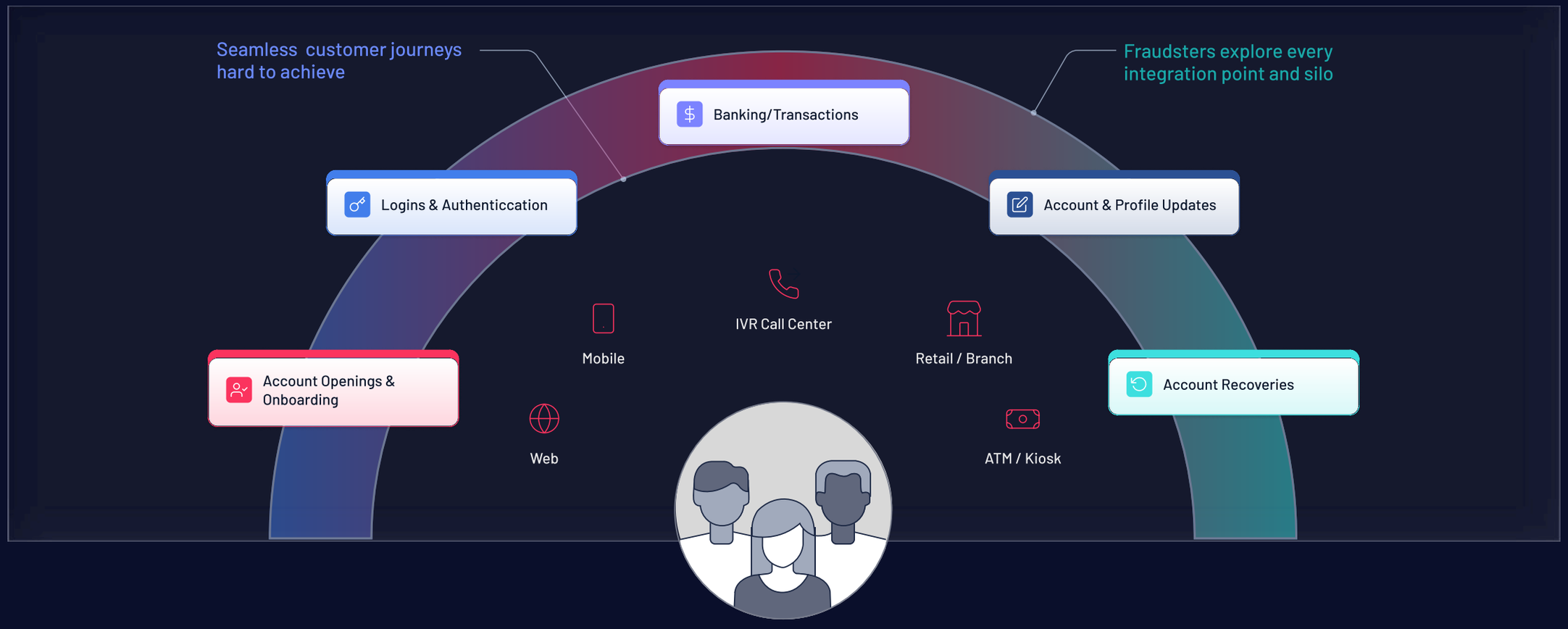
A QUICK NOTE

This paper will discuss:

1. **The major attack vectors**, including the more recent use of Gen AI for attacks
2. **The appropriate controls** to mitigate these attacks across the customer life cycle. The customer life cycle will involve web and mobile online as well as portions of branches and call center activities.

Some discussion beyond financial institutions will be covered.
The mitigations will span the customer life cycle, with a specific focus on 'left of boom' controls.

FIGURE 1: CUSTOMER LIFE CYCLE



HOW TRANSACTIONS ARE CREATED

Five ways a transaction reaches your platform in 2026.

In planning for online security, it's important to understand how transactions (including online traffic) can be created today. Each source has different risk characteristics and requires different controls.

SOURCE / 01

Human bank customer

A real customer executing a transaction. This may be an intentional activity, or done under some level of psychological or physical duress (the scam case).

SOURCE / 02

Bank customer agentic AI agent

A real customer delegates a series of transactions to an AI agent (e.g. find a flight to Heathrow, lock it in, pay with a credit card). Unfortunately, security for agentic AI is being built out as these agents are deployed. As a result, some could be weaponized as 'rogue' customer AI agents.

SOURCE / 03

Human non-bank customer

A human who is not the customer or real applicant, executing the transaction for fraudulent purposes. This could include people 'recruited' to open accounts at a bank branch.

SOURCE / 04

Computer bot

An automated bot sending traffic at low volume (to avoid detection) or higher/extreme volume. An example of this is credential stuffing where the fraudster has millions of credentials he wants to validate.

SOURCE / 05 · NEW IN 2026

AI computer agentic bot

Used in some of the most advanced attacks in 2026. An example of this involves phishing where the AI agent completes multiple tasks normally done by a human to set up an attack. The fraudster directs the bot to execute a complete phishing campaign, including:

1. **collecting information** on hundreds of phishing victims for a 1,000-person spear phishing campaign (targeted attack against one person with specific personal information)
2. **registering a phishing domain** name (e.g. `bankofh0mewood.com`)
3. **crafting the 1,000 personalized phishing emails**
4. **sending the emails** for the campaign

The Current Threat Landscape

*This section contains a summary of attack vectors used against banks and other entities. There is enough information included in this section to allow the reader to understand the next section, "**How to Mitigate Fraud During the Lifecycle of a Bank Customer.**"*

In a February 2026 article for the American Banker, David Maimon (Professor at Georgia State University teaching cybercrime research) provided an interesting summary on the fraudster we face. His thoughts were important enough, they provide the lead in to this section. This is based on his deep research of the dark web and the Telegram platform.

KEY THOUGHTS FROM DAVID MAIMON

"Fraud is no longer the province of individual scammers, but the product of a burgeoning industry of its own. What once looked fragmented now resembles a supply chain (with sophisticated tools and organization): identity sourcing, grooming, monetization and laundering, often handled by different actors specializing in each stage. "

- Operational defenses risk being overwhelmed without significant changes.
- Bankers need to take a clear-eyed look at the new fraud landscape and adjust their systems to reflect the new level of danger.
- Mail theft and more stolen checks in circulation, as much as data breaches, led to more identity theft attempts in the following weeks.
- Fraud succeeds not because every actor was brilliant, but because the system was organized.
- Customer education/messaging must be treated with the same rigor as transaction monitoring. **Messaging is a control. And controls should be tested.**

Overarching many attacks in 2026 is the use of Generative AI (GenAI), the new powerful capability that came onstream in late 2022. In the past 12 months, we have seen the fraudsters make use of this powerful tool to create new attacks or more powerful existing attacks (e.g. phishing). All of the power you read about GenAI for the good (productivity, coding, video, graphics and more) is available for the fraudsters. An entire white paper could be written about the severe threat of this 'evil twin'. Suffice it to say, fraudster use of GenAI makes 2026 the most dangerous year for online security since maybe the 2005-6 period.

SECTION 1

Peripheral Attacks: The Perimeter has Gone Industrial.

THREAT DEFINITION

Phishing & Quishing

General, or targeted attacks against existing customers, usually via email or text message, with the primary purpose of collecting user logon credentials.

Fraudsters have become very sophisticated in creating effective phishing attacks. It used to be you could look for weak structure in the emails or text messages or spelling errors. These days, you will rarely see that. Instead, you can see more targeted emails that create a sense that the message is real and needs some action taken. Text in the message can relate to you personally. Plus, it is quite easy for fraudsters to 'spin' up cloned web sites to convince the customer they have gone to a real site to complete the activity requested in the email. Phishing can be done via email, text messages and even with QR codes (Quishing).

Phishers have developed phishing-as-a-service where the components to create a phishing attack can be ordered. This can include: pre-built email/SMS templates, email/SMS distribution, cloned websites, admin panels to manage the phishing campaigns, reverse proxy modules such as Evilginx, possible hosting infrastructure and APIs/bots to help execute the campaigns.

Many of these phishing services can also be fully automated by Gen AI 'agents'. These agents can collect targeted information on individuals at scale (think 5,000 spear phishing emails), craft personalized emails for each target, create cloned web sites for several entities, register the domains, and initiate the campaigns.

In effect, much of phishing-as-a-service can be automated by Gen AI agents. This allows less-technical individuals to become phishers and be able to do it at scale. Running ten phishing campaigns at once is now a relatively easy activity. So, expect more phishing and more effective phishing campaigns.

Watch for more examples of fraudsters using AI agents in different attack scenarios.

Phishers have also gotten more sophisticated with the use of reverse proxy phishing to capture one-time passcodes in addition to the user ID and password. This is where the customer is taken to a cloned web site that is connecting to the real site. As the customer logs into the clone site, the phisher is simultaneously logging into the real site. As the real site asks for a one-time passcode (OTP), the cloned site is asking for the same OTP code. When the customer receives the OTP code, they enter it into the clone site. Then the clone site re-enters the OTP into the real site and authentication is complete on the real site. The fraudster can then complete unauthorized transactions. Brian Krebs identified a new reverse proxy service in 2026 called 'StarKiller.'

See Figure 2 for how reverse proxy phishing works.

FIGURE 2: HOW REVERSE-PROXY PHISHING WORKS

STEP 01 · CUSTOMER

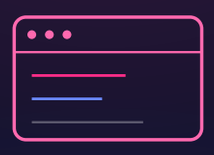
Phishing message

EMAIL · SMS · QR

The user follows a link from a malicious email, SMS or QR code that looks authentic.

STEP 02 · ATTACKER

Reverse-proxy clone site



Looks identical to the real bank. Credentials and MFA tokens are collected in real time during the live bank session the attacker is running in parallel.

STEP 03 · BANK

Legitimate site

Bank issues OTP challenge that the bots immediately relay back through the customer. The bank now believes the attacker is the customer.

WHY PHISHING STILL WORKS

Phishing continues to work well because most entities still rely on user ID, password and one-time passcode. For the most part, they are still not using phishing resistant authentication, which we will discuss more in the upcoming mitigation section.

THREAT DEFINITION

Cloned bank mobile apps

This is where fraudsters will take a legitimate mobile app, add malware to it and rewrap it as a legitimate mobile app and place it on third party mobile app stores. Once a customer downloads this app (e.g. via sideloading), the rogue app can copy screens, credentials and more.

THREAT DEFINITION

Fraudulent advertising

Fraudster buys ads from digital platforms to promote bogus bank products. The goal is to get customer/ prospects to click these ads and then get compromised.

THREAT DEFINITION

Credential stuffing

where fraudsters obtain stolen credentials and then test them against various legitimate web sites for matches.

The fraudster will conduct low- and high-volume attacks using these stolen customer user IDs and passwords from many locations (eCommerce sites, healthcare, gaming, phishing, etc.) to see if they match bank customer credentials (under the assumption the customer uses the same credentials multiple times).

Done at volume, this can seriously impact a customer's web site and the customers. Multiple attempts against the same customer can cause the account to be locked, causing customer concern on what is happening to their account.

THREAT DEFINITION

Bots & DDoS attacks

A high volume of automated/bot traffic focused at web sites, often for the purpose of causing the web site to shut down. High volume DDoS attacks often use compromised PCs, even TVs and Internet of Things (IoT) devices to creates the massive attacks.

In the fourth quarter 2025, Cloudflare saw "a record-breaking **31.4 Terabits per second (Tbps)** attack." Cloudflare said DDoS attacks doubled in 2025. DDoS attacks against bank web ***and*** mobile can shut down these services.

Even without DDoS attacks, volume automated 'junk' traffic can be a noticeable percentage of web traffic.

THREAT DEFINITION

Fake browser extensions

The use of browser extensions to conduct malicious activities, including extracting credentials, browsing information and emails. This information could be used for subsequent phishing or ATO attacks.

THREAT DEFINITION

Local SMS blaster

Custom made devices that emulate cell towers for the purpose of sending scam text messages (asking individuals to enter personal data) to a local audience. In one case, London England police arrested two individuals carrying a portable suitcase with an SMS blaster inside. This type of attack can also be done with a remote **SIM farm** (hundreds of SIM cards in a SIM box). *See Figure 3.*

THREAT DEFINITION

API attacks

This is where fraudsters use known entity APIs to attack institutions (both web and mobile) to obtain customer information or access to the institution. As an example, these can be APIs used by the mobile app to connect to bank servers at the main data center. Bank Administration Institute (BAI) says: "on average, financial services manage approximately 601 APIs." This a growing threat surface with the largest banks having thousands of APIs to protect.

To show how easy this can be to break APIs, there was a recent effort to show how easy it was breach a major consulting firm. A security company pointed its new autonomous offensive agent at this firm. "The agent mapped the attack surface and found the API documentation publicly exposed — over 200 endpoints, fully documented. Most required authentication. Twenty-two didn't." Significant amounts of data could have been exfiltrated.

NOTE: *The security firm disclosed the breaches to the consulting firm and all weaknesses were fixed before this information was publicly disclosed.*

FIGURE 3: SIM FARM
Source: PBS and United States Secret Service

THREAT DEFINITION

Remote access tool

Software that is placed on a customer's PC or mobile device, via subterfuge (e.g. support for a help desk scam, phishing with a link that when clicked will download this software). Remote access tool is used to remotely access the customer's device to execute financial transactions. Tool could be a generic product (e.g. Team Viewer) or one used solely by criminals. This is used for both retail and commercial fraud.

THREAT DEFINITION

Malware

Malware is any software, firmware, or code that is intentionally designed to disrupt, damage, or gain unauthorized access to computer systems, networks, or data, usually compromising their confidentiality, integrity, or availability. Usually, it can be downloaded on customers PC or mobile device, based on deceptive instructions, to be used by fraudsters to execute online transactions. This was a very active attack vector between 2010-2022 (popular names were Zeus, Dridex, Nyman, etc.) and can still be used in 2026, especially on mobile devices via side loaded apps and even from legitimate app stores.

THREAT DEFINITION

Supply chain attacks

This is where a fraudster uses a customer's supplier to access the customer site. The fraudster breaks into the supplier and uses this to tunnel into the customer's data center. So, how does this apply to bank security? There is a recent case where a US federal credit union sued its platform supplier, a Banking Services Processor. The lawsuit claims that hackers breached one of the supplier's systems used by the credit union and "took control of some credit union customers' accounts to steal money." As banks and others use many online security providers, including from some using new Gen AI capabilities, the potential for a breach affecting the security controls is no longer an unrealistic concern.

THREAT DEFINITION

Rogue AI agent attack

As many companies start to use Gen AI agents to do fraud work or interact with customers (chatbots), there is the potential that these agents can be hijacked by illicit prompts. These illicit prompts can then direct the agent to exfiltrate information, possibly even customer information if it is a chatbot. Since the chatbot is customer facing, a fraudster can easily enter an illicit prompt causing data exfiltration. An internal AI agent in a process flow (e.g. doing fraud analyst work) could be redirected by an insider attempting to access data.

THREAT DEFINITION

AI Agents Scanning for Vulnerabilities

Fraudster use of fully autonomous AI agents (hackbot) that can automatically scan web sites for vulnerabilities. One such bot is Shannon, an open source fully autonomous AI hackbot. As an example, this could be very effective in finding MFA design flaws that could allow for MFA Bypass. The Wall Street Journal reported on how **"Anthropic's Claude Opus 4.6 AI model found more high-severity bugs in Firefox in two weeks than typically reported in two months."** The article went on to say: "Tools powered by AI are increasingly adept at spotting vulnerabilities and are beginning to rival the talents of seasoned security experts." And now with Claude Mythos, AI Agent 'scanning for vulnerabilities' will become a more severe threat. **This may soon become the number one threat to financial institutions.** And smaller institutions will need help quickly patching vulnerabilities. This could become a 'survival' threat to smaller institutions.

THREAT DEFINITION

Crime-as-a-service

It is now possible to hire criminals to commit crimes against customers who have large deposits. And this can be done deceptively.

In a recent case, two teenagers were allegedly forced to commit a crime in Arizona. They went to a house in Arizona USA and tried to rob the owner of $66 million in a cryptocurrency account. The robbery was stopped, but the two teenagers claimed they were extorted into doing this robbery.

How can an institution prevent this kind of crime?

THREAT DEFINITION

First Party Fraud

When the actual customer is committing the fraud, this is first party fraud. This is becoming a growing trend. Here are some examples:

- Customer does a faster payment in the UK (authorized push payment) and then ‘claims’ they were coerced to do the transaction.
- Customer orders a TV from Amazon and then claims they never received it,
- Customer orders a product from an eCommerce company, then two years later orders the most recent version of the same product. Customer then sends the two-year-old product back (in original wrapping) for refund as though it was the new product.

SECTION 2

Direct Attacks Against the Bank Customer Life Cycle

This section will discuss the threats across the life cycle of a bank customer. Many of these threats also occur in the customer life cycle of other businesses. The fraudsters do not care which business they attack, as long as it is not too difficult and there is a financial reward.

Overarching Threats Across the Entire Lifecycle

This next set of threats run across the entire customer lifecycle.

THREAT DEFINITION

Call center threats

Fraudster attacks done against either the call center automated IVR or against the live call agents. According to Pindrop Security, **"based on internal data, AI fraud (non-live) surged 1210% by December 2025.**

This is allowing cheaper, faster, harder to detect and startingly scalable attacks." Some of these threats could also occur against live staff in other parts of a company (e.g. a wealth manager or investment advisor authenticating a customer).

This is becoming common in financial, healthcare and retail sectors.

THREAT DEFINITION

Capture reconnaissance info from IVR

The fraudster will use some of the customer stolen PII to defeat the basic IVR authentication in order to obtain information on the account which can be used in subsequent phone calls to the live call agent.

THREAT DEFINITION

Defeat voice authentication

The fraudster is easily able to get a recording of the customer's voice and can then replicate it for 'my voice is my password' type of authentication and pass this authentication. With Gen AI, the quality of the voice replication is quite high.

THREAT DEFINITION

Defeat KBA and OTP

The fraudster has the ability to collect so much information on customers that they often know the answers to KBA questions better than the customer. Bank impersonation allows the fraudster to contact the customer and to deceive the customer into providing the OTP that could be required by the call center as authentication.

THREAT DEFINITION

Fraudster AI Agent to attack IVR and live call agent

With Gen AI the fraudster can now create AI agents that have the ability to interactively communicate with either the IVR or a real call agent. To understand the interactive capability of AI agents, simply look at the Apate solution that uses AI agents to interactively talk with scammers. The purpose would be to collect more customer information and possibly request a transaction be executed. The use of AI agents can be done at scale. In eCommerce, there have been examples of fraudsters using AI agents to interact with staff dealing with returns.

THREAT DEFINITION

Money Mule Accounts

The use of a bank account or a cryptocurrency account to receive fraudulent or scam transaction money. The money is received in the money mule account and quickly moved (often within minutes) to many money mule accounts to make it difficult to track and recover funds. These accounts are also used to facilitate check fraud. Money mule accounts can be set up by scammers online, in branch or sometimes via the call center. Often scammers, for a fee of several hundred dollars, will enlist young students, existing customers and the homeless/drug addicts to open accounts. Romance scam victims are also coerced to opening money mule accounts.

THREAT DEFINITION

Fraudster/fraudster AI agent infects call center AI agent

Some entities will start to use AI agents to interact with customers. AI agents are subject to attacks, such as prompt injection. If the fraudster is able to take some level of control of the company AI agent, then they could possibly obtain customer information on one or more customers or even get the AI agent to initiate a transaction.

LIFECYCLE STEP 1 · ACCOUNT OPENING

In 2026, many controls for account opening have been broken.

Gen AI has proven to be quite effective as an attack method. The three primary areas that Gen AI has severely weakened controls are around **identity verification**, **document verification**, and at scale use of **stolen identities and synthetic IDs**.

Plus, Frank on Fraud reports that:

01 Bots collect large batches of stolen identity documents, including passports, driver's licenses and national ID cards, from dark web forums and data breach dumps.

02 Then, for each stolen ID, the AI searches social media and the open web for people with similar facial features.

03 The bots stitch the look-alike's photo together with the stolen ID document into a single composite image.

04 The bots automatically fire these composites at company verification systems. At scale, even a false acceptance rate of 0.1% becomes a reliable way in. Submit 10,000 fakes, and roughly 10 will pass.

THREAT DEFINITION

Stolen Personally Identifiable Information (PII)

Real customer PII that has been obtained from data breaches. There continues to be massive amounts of stolen customer PII available. Fraudsters use this information to open up accounts based on real people's data. This PII data will pass many control tests that will validate it is legitimate. This data can also be used as input for social engineering and ATOs.

THREAT DEFINITION

Synthetic ID Data

Synthetic IDs deliberately blend real and fabricated attributes to create scalable, fictitious identities with plausible and consistent identity data. Fraudsters can spend months and years building up the financial support data (history and documentation) to make synthetic IDs appear real. Gen AI allows fraudsters to scale up the creation of synthetic IDs to allow for high-volume account opening/credit card applications.

THREAT DEFINITION

Bogus account opening documentation

The creation of bogus government documents (driver's license, passport) and any supporting documentation to validate identity. Gen AI allows the quick creation of bogus complex government documents and supporting documents. A fraudster can simply pull a homeless person off the street and create a valid looking driver's license with a photo of the homeless person to be used to open up an account at a branch or online.

A recent example of fraudsters using people off the street to commit bank crimes occurred in a check cashing fraud case as report by Sylvia Gallo (below) on a 2023 case in the US Southern District of New York:

- To actually cash the cheques, the ring recruited people they called "walkers." These were often elderly, disabled, or homeless individuals who were paid $200 to $300 each to walk into a bank and make a deposit
- They (the fraudsters) dressed them, took them to hair salons, gave them fake IDs matching the account names on the cheques, and coached them on how to behave at the (bank) counter.
- The operators understood that the single most important moment in the entire scheme is the thirty seconds a person spends standing at a bank teller's counter.
- They understood that tellers are trained to assess legitimacy based on appearance, behavior, and documentation.

This is why the branch is just as vulnerable as the online for new account opening. The only difference is the fraudsters can do the online attacks at scale.

THREAT DEFINITION

Emulation of PC, Android and iOS mobile phones and cameras

The ability to completely emulate PCs, any mobile phone and the associated cameras. The fraudsters can inject photos and videos, with live motion activity to simulate a human.

THREAT DEFINITION

Deepfakes

The creation of videos to simulate a real person, based on as little as a photo. This is one of the advanced features of Gen AI used by fraudsters. They can have stolen PII, maybe a stolen driver's license and simply use the photo from the legitimate driver's license to create a video interactive session with the bank to 'prove' liveness.

THREAT DEFINITION

Defeating liveness testing

The ability to bypass the validation of a human holding an identity document or to defeat the actual liveness test of a human. This is done by emulating the device camera and injecting deepfake video into the presentation flow. In the past 12 months, there has been significant improvement in the quality of imagery injection.

THREAT DEFINITION

Defeating location

The use of the IP address from a nearby compromised PC to demonstrate proximity to bank's footprint or fooling the location signals on a mobile phone.

LIFECYCLE STEP 2 · LOGON

Defeating the front door.

THREAT DEFINITION

Defeat User ID and Password

This is where the fraudster has purchased and verified a stolen logon credential (user ID and password) and then uses these credentials to successfully login.

THREAT DEFINITION

Defeat One-time Passcode (OTP)

The ability of the fraudster to be able to use the one-time passcode as part of a fraudulent logon. This can be accomplished several ways, including:

1. Making a direct call to the customer impersonating a bank official and telling the customer as part of this phone call authentication, the customer will receive a text message and to provide the OTP code from the text message to the caller (while at the same time the caller logged into the customer's account online and the bank is sending the OTP to the customer as part of the real logon) and
2. Complete a SIM swap at the mobile store of the customer, claiming they lost their phone, need to buy a new phone with a new SIM. This allows the fraudster to log into the bank account with stolen credentials and receive the OTP directly.

Another way to defeat a one-time passcode is to compromise a soft token app on the mobile phone. In March 2026, Microsoft identified "a vulnerability in Microsoft Authenticator for both iOS and Android could leak your one-time sign-in codes or authentication deep links to a malicious app on the same device.

For the vulnerability to be exploited, the user would first need to install a malicious app on their device and then accidentally choose that app to handle a sign-in deep link."

A reminder of the sophistication of application scams (account opening, loan applications, eCommerce applications) from **Frank McKenna**, Point Predictive — via LinkedIn 03/04/26

> "This was one of the most sophisticated bust out rings that operated last year called the "South Beach Bust Out Syndicate" A broker would recruit straw borrowers who would go into the (car) dealership where the Finance manager would have them sign pre-filled applications. Fake employment and income on all of them."

This involved a former finance manager (insider fraud) and caused a loss of $1.5 million in fraudulent loans.

If you don't think this can happen with online account opening at a bank, think again.

KEY TAKEAWAY

1. At account opening, documents, PII and photo/videos can no longer be considered reliable truth.
2. Watch for account opening fraud at scale.

THREAT DEFINITION

Bypass multi-factor authentication

The ability of the fraudster to bypass one or more controls around the logon authentication, such as the OTP, a biometric, a token or push notification. The bypass can also take place during an active online session and can be done by getting around insecure application code. This typically occurs with weak security around the logon, maybe at the time of new controls being added. This allows authentication to be complete with just one factor of authentication.

KEY TAKEAWAY

This type of bypass can also occur as part of any step-up authentication during the online session

LIFECYCLE STEP 3 · CONTACT CHANGES

Takeover mode: Initiated.

THREAT DEFINITION

New account setup changes

Once the new account is set up, the fraudster change contact information to the fraudster contact information.

THREAT DEFINITION

Existing account changes

The fraudster will change contact information that allows him to pass subsequent authentication. One key change will be the phone number in order to receive the authentication OTP.

Oftentimes these changes can be made by the fraudster because the company has less security around contact information changes.

UK CIFAS · 2025

Unauthorised SIM swaps up 38%

SIM swap is one of the primary mechanisms used to take over bank and crypto accounts in 2025–26.

LIFECYCLE STEP 4 · TRANSACTION PROCESSING

The first move after the takeover.

THREAT DEFINITION

Malware

(See previous Malware definition)

THREAT DEFINITION

Bypass Step Up Authentication

(See previous Bypass Step Up Authentication definition)

THREAT DEFINITION

Account takeover transactions

These are the unauthorized transactions done by the fraudster after they have compromised the customer's online credentials (user ID, password, OTP or other authenticators). This also occurs when the fraudster has been able to place malware or remote access software on the customer's device.

One of the ways accounts can be taken is for fraudster to do a SIM swap to the customer's mobile phone in order to receive the authentication OTP code. UKs CIFAS has reported "a notable rise in unauthorized SIM swaps (+38%)." This is happening for bank and crypto account takeovers.

THREAT DEFINITION

Email, video and voice compromise

The fraudster using email, video or voice to fraudulently direct a corporate or consumer customer to send money to an account controlled by the fraudster. Listed below are the scenarios:

1. A member of the finance department receives an email purportedly from the CEO directing the finance staff to send a wire to account to support a confidential acquisition. They may even be told that the 'attorney involved' will call with more information.
2. A member of the finance department will receive an email purportedly from a current vendor with information that there is a new bank account for future invoice payments.
3. A member of the finance department will get a request to join a Zoom call which appears to have the CEO and several other employees on the call. The "CEO" will direct the employee to wire $10 million to a specific account. (real case)
4. A consumer buying a home will receive an email purportedly from the escrow company (US) or attorney involved in the purchase advising the buyer of a new bank account number to send the home down payment to.

These attacks go by the name of business email/video compromise, vendor email compromise and real estate email compromise.

THREAT DEFINITION

Rogue AI agent transactions

This is where a legitimate customer AI agent gets compromised by one of many AI Agent attack vectors. The most obvious way is via prompt injection that directs the customer's agent to execute in effect 'unauthorized' transactions. This could involve buying goods that the customer had no idea about, or even doing an online banking transaction. What is very unclear in these types of rogue banking transactions is whether the bank regulators will consider these transactions as truly 'authorized' (limited chance of reimbursement) or as effectively 'unauthorized' transactions (more often reimbursed). The rogue agent could also be 'forced' to give up the banking credentials (or other credentials such as eCommerce credentials), allowing fraudsters to do illegitimate transactions on their own.

THREAT DEFINITION

Consumer scams

The fraudster deceiving the customer to execute and authorized a transaction (e.g. wire, faster payment, ACH, etc.). There are so many scams the customer can be involved in, including romance scams, investment scams, impersonation scams (police, bank officials, etc.), grandparent scams and more. The long con scams (romance scams, investment scams) can involve the customer sending dozens of transactions over months, sometimes totaling over $1 million. The short cons (impersonation scams, grandparent scams) may often involve only one payment, often under psychological duress, as a 'one and done'.

These scams often start with a text message or WhatsApp type message or on a digital platform (Facebook Marketplace or dating app). What is most troubling about the messaging starting point is that fraudsters are moving from unencrypted messages to encrypted messaging, making it impossible for the telecom providers to detect the message. Ian Matthews, Founder and President of WMC Global says it best (from a Linked in comment in March 2026.

Ian Matthews, Founder and President of WMC Global says it best — via LinkedIn comment March 2026

> "Fraudsters are moving away from SMS messages to RCS and iMessage, which are end-to-end encrypted. Carriers cannot inspect the content. Third-party security vendors have no network-level access. The entire filtering stack that was built around SMS does not apply."

This puts more responsibility on the mobile platform vendors, Apple and Google, to be able to detect scam messages.

THREAT DEFINITION

Consumer scams - *continued*

As the scam gets under way, fraudsters deceive the customer to send money from their bank to typically another bank account. Any account the money is sent to is called a money mule account. Sometimes, and more often so, fraudsters will convince the customer to withdraw cash, purchase gold bullion and package it for a representative of the scammer to pick up (more common than you think). Other times the customer is directed to send the withdrawn cash to the scammer via a crypto ATM. A newer way, as part of investment scams, is for the victim to send money from their bank to a crypto exchange.

There are more losses in consumer scams than account takeovers. Even worse than the financial loss is the massive emotional toll on the scam victims, sometimes to the point of suicides. The fraudsters are using Gen AI in consumer scams, including deepfake videos/voice for romance scams to 'prove' it is a real person on the other end of the call, with automated translation capability, and the use of AI bots for impersonation calls (which allows for at scale scam calling.)

KEY TAKEAWAY

As financial institutions and other entities have tightened up online security, the fraudsters have successfully gone to the weak link—the consumer.

A REMINDER

"My voice is my password" is broken

With GenAI, voice replication is high-quality and trivial to produce. Voice-only call-center authentication should be considered a legacy control.

LIFECYCLE STEP 5 · OTHER SERVICES APPLICATIONS

Different service, same threat.

Other services that require an application process, such as mortgage, automobile and credit card applications, are suffering the same threats as bank account opening. Using stolen personally identifiable information and bogus government documents to obtain loans is becoming commonplace. AI can basically create any required loan documents that are fooling underwriters and cannot be detected by the human eye.

In fact, in February 2026 Australia's Commonwealth Bank just reported **over $1 billion in fraudulent mortgage loans "with documents created by AI."** The article also stated "the exposed loans could amount to $1 billion, which would be the biggest fraud ever committed against an Australian bank."

Frank McKenna, CEO at Point Predictive said: "Auto Lenders and dealerships are experiencing a rash of new AI schemes to assist fraudsters with creating flawless paystubs and bank statements. We also experienced a massive spike in dealership cloning where scammers use AI to create fake websites and impersonate dealerships to get wire transfers from customers"

Two More Sources of Attack Vectors

To go deeper into online security attack vectors, there are two additional sources. These sources list almost every known fraud and scam attack vector in an organized manner.

01 **Fraud Kill Chain:** The authors, experienced fraud fighters who also obtained additional input from dozens of banks, describe it as: "The Fraud Kill Chain is specifically developed for fraud and scams. It breaks them down into distinct stages—from reconnaissance and attack planning, to psychological manipulation, monetization and money laundering."

02 **MITRE Fight Fraud Framework:** This framework was just released on April 2026. MITRE describes it: "The MITRE Fight Fraud Framework™ (F3) is a curated knowledge base of tactics and techniques used by financial fraud actors, derived from real-world observations of cyber fraud incidents."

How to Mitigate Fraud During the Lifecycle of a Bank Customer

This is the core part of this paper. In discussing mitigation, we will look at identity, authentication, transaction anomaly detection, bot detection, location assessment, customer education and customer interdiction.

In defining a strong security mitigation program for financial institutions, there needs to be a clear and defined strategy. Here is a recommended high-level approach to consider:

01 Reduce all unnecessary traffic to the website and APIs.

02 Target 'left of boom' for serious controls to prevent fraud and scams before the financial transactions ever occur. 'Boom' is when the fraudulent transaction executes. 'Left of boom' is what occurs before 'Boom'. Effort spent 'left of boom' is significant because it helps to prevent fraud and scam losses and all of the associated operational expense associated with alert handling, funds recovery, SAR reporting, etc.

03 Have mitigation across the entire life cycle.

04 Have a strong program for quickly processing new customer account openings, while aggressively blocking fraudulent account openings.

05 Have multi-channel transaction anomaly detection.

06 Have consumer scam prevention program.

07 Have money mule mitigation program.

08 Have meaningful and actionable customer education. This should not be a check the box exercise, but education that causes the consumer to really think about their actions to prevent fraud and scams.

09 Have a meaningful program for customer interdiction when alerts are generated. This is especially important for business and consumer scam transactions—a major and growing issue for financial institutions in 2026. For consumer and business scam interdiction, the measure of success is 'did the money stay in the bank?'

10 Have a strong red team approach to constantly test for security gaps. 'Red team' is a combination of people/tools that are used (either internal team or outside group) to constantly validate security solution capabilities and identify security weaknesses in all channels (web/mobile/branch/call center).

*See **Figure 4** for a graphic portrayal of this security mitigation strategy.*

~80%

Estimated share of website traffic that may be unnecessary or fraudulent — bots, scrapers, attack agents.

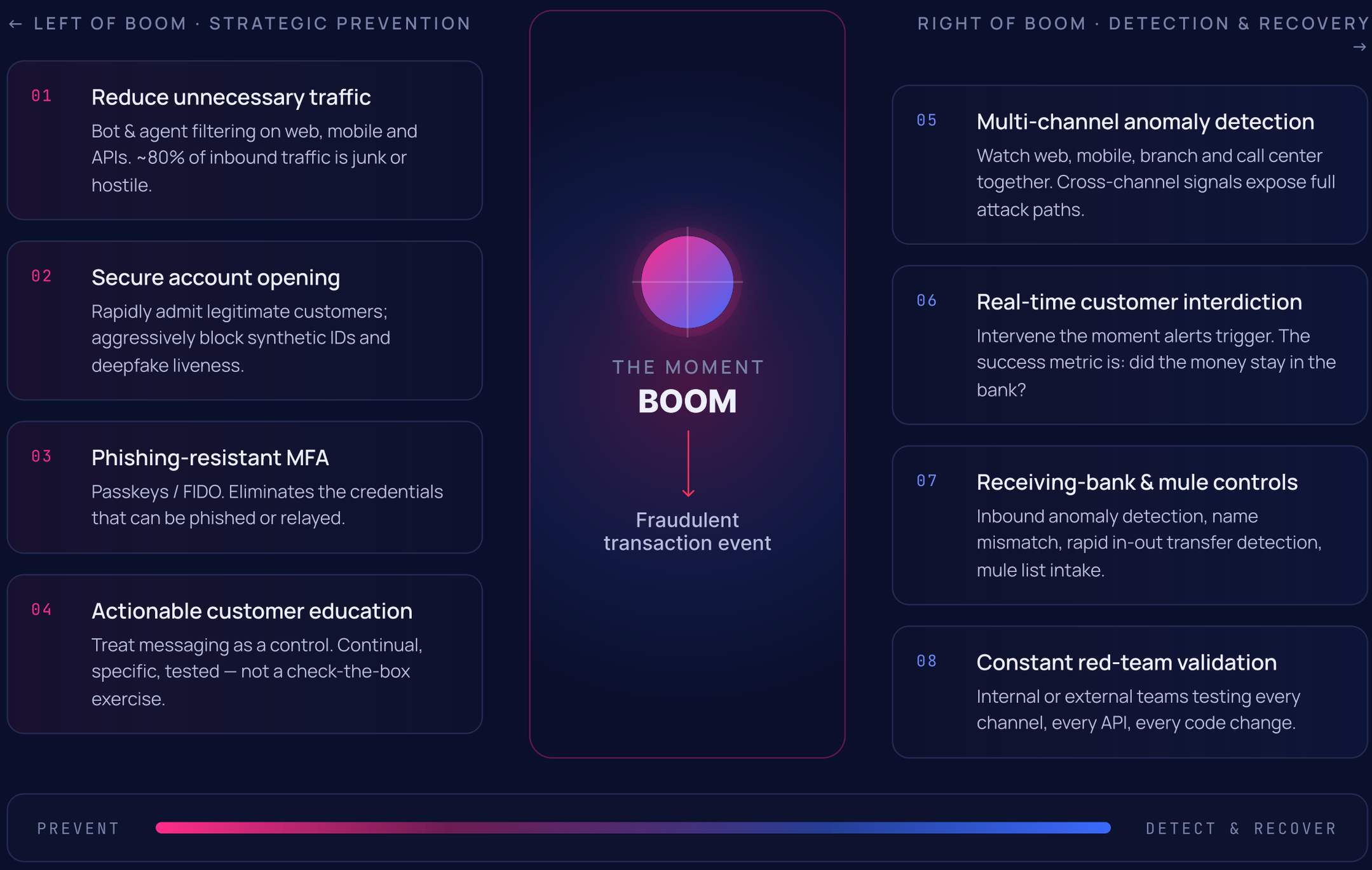
$25K

Suggested threshold above which retail wires should trigger a delay for further review.

60s

Practical undo window for high-risk consumer transactions where supported by the platform.

FIGURE 4: FINANCIAL INSTITUTION ONLINE SECURITY STRATEGY



For non-financial institutions, there are additional components to the strategy, such as:

01 Strong program for customer return management. (eCommerce)

02 Strong location confirmation for state licensed gambling.

The Controls for Online Security

The controls section is broken into three areas

SECTION 1 **Basic overarching controls**, including special notes that cover multiple solutions in more than one section.

SECTION 2 **Lifecycle controls**

SECTION 3 **Receiving Bank controls**

SECTION 1

Basic overarching controls

Notes on Identity and Authentication Assurances

To help understand the levels of assurance for identity and authentication required for online security, the best resource is the National Institute of Standards and Technology (NIST). NIST has a series of publications that go into depth on authentication:

- *NIST SP 800-63-4 Digital Identity Guidelines*
- *NIST SP 800-63A-4 Digital Identity Guidelines Identity Proofing and Enrollment*
- *NIST SP 800-63B-4 Digital Identity Guidelines Authentication and Authenticator Management*
- *NIST SP 800-63C-4 Digital Identity Guidelines Federation and Assertions*

NIST has precise identity and authentication assurance levels that help define the amount of security required for both identity and authentication.

See Appendix for more information on NIST Guidelines.

Blocking Bot Activity

An overarching control must be in place to block fraudulent bot transaction activity. There are many estimates that web site traffic can contain up to 80% of unnecessary or fraudulent traffic. So, the first order of business is to have controls that block illegal bot traffic. What is more difficult in 2026 is that there is now legitimate bot activity, known as customer AI agents. Plus, the fraudulent bot traffic can be made to look human. The solution must be able to properly distinguish between good and bad bot traffic. GenAI by the fraudster is making this assessment more difficult. Bot detection has to cover the following situations:

- Bots/AI agents used to conduct credential stuffing.
- Bot s/AI agents used to flood traffic to web site and mobile applications, including APIs used as part of web and mobile.
- DDoS attacks
- Bots/AI agents used for online account opening at scale.
- Bots/AI agents used to attack call center IVR and customer agents

Blocking phony bot traffic can reduce the cost of servers required to support web and mobile activities.

NOTE ON GEN AI SOLUTIONS

Several solutions in 2026 can incorporate the use of Gen AI components. It is important to remember that security is not built into Gen AI products by default. As a result, before any Gen AI components, such as AI agents and chatbots, are added to the security stack, a serious assessment and testing of the security of these components must be completed. Plus, full Customer Identity and Access Management (CIAM) has to be considered and deployed for AI Agents. Mill Pond Research warns that too often companies are putting AI agents into production with too much authority (access) and too little constraint on what they can do. The article went on to say: “we are entering the era of personal accountability for algorithmic actions.” Look for the concept of ‘guardian agents’ to become the norm to help mitigate AI agent security risks.

> **“We are entering the era of personal accountability for algorithmic actions.”**
>
> **Mill Pond Research**
>
> *On the risks of putting AI agents into production with too much authority and too little constraint.*

Block Phishing

A primary goal is to stop customer phishing attacks. This is paramount for two reasons:

1. Phishing attacks are too easy to create and they are scalable.
2. Most companies have not yet deployed phishing-resistant authentication.

This solution involves a multi-pronged attack. Below are the key solution components to prevent phishing:

1. Constantly search for clone web sites that can be used for phishing (and other attacks). In fact, Amazon recently launched SENTRIX . This robust program uses AI to analyze more than 50,000 suspicious URLs every week. This is a continuous sweeping of the internet to find these suspicious URLs. Continuous sweeping is essential because of the scale that fraudsters deploy to generate phishing sites. SENTRIX has three components: 1) advanced monitoring to scan for URLs (including using a third-party search API to monitor search results from several search engines), 2) risk assessment (including both textual and visual graphics/ logos), and 3) streamlined URL takedown.
2. Protect against reverse proxy phishing.
3. Have quick phishing site takedown capability, with ability to alert browsers (e.g. Chrome, Edge) to bogus sites before takedown is effective.
4. Prevent credential stuffing on web and mobile channels.
5. Be able to identify and report if a customer has entered credentials into a phishing site.

Mitigating Fraudulent Company Ads

It is important to detect and remove fraudulent ads promoting the company's services, especially financial services. There are solutions that effectively address this specific issue.

Call Center Controls

There are a number of call center controls that can be deployed. Some of these include:

1. Use a call print to detect where a call is coming from.
2. Create data base of marked fraudulent to calls to be able to identify fraudster in subsequent calls.
3. Detect call is an AI bot call.
4. Detect call is from person, but impersonating customer. This can detect impersonation of "my voice is my password" authentication.
5. Sound protection of IVR unit to prevent fraudsters from intelligence gathering.
6. Risk score inbound phone number.
7. Use mobile app to support authentication of caller.
8. Use models and IVR/Contact center touchpoints to help 'inform' on online banking transaction controls

Red Team Testing

As discussed above, red team testing is a critical component of a fraud /scam mitigation strategy. The best solutions are a combination of people and tools. Oftentimes the tools today include using Gen AI fraud capabilities like a fraudster to see if the current control stack is really working. This can also be used to red team test a new vendor before selection. A good red team will create fake: voices, documents, facial images, voices, liveness proofs. They will also create AI agent to interact with call center IVR and staff to test what information can be extracted. The red team basically tests the entire control stack.

Customer Agentic AI Controls

The concept of a customer using an AI agent (based on GenAI Large Language Models) to conduct financial transactions is coming into play in 2026. The AI agent may:

- Complete online account opening
- Apply for a credit card
- Complete an online financial transaction
- Buy a TV from Amazon
- Buy tickets for a concert
- Make airline reservations and pay for the ticket

A big open question is how to authenticate and validate this AI agent belongs to the customer and is doing what the customer expected. And given that AI agents can be compromised (e.g. prompt injection), how to know if the agent is compromised. Solving this problem is still a work in process. Several companies are initiating solutions to help control and protect customer AI agents. Here are some ways:

1. **Mastercard announced Verifiable Intent** described as "the layer creates a tamper-resistant record of what a user authorized when an AI agent acts on their behalf, establishing a shared source of truth across the ecosystem." The announcement went on to say "It is built on widely adopted specifications from the Fido Alliance, EMVCo, the Internet Engineering Task Force, and the World Wide Web Consortium. It is designed to work across agentic protocols, devices, wallets, platforms and even other payments networks." It will be interesting to see how this could apply to online banking transactions. In March 2026, Banco do Brasil and Visa conducted Brazil's first AI agent-initiated payment using the Visa Intelligent Commerce platform and tokenization. So, AI agent transactions are underway.
2. Use of customer identity and access management (CIAM). Banks should allow customer's agents to have different permissions (e.g. can do a wire, but only under $500). This is similar to how banks set up corporate online users (e.g. access to wires and ACH, but only up to $100,000). The authentication of a customer AI agent will need to be different than for a human customer. And banks need to define how will the AI agent be linked to/authenticated with the customer.

What is unclear is how to detect a bad AI agent from a good one. What is also unclear is if the customer AI Agent goes rogue (either from a prompt injection or just hallucinating) and executes a transaction not authorized by the customer, who is liable. Authorization and authentication will be key, along with the validation of device, network and behavioral elements. Also, will these agent transactions be considered unauthorized or authorized transaction, potentially impacting bank liability for reimbursement.

In the eCommerce payment space, there are a number of protocols being created to help with eCommerce payments. These include OpenAI and Stripes Agentic Commerce Protocol (ACP). Google has developed the Agents Payment Protocol (AP2) and the Universal Commerce Protocol (UCP). VISA introduced the Trusted Agent Protocol (TAP). I expect we might see some similar protocols for banking payment agents.

Securing data used by an AI agent/chatbot is also critical. In March 2026, Sears had to announce that one of its chatbots had left customer data (chatlogs and audio files) it collected unprotected. Prompt injections could cause the same issue.

This is why strong guardrails, strong privacy controls and detailed inventories, permissioning, auditability, authorization and authentication is essential for AI agents.

Consortium Data

Another overarching control is consortium data offered by the best solution vendors. This consortium data can be used across the customer life cycle to identify fraudulent data. It can run from identifying common device fingerprinting data or IP and network analysis at logon and transaction session to document verification items or other PII and synthetic data used multiple times at account opening against many banks. In fact, Point Predictive recently completed an analysis (Justin Davis, Point Predictive on LinkedIn post February 25, 2026) and showed:

- Fraudster A filed 47 apps using 6 SSNs.
- Fraudster C hit 35 apps across 10 SSNs.

These same applications were submitted to multiple entities. Without consortium data, these duplicate account applications would not be easily detected.

The Global Signal Exchange (GSE) is a new source for data sharing. Google and the Global Anti-Scam Alliance (GASA) came together to form the GSE in 2025.

According to the GSE website, the GSE platform provides tools for trusted parties to share threat intelligence it a way that fits with their organizational constraints, challenges and objectives. The GSE is currently working with large banks in the US and the UK on data sharing pilots.

Zero Trust Controls

What is unclear is how to detect a bad AI agent from a good one. Overarching controls for online security should follow the concept of Zero Trust. This has been an evolving concept for cybersecurity for companies. Here is the NIST definition (from the NIST Special Publication 800-207 Zero Trust Architecture) .

> **ZERO TRUST (ZT)**
>
> provides a collection of concepts and ideas designed to minimize uncertainty in enforcing accurate, least privilege per-request access decisions in information systems and services in the face of a network viewed as compromised.

In December 2025, Paulo Fernandes Biao wrote a research paper on extending Zero Trust to High Assurance Banking Systems. Online security controls have to assume that many compromises have occurred in the life cycle of online banking. Some of the many compromises are:

- Billions of authentication credentials have been stolen.
- Customers have been phished via reverse proxies.
- Customers have been coerced to execute transactions.
- PCs and mobile phones have been compromised with malware.
- Fraudsters hijack sessions in progress.
- Fraudsters create bogus verification documents

In this environment, there really is limited trust. And this mindset needs to be ‘front and center’ when deciding what controls need to be deployed for effective online security.

Each online security vendor used must also be subjected to Zero Trust thinking, as we continue to see supplier compromises.

Special attention should also be focused on the many APIs used for online banking.

SECTION 2

Lifecycle controls

LIFECYCLE STEP 1 · ACCOUNT OPENING

Stop the mule pipeline.

Online account opening is a major attack vector for criminals. Why? Because having a plentiful set of bank accounts allows fraudsters to create an unlimited number of money mule accounts. And money mule accounts are the fuel for account takeover fraud and consumer scams. Most countries have no idea how many bank accounts are currently money mule accounts. Recently, Thailand closed down hundreds of thousands of money mule accounts.

Strong online security has to stop fraudsters, but while easily allowing legitimate customers to set up new bank accounts—either online or at the branch.

THREAT DEFINITION

Device ID Assessment

Online account opening has to be able to assess the device identification data. Often there are indicators within the device ID that can alert to a suspicious device. With AI agents and bots also entering online applications, the solution needs to specifically assess the devices for these transactions. Consortium data can be helpful in catching the same fraudster across multiple entities.

Device identification for the customer life cycle begins at online account opening, if that is how the account is opened. This device information must carry forward through the other life cycle steps.

THREAT DEFINITION

Location

The solution should be able to identify the proximate location from where the applicant/fraudster is at. The solution needs to pierce proxy servers. The location component should be as good as required by US state gaming commissions that require online gamblers for a specific state be absolutely, positively within that state. Detection should look for a number of applications coming from similar locations.
In most cases, the company may not want any foreign applications, or applications out of state (or out of province). In the US applications must not come from certain sanctioned countries.

THREAT DEFINITION

Data entry behavior

The solution should be able identify behavior of data entry, including copy paste, slow entry of familiar numbers such as a US social security number. It should also detect if data is entered via remote access. There are literally hundreds of signals behavioral biometrics can synthesize and be available for risk review and analysts. And it can really be used throughout the customer lifecycle.

THREAT DEFINITION

Validate PII is for Applicant

Protecting account opening requires new approaches to insure the PII data is truly for the person applying for the account. There are techniques that can analyze the PII data in conjunction with telephone numbers, email addresses, even the age of the applicant.

THREAT DEFINITION

Detect Synthetic ID data

The ability to detect non-real, but realistically created PII data. This is one of the most challenging problems to solve as fraudsters are creating synthetic IDs at scale combining real and synthetic data, building trade lines and more to make this PII data look legitimate.

THREAT DEFINITION

Duplicate PII and Synthetic ID data

Look for the same PII data, phone numbers and email addresses being reused. Consortium data can be quite valuable for this detection activity.

THREAT DEFINITION

Identity verification

The requirement for knowing your customer requires a level of validating the customer is who they say they are. NIST offers several suggestions for effective identity verification.

THREAT DEFINITION

Liveness testing

The ability to validate the person applying matches to the document submitted. There is a comparison between the photo on the document and a live video session with the person. The solution has to be able to detect video injection. One test that may still be working is to ask the person to put three fingers directly in front of their face (recommendation from scam expert Jim Browning). But with the new Deepfake products such as Decart Lucy 2 discussed by Perry Carpenter in March 2026, even this test will no longer work.

For both document verification and liveness testing, there is some new research recently presented at the eFraud Global Forum San Francisco that discusses how the physics of photography and the pixilation of a human vs an AI image/video can really help to distinguish real vs fake images. Another solution forces light variation from the mobile phone on the person's face and is used to distinguish real from deepfake. Consider assessing solutions with injection attack detection components using CEN TS 18099 biometric data injection attack detection standard.

Red team testing is essential for liveness testing and document verification controls to validate the security solution can really catch fake documents and fake liveness testing. Gen AI use by the fraudsters is constantly making liveness testing a difficult control.

THREAT DEFINITION

Document verification

A common way to validate identity is via the use of a driver's license, government ID card or passport. Today's solutions have to take into account that GEN AI can easily create these documents for fraudsters. So, the solution has to be more advanced than what was acceptable even last year. And this is true for online or in the branch. One vendor in the US will use a driver's license for identity, AND matches it against the US American Association of Motor Vehicle Administrators (AAMVA) for verification. Unfortunately, not all states allow this access.

Also, in the US there are 24 states with mobile driver's licenses that can be used for verification. One challenge with this solution is there are currently 17 wallets used by these states. It may still be possible on a one off for a fraudster to create bogus documents and go to the state department of motor vehicles and obtain a valid driver's license or mobile driver's license (mDL).

We could be getting close to where government documents are just not trust worthy, without very strong verification tools (e.g. physical document scanner). By the way, every bank branch should have a physical document scanner for verification.

THREAT DEFINITION

Email and phone number verification

This is verifying that the entered email addresses and phone numbers belong to the submitted PII data. The fraudsters will often use stolen PII data and their own email address and phone number. There are many ways to:

- Check if a phone number is legitimate, and assigned, and does it match to the PII data.
- Check if an email address is legitimate, or comes from a suspicious domain, or the address itself is suspicious (Jones1, jones 12, jones123@yopmail.com) and does it match to the PII data.

The telcos have a number of security tests available from the GSMA CAMARA APIs data set.

LIFECYCLE STEP 2 · LOGON

Phishing-resistant MFA. Full stop.

THREAT DEFINITION

Authentication

In 2026, customer authentication should consist of phishing resistant multi-factor authentication (MFA). Full stop. Yes, it can be a challenge and costly to move the customer based from User ID/Password/ OTP or push notification to phishing-resistant MFA. But 'prudent-person' security leaves companies no choice, especially for financial entities. Phishing-resistant MFA should be the commercially reasonable standard in 2026. NIST SP 800-63B-4 (July 2025), page 35 contains a detailed discussion of phishing resistant MFA. There is also good information at the FIDO Alliance on phishing resistant authentication.

One of the key benefits of phishing resistant MFA is that the MFA cannot be phished. This dramatically reduces fraud attacks and losses. Three other benefits include the ability to:

- Support password resets
- Support multi-device access, to include multiple PCs and mobile devices, with multiple browsers, and additional devices such as TVs, gaming consoles and more
- Add a new device

Other good alternatives are physical Yubico keys, which are included in the FIDO phishing resistant offerings. Another good option, short of deploying phishing resistant MFA, is the physical hard token, by RSA and other vendors. Although the hard token offers an 'air gap' to PCs and mobile devices, it is not 100% phishing resistant. Some companies have replaced the hard token with a soft token, a solution that resides on a mobile phone as a mobile app. Both hard and soft tokens work the same way—a somewhat random 6-digit number appears on a screen. A soft token is also not phishing resistant.

Any hard or soft token can still be attacked via social engineering, but nowhere near as easy as OTP (simply by changing contact information or via a SIM swap). Soft tokens, essentially mobile apps, can possibly be attacked via mobile malware.

Another key consideration is to offer multiple authentication options. High net worth customers, or those with large investments accounts need a higher level of authentication security as opposed to a customer with a $1000/£1000/€1000 checking/ current account.

THREAT DEFINITION

Device ID

The device fingerprint assessment is important at logon. Since fraudsters can emulate device fingerprints, the control has to be sophisticated enough to detect false device fingerprints. Device identification needs to also take into account device changes that do not affect the actual assessment of the device such as browser updates (e.g. a new browser update vs reverting to an older version, operating system changes). Here are some additional controls around device fingerprint:

- Pierce proxy servers to obtain actual IP address of person/bot logging in.
- Identifying characteristics of the device fingerprints of good AI agents.
- Understanding the many devices the customer uses and the associated browsers. It would not be uncommon for one customer to use 3-5 devices and as many browsers.

There has been significant improvement in device identification in the past year. So, if you have a device identification solution greater than three years old, it is worth revisiting some of these new capabilities.

THREAT DEFINITION

Device Location

There are several controls around location that are critical.

- **Limit online access by country.** In the US this required as part of OFAC regulations for sanctioned countries.
- **Know exact location** of use device or use geo-fencing. There are solutions than can ensure location is within a state/province/country.

The strongest mobile device location detection may require an SDK to be added to the bank's mobile app, which is the practice of effective vendors such as the sponsor of this white paper.

Ask yourself—Do I need device location as strong as US states for gambling controls?

THREAT DEFINITION

MFA Bypass

To protect against MFA Bypass (coding weaknesses) requires red team testing wherever credentials are required. This testing, either internal or external, needs to be done frequently, but especially as part of any code changes. This problem is most often caused by weak security around MFA deployment, not from the vendor product.

THREAT DEFINITION

Presence of Malware and Remote Access

At logon, and throughout the customer life cycle, have the capability to detect for malware and remote access on both PCs and mobile devices. Determine how much to limit the customer activity when malware and remote access are detected. For the mobile device detections, it will require one or more SDKs to be added to the bank mobile app.

Some solutions will scan for other suspicious mobile apps and if the bank app itself being used is the legitimate bank app or a modified (e.g. malware laden) version.

LIFECYCLE STEP 3 · CONTACT INFO CHANGES

Serious challenge control needed.

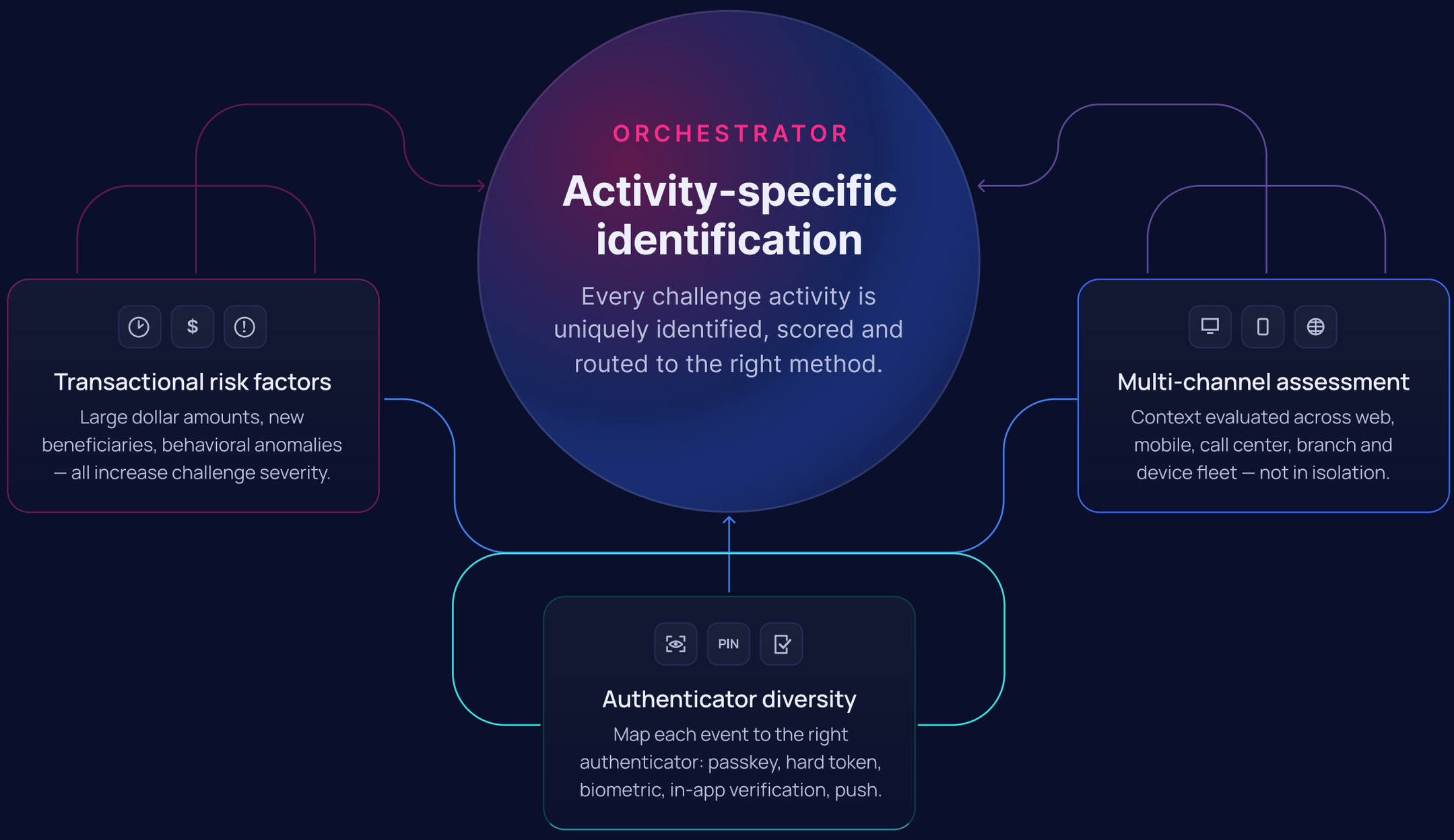
There needs to be serious challenge control for all important customer online activity —starting with contact information changes. One of the first steps of account takeover involves changing contact information.

To support challenge control (across the entire life cycle), one of the most effective components is what is called ‘challenge orchestration’. Challenge orchestration allows every challenge activity (e.g. logon, change phone number, initiate a wire) to be identified for unique treatment.

There can be a challenge method assigned to each challenge activity, along with consideration of the severity of the event (e.g. $100 vs $500,000 wire). The severity of the challenge method defines what level of authentication assurance is needed. Using the NIST guidelines for authentication levels of assurance is one way to help make this determination. The challenge orchestration needs to be in place for the customer life cycle.

See Figure 5 for a graphical description of challenge orchestration.

FIGURE 5: UNDERSTANDING CHALLENGE ORCHESTRATION



A special control around contact changes is to alert on contact changes made immediately after account opening. If the fraudster is able to bypass account opening controls, one of the first things they will do is change the contact information used for account opening (often part of the stolen PII) to the fraudster accessible contact information.

Plus, the customer should be notified of all contact information changes.

LIFECYCLE STEP 4 · TRANSACTION PROCESSING

Step up to safer transactions.

Using the challenge orchestration component, there should be appropriate step-up authentication for transactions. These can be financial transactions for banks and for purchases at eCommerce sites. Step-up authentication can be based on the type of transaction, dollar amount, time of day, etc.

Anomaly detection

A major control for transactions is anomaly detection. This can be based on both device information and transaction history by type of transaction. Every payment type should have anomaly detection in place. This includes bill payment, faster payments, wires, and ACH (or equivalent), etc. The analysis should be looking at the sending amount and the beneficiary account at the receiving bank. Better solution providers offer 'first-time beneficiary seen' feature. Plus, the solution will allow the customer to add models as part of the assessment. The anomaly detection must detect remote access.

An example of a good anomaly control when the fraudster commits a cookie or session hijack during a live customer session is as follows:

- Watch for anomalous activity *during* this session (e.g. customer makes payment, buys some goods, then fraudster changes contact information and does more transactions).
- Watch the behavioral biometric signals change during the session (e.g. there are now two different users during the same session).

Anomaly detection systems will generate alerts.

For commercial transactions, the alerts may just be for the fraud team (e.g. allow the wire transaction to complete online process, but hold the transaction). For suspicious commercial customer transactions, the anomaly detection system may have an integrated built-in delay, allowing the fraud analyst to investigate the transaction before it is released or cancelled.

For retail customers, these alerts, with contextual specific content, can be provided to customers online for real-time for decisioning, or if highly suspicious or high dollar (e.g. any retail wire over $25,000), the transaction can be delayed for further review. Alerts will also be generated for the fraud team.

The anomaly detection solution should contain link analysis capability. This allows the fraud analyst to find more fraud based on the data points from a newly confirmed fraud case.

An advanced use case for anomaly detection is to incorporate cross-channel activity into the anomaly detection engine. The anomaly detection solution can include logon, contact changes, call center activity and financial transactions. This can uncover a suspicious path the fraudster may have taken before they tried to execute the actual financial transaction (e.g. changed the contact phone number 3 days before, obtained balance information from the call center two days ago and now doing a large transaction).

Further, having specific capabilities around the aggregation of attacks via AI analysis can be a significant advantage in identifying and resolving these events, by understanding the control gaps and drivers that led to the incident.

We are starting to see additional evidence for industry Gen AI solutions for anomaly detection. In March 2026, Mastercard announced it has been "researching and building a new foundation model, which is a large-scale AI model that can be used as a basis for a wide range of applications…. Our new foundation model is a different kind of deep learning neural network (than standard Large Language Models- LLMs), called a large tabular model, or LTM, which is trained on structured data, such as large-scale tables or datasets." In this case Mastercard is choosing in effect to create a custom version of an LLM, but built on its own billions of records. We will also see Gen AI fraud solutions built using standard LLMs such as Anthropic's Claude or Open AI's ChatGPT. GenAI Copilots within analytics tools are currently present and should be leveraged to improve speed, visibility and reliability in investigation outcomes.

ISO 20022 Payment Data

The new ISO 20022 payment standard allows sending banks to add many new data fields to the payment transaction. This data will become valuable to the anomaly detection system.

BEST RECOMMENDATIONS FOR AI AGENT SOLUTIONS

1. Crawl before you walk.
2. Make sure the guardrails, along with security, are strong: Inventory the desired agents, assign them permissions, create and define authentication and authorization controls
3. Make sure you can audit where the AI agent went and what data they collected and what were the reasons for any decisions they made.
4. Make sure the AI Agent follows regulatory guidance and it is not perceived as biased.
5. Fully autonomous AI agents should be a final activity only after successfully completing semi-autonomous AI agent activity and careful consideration of the risks of fully autonomous AI agents.

Two Princeton University professors are developing a ‘reliability index’ for AI Agents. They want to quantify the reliability of an AI agent. In the Financial Times, they summarized four criteria for this assessment:

1. **CONSISTENCY**
   They get it right consistently, not right today and wrong tomorrow on the same thing
2. **ROBUSTNESS**
   They don’t fall apart when conditions aren’t perfect
3. **CALIBRATION**
   They tell you when they’re unsure rather than confidently guessing
4. **SAFETY**
   When they do mess up, their mistakes are more likely to be fixable than catastrophic

They see about 80% reliability in some studies they have done. Weaknesses are hallucinations and ambiguity. The ‘human in the loop’ is probably still key in using AI Agents for fraud detection and automation of case management summaries.

Other transaction controls:

- **Restrict mobile transactions** if suspicious activity is detected on the mobile phone (e.g. remote access, screen overlay or mobile malware detected.)
- Have strong mobile deposit controls.
- **Behavioral biometrics**: This is a powerful solution for sending banks to better understand the transaction and whether it is being done by the customer/AI agent or a fraudster/AI agent. Can also be used on the receiving bank side.
- Allow customers to set locations where high-risk transactions can be performed. Any other location can have a built-in time delay of x hours. This is very helpful when mobile phones are stolen.
- **Receive receiving bank account risk score:** Sending bank can request risk signal information from the receiving bank in some situations today.
- **Third party receiving bank information:** There are several solutions where the sending bank can obtain limited information on the receiving bank and receiving bank account from third parties.
- **Payment operator network risk signals:** Several faster payment system operators provide sending banks receiving bank risk signals (e.g. EWS/Zelle, FedNow, The Clearing House RTP, UK and Philippines Faster Payments). It is valuable to use network signals, when available, to stop bogus transactions. Watch for more network signals being made available.
- **ISO 20022 Pre-Validation:** ISO 20022 is the new transaction format for payments. Sending banks will be able to request information from the receiving bank before the transaction is executed. This is still in the early stages of deployment.
- **Confirmation of payee:** Sending bank can send receiving bank account name and account number to receiving bank for matching verification. This is becoming common in Australia, the EU, New Zealand and the UK. It can help prevent consumer scams and customer 'fat fingering' of input data.
- Sharing money mule account information among banks.
- **Customer can confirm bank is calling customer.** To help prevent bank impersonation, several banks have developed a way for the customer to confirm it is the bank official/fraud analyst actually calling the customer, by simply having the customer open the mobile app and see a screen that shows if the bank is/is not calling the customer at this moment. This can prevent fraud and scam losses (e.g. being asked for the OTP or being told to move money from an account for security reasons).
- GenAI chatbots for scam resolution, fed by transaction risk models to interrogate possible victims on the potential red flags and red path/yellow path/green path customers.

Commercial transaction controls for email, video and voice compromise fraud transactions

For Email, video and voice compromise fraud transactions, the best solution on the commercial side is anomaly detection and first new beneficiary to a new country or bank account. As most consumers do few large amount transactions, anomaly detection will not help per se. Consider placing holds on high dollar consumer transactions.

What is key is the bank must provide strong and continuous education to its commercial customers about this threat, including to real estate commercial customers involved in home purchase transactions.

The key point is never trust and email or inbound phone call. And **ALWAYS** verify with the requestor using a channel different from where the request came from. Example: if the request comes from an email, use a telephone number from a source other than that actual email. This threat is over 15 years old and the annual losses are greater than ever—in the $ billions per year in the US alone.

Consortium Data

Consortium data can be extremely helpful in preventing transaction fraud losses. This can provide some of the best alert signals with strong false:positive ratios. Consortium data may provide some of the best fraud signals.

Preventing Scam Transactions

In addition, there should be a series of controls for banks to specifically help prevent consumer scams. Consumer scam transactions are transactions authorized by the customer (authorized transaction)—but under deception. So, the customer believes they need to complete these transactions (based on impersonation of bank or government official, a romance scam or investment scam, etc.). Although these scams start outside the bank, the bank has a role to help the customer and see that the potential scam money stays in the bank. Below are some of the controls recommended for banks to deploy.

- Default for new faster payment services should be 'requires activation'.
- Any activation of faster payment services should have a built-in overnight delay.
- Build in friction and delays for suspicious transactions.
- Use behavioral biometrics to identify odd customer transaction activity (examples: doodling on screen, abnormal delays and session length during the transaction steps).
- Share data between sending and receiving bank to possibly alert on a suspicious receiving bank account.
- Analyze if the customer is on an active voice call while completing a financial transaction.
- Remote Access Tool or screen broadcasting
- Limit payments to high-risk channels, including crypto exchanges.
- Allow customers to lock the account if they become concerned about a fraud or scam.
- Allow customers to have their own additional security controls
  - Delay faster payments by X hours.
  - Have low faster payment amount limits
  - Limit the amount of money in the customer's account that can be transferred online.
- Have a trusted contact on senior accounts.
- Hold suspicious transactions for customers over 60 (as allowed by law).
- Allow transactions to be undone. This 'undone' capability can take place in the x seconds (e.g. 60-120 seconds) after the transaction was executed.
- Buy list of money mule accounts. There are several vendors that have 'honeypot' techniques to collect money mule bank accounts and cryptocurrency money mule wallet accounts. This is valid for fraud prevention as well.
- Use AI agent with LLM to help fraud staff effectively interactively interdict with potential scam victims. Use AI agent to listen to customer response and then help identify the 'next best question' to ask to help 'break the spell'.
- Have dedicated scam prevention team.
- Conduct staff training on the psychology of scams (understanding the mindset of a scam victim and how to effectively communicate with such a person under the influence of a scammer) to allow staff to be able to best stop the funds from leaving the bank once a transaction alert is generated (or the customer is in the branch wanting to withdraw a large amount of cash or do a large wire or transfer to a crypto exchange). Staff interdiction with a potential scam victim is an art that banks must master as part of scam prevention.
- Offer customer scam messaging detection controls based on specific scam models and third party (i.e. telco) integrations.
  - There are several solutions that allow customers to help determine if a message they receive (text, WhatsApp, etc.) is a potential scam.

Overarching the specific scam controls, there should be a written scam prevention strategy. Many banks have already put in place a well-defined scam prevention strategy. *See Figure 6 for a composite of these consumer scam prevention controls.*

FIGURE 5: STRATEGIES TO PREVENT CONSUMER SCAMS

Consumer scams produce authorised transactions under deception. Defense requires two parallel tracks: systemic friction & intelligence on the platform side, and human-centric protection & intervention on the customer side.

01 Systemic friction & data intelligence

Strategic delays & activation
Default new faster-payment services to "requires activation" with built-in overnight delay. Build friction for suspicious transactions.

Behavioral biometrics
Identify on-screen doodling, abnormal pauses, copy-paste of familiar numbers, active voice call during transaction.

Cross-bank data sharing
Exchange data with receiving banks to flag suspicious beneficiary accounts and mule networks before the money moves.

Network & consortium signals
Use FedNow / RTP / Faster Payments network signals.
Buy mule lists. Apply consortium data to find at-scale fraudsters.

02 Human-centric protection & intervention

Empower customer controls
Account lock, low per-channel limits, undo-window (60–120s) on high-risk payments, custom location rules.

Trained interdiction staff
Train on the psychology of scam victims. AI agent suggests the "next best question" to break the spell during the call.

High-risk safeguards
Hold suspicious transactions for customers over 60 (as law allows). Trusted contact on senior accounts. Crypto-channel limits.

Customer scam-message detection
Offer customers an app to identify whether a text or WhatsApp message is a likely scam — before they ever respond.

Data Sharing

Data sharing between financial institutions is another valuable control to prevent fraud losses. There is no consistency in how and if this is done around the world. Regulation, or the lack of directed regulation, limits data sharing. Banks need to work with other banks, trade associations and regulators to allow for full data sharing, subject to privacy. The UK has some excellent data sharing. Members of the National Cyber Forensic Training Alliance (NCFTA) in the US are involved in data sharing. In the EU AMLR Article 75 and the new Payment Services Directive 3 (PSD3) require data sharing. Data sharing should be part of your online security stack.

LIFECYCLE STEP 5 · OTHER SERVICES APPLICATIONS

Many of the same comments around controls for online account opening apply for other applications. Even though there are often more documents required for these applications, the banks need to realize they can also be AI generated.

Another more prevalent solution is a number of consortiums around different types of applications. There are consortiums for automobile loans and consortiums for personal loans. These application consortiums are very helpful to identify the same fraudsters using the same devices, loan documents and application data.

SECTION 3

Receiving bank controls

The newest area for formal controls revolves around receiving bank accounts (beneficiary accounts). Controls need to be placed on receiving bank accounts because these accounts are where fraud and scam monies go. And it is known fact there are way too many money mule accounts in existence across the banking ecosystem, it is just too easy for fraudsters to have access to money mule accounts.

As a reminder, there are different ways that money mule accounts are established:

1. New account opening.
2. Existing customer loans their account out for a fee (e.g. solicitation for accounts from a Telegram channel or other social media). In a 2022 UK survey, "one in 10 young people would agree to move money through their bank account in return for cash".
3. Foreign students who come to a country for the express purpose of opening an account for money muling or once they leave, they sell the account.
4. Romance scam victim (existing customer) who becomes complicit (based on the psychological grooming of the fraudster) in money mule activity.
5. Pay an economically challenged person a fee to open a bank account. This is also used for check cashing.

Here are the controls required to detect and mitigate money mule accounts.

- Strong online account opening controls. These controls are listed above and included here as a reminder that strong controls at account opening (left of boom) can help to prevent the establishment of money mule accounts. Use of shared consortium data across many banks, eCommerce, crypto exchanges, lenders will really increase the leverage to identify money mule networks.
- Inbound anomaly detection on wires, ACH and other inbound transactions. In the US, NACHA currently requires anomaly detection for inbound ACH transactions.

  Key inbound analysis includes:

  - Looking at inbound transactions in conjunction with previous activity to also detect existing customer who has become a money mule.
  - Looking at brand new accounts with high dollar inbound transactions or relatively dormant account receiving high dollar inbound transactions.
  - Link analysis once one money mule account is detected to use the data on that account to find 'similar' accounts. For this control to work, it is important to save all of the internet/device data associated with new account opening.
  - Use of payment operator network signals have been very effective in the UK to find money mules by tracking a series of sequential inbound/outbound/inbound activity among banks.
  - Robust data sharing among banks will provide material signals to detect and remove money mule accounts.
- Mismatch of name on inbound transaction vs name on beneficiary account.
- Inbound transaction with corporate name going to consumer account.
- Holding suspicious inbound transactions (subject to law and regulation).
- Behavioral biometrics to detect anomalous log in activity by the money mule (e.g. constant logging in to see if the money has arrived)
- Analysis of rapid outbound transfer of recent inbound transactions.
- Staff training to be better able to detect and prevent money mule activity.
- Buy list of money mule accounts. There are several vendors that have 'honeypot' techniques to collect money mule bank accounts. Use these lists to help find money mule accounts within the organization.
- Educate consumers about the inherent risks involved in being a money mule.

Work with your legal department to understand the addition of receiving bank controls and any possible UCC 4A concerns around liability. *See Figure 7 summarizes the controls for money mule management.*

FIGURE 7: COMBATING THE MONEY MULE

SOURCES

How money mule accounts are established

01 **New fraudulent account opening**
Synthetic IDs, stolen PII, AI-generated documents and deepfake liveness — at scale.

02 **Existing customer loans their account**
Solicited on Telegram & social media for a fee. "1 in 10 young people would do it for cash." — UK survey, 2022

03 **Foreign students & transient accounts**
Open an account purely for muling, or sell the account on leaving the country.

04 **Complicit romance-scam victim**
Psychologically groomed into facilitating mule activity from their own existing account.

05 **Paid "walkers"**
Homeless, elderly or vulnerable individuals paid $200–$300 to open an account or cash a check.

DETECT & REMOVE

Receiving-bank controls for accounts already open

01 **Inbound anomaly detection**
High-dollar transfers into new or dormant accounts. NACHA already requires inbound ACH anomaly detection in the U.S.

02 **Name mismatch**
Name on inbound transaction does not match the beneficiary account. Corporate-named transactions hitting consumer accounts.

03 **Rapid in-out detection**
Funds arrive and are pushed out within minutes — often to multiple downstream mule accounts.

04 **Behavioral biometrics for mules**
Constant logins to check whether funds have arrived; repeated session patterns inconsistent with normal customer behavior.

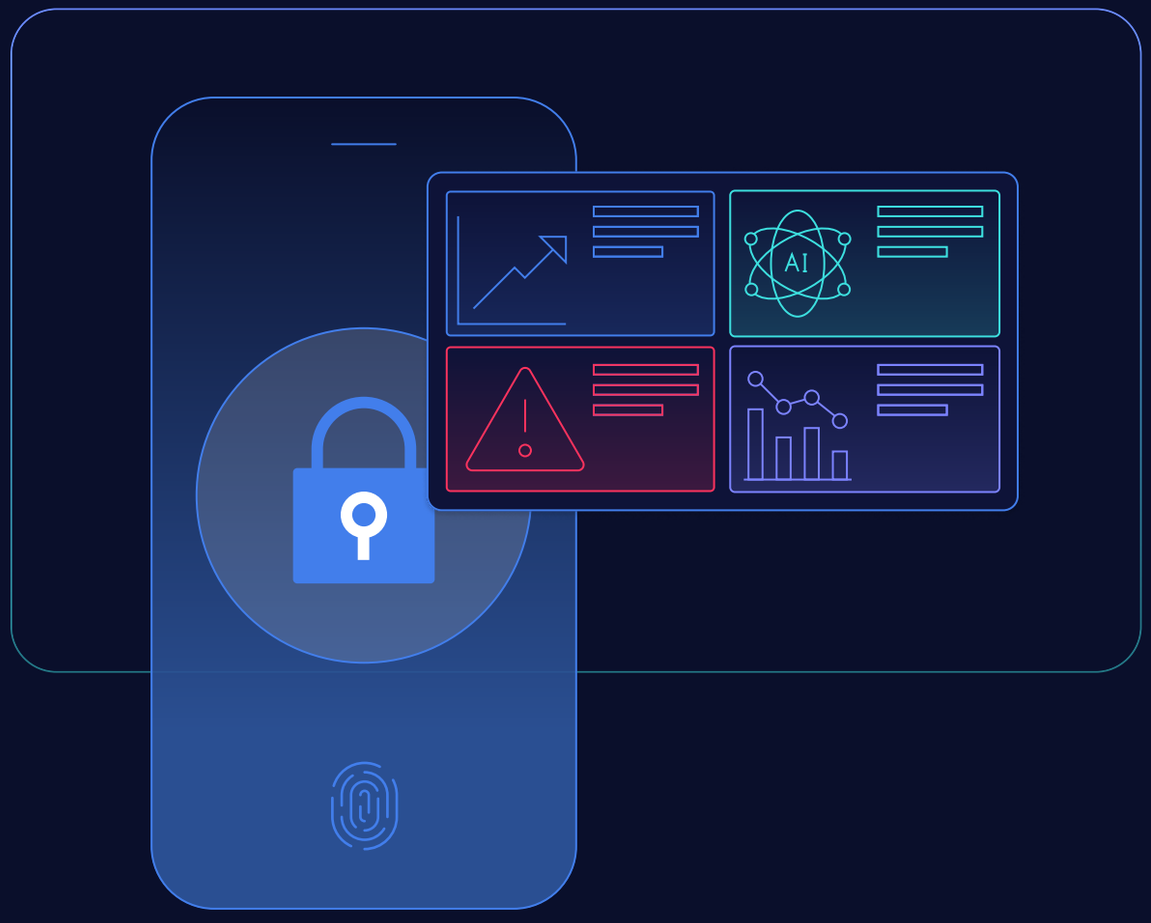
05 **Link analysis & mule lists**
One confirmed mule reveals similar accounts via device, PII and metadata links. Honeypot vendors sell active mule lists.

LEFT OF BOOM

Strong account-opening controls (consortium PII, device, biometrics, document scanning) prevent mule accounts from existing in the first place.

FCA FINDING

"Where firms are outliers — more reported mule accounts than peers — there is a lack of senior management oversight."
Treat mule mitigation as an executive-level program.



NOTE

the UK's Financial Conduct Authority (FCA) has found that disinterested senior management in addressing money mules often leads to poor money mule management programs. "A proactive strategy for tackling money mules helps protect a firm's regulatory compliance, reputation and customers. We (FCA) found that where firms are outliers, in that they have more reported mule accounts than their peers, there is a lack of senior management oversight."

SUMMARY

Prepare for a constant wave, not a blizzard.

Banks and other online entities have to prepare for the wave of AI fraud and scams. It is not a one and done blizzard. It will be constant. There was a recent article by fraud expert Gavin Holland which summarized the new constant with AI. He calls it Fraud 6.0. Here is his definition of Fraud 6.0:

F6 / 01
Agentic phishing & social engineering at scale

F6 / 02
Real-time deepfake voice and video

F6 / 03
Next-generation synthetic identity fraud

F6 / 04
Autonomous account-takeover (ATO) agents

F6 / 05
AI-orchestrated money mule networks

F6 / 06
Adversarial AI actively probing bank defenses

Now, all of these capabilities may not exist today, but there is a good probability they will exist tomorrow- and be extremely good. These are no longer 'black swan' events.

There needs to be more data sharing between banks, eCommerce, telcos, digital platforms and other to help in this fight. Plus, fraud fighters need to push payment system operators to provide more transaction data signals. Network signals are provided in the UK and the Philippines, with Zelle in the US and there are pilots in the US for FedNow and The Clearing House RTP transactions. When Canada rolls out real time payments, there will be network payment signals included. Payment system operators need to also include network signal on wires and other payment types.

Consortium data will continue to be most valuable to fight organized crime.

In the next few years, we should start to see more Gen AI solutions to fight fraud and scams. We need to have fewer alerts, with more precision in positive fraud detection. We will also see AI agents helping fraud analyst productivity and changing the paradigm of alerting and customer interdiction.

A sound online security strategy should include protection for 1) the consumer across the entire consumer life cycle, including consumer scam prevention – with effective customer interdiction, protection against any channel used in the life cycle and 2) a money mule management program for banking safety and soundness.

At the March 2026 eFraud Global Forum, Erika Sanchez at BanCoppel in Mexico had four very relevant quotes for online security:

01

Trust must be orchestrated across signals, channels and time — not assumed from one convincing interaction.

02

Fraud maturity is no longer defined by how much you detect. **It is defined by the quality of the decision in the last reversible moment.**

01

AI contributes speed, scale and hidden pattern discovery; the human (staff) is responsible for context, edge cases, ethics, material overrides and accountability.

02

The closing question is no longer “Did the transaction look legitimate?” **The real question is: Did we verify the truth before the money moved?**

As shown in this document, there are dozens of controls required for the customer life cycle. These controls must address both fraud and scams. It takes careful thinking to select and deploy these controls. And it takes a constant assessment of your security position, probably annual at minimum, to be able to identify new security gaps to be addressed. Red Team testing should be part of this on-going assessment exercise, including constant testing of all online related APIs.

With the heavy losses from consumer scams, banks need to be actively working with telecom providers and digital platforms to help eliminate the start of these scams. In the US, several telcos are adding scam prevention controls. Match Group is adding bank level customer authentication for new accounts. Apple and Google are adding scam prevention controls at the mobile phone platform level. A number of vendors are offering scam message detection apps that banks can make available to its customers.

When assessing vendors to help complete the online security stack, remember we are going through a capability paradigm shift in 2026. What was recently impossible to have as an online security control may very well be available in 2026-7. We are already seeing impossible controls (and attacks as well) that were not believable 12-18 months ago. So. Really be creative and thoughtful when assessing vendors.

CLOSING

Did we verify the truth before the money moved?

The threat of Gen AI to disrupt security is real. The Anthropic Mythos model is a ‘shot across the bow’ to warn that many black swan events we used to think of are now way more real and near-term threats. The Wall Street Journal best summarized the potential threats of the Mythos model with these two quotes:

> **“The U.K.’s AI Security Institute, after testing Mythos, found the model could exploit vulnerabilities on its own, quickly executing tasks that would take a human days.”**

> **“Whether or not Mythos is a hacker superweapon really is immaterial to the conversation. If it’s not this model, it’ll be another one in five minutes.”**
>
> **Dave Lewis**
> Global Advisory, Chief Information Security Officer, 1Password

The online security solutions need to be “all hands-on deck”. Network signals, consortium data, anomaly detection and ‘Gen AI fighting Gen AI’ will be the norm. We also need more partnerships between government, banks, telcos and digital platforms— where some countries have addressed this is with a National Anti-Fraud Center to co-ordinate all these players.

Don’t be the one caught in this blizzard all by yourself.

About Transmit Security

Transmit Security delivers future-proof customer identity experiences in a world where AI is accelerating change across both fraud and user access. We do this by fusing customer identity, fraud prevention, and identity verification into a single, unified system, eliminating silos and enabling rapid adaptation.

At the core is our Predictive AI, continuously learning from real-time signals to detect intent, uncover emerging fraud patterns, and make accurate decisions before damage is done. This fusion-first approach allows leading enterprises to stay ahead of evolving threats while delivering seamless, secure experiences to their customers.

Appendix

National Institute of Standards and Technology (NIST)
SP Digital Identity Guidelines-SP 800-63-4 | VERSION: 4 July 2025

From NIST website / https://pages.nist.gov/800-63-4/

These guidelines cover the identity proofing, authentication, and federation of users (e.g., employees, contractors, or private individuals) who interact with government information systems over networks. They define technical requirements in each of the areas of identity proofing, enrollment, authenticators, management processes, authentication protocols, federation, and related assertions. They also offer technical recommendations and other informative text as helpful suggestions.

There are four sections to SP 800-63:

SP 800-63-4: Digital Identity Guidelines

SP 800-63: describes the digital identity models, risk assessment methodology, and processes for selecting assurance levels for identity proofing, authentication, and federation.

SP 800-63A-4: Identity Proofing and Enrollment

SP800-63A: provides requirements for identity proofing and the remote or in-person enrollment of applicants, who wish to gain access to resources at each of the three IALs. It details the responsibilities of Credential Service Providers (CSPs)with respect to establishing and maintaining subscriber accounts and binding CSP-issued or subscriber provided authenticators to the subscriber account (enrolled in the service.

SP 800-63B-4: Authentication and Authenticator Management

SP800-63B: provides requirements for authentication processes that can be used at each of the three AALs, including choices of authenticators. It also provides recommendations on events that can occur during the lifetime of authenticators (e.g., invalidation in the event of loss or theft).

SP 800-63-C-4: Federations and Assertions

SP800-63C: provides requirements on the use of federated identity architectures and assertions to convey the results of authentication processes and relevant identity information to an organization's application.