StrongestLayer

The DocuSign Detection Gap

Why Signature-Based Email Security Fails Against Modern Impersonation Attacks

StrongestLayer Threat Research | December 2025

TABLE OF CONTENTS

EXECUTIVE SUMMARY

StrongestLayer threat researchers analyzed over 2,500 advanced email attacks from Q4 2025—every one of which had bypassed Microsoft E3/E5 or leading secure email gateways (the top 3) before detection. Among these sophisticated threats, DocuSign impersonation emerged as the dominant attack vector, accounting for 13.8% of all advanced attacks analyzed.
Key Findings:

- 281 DocuSign credential harvesting attacks detected across monitored environments in Q4 2025
- 0.458 average Jaccard similarity among DocuSign attacks—meaning each attack shares less than half its features with other variants, rendering signature-based detection ineffective
- 68% of all advanced attacks fell below the 0.30 Jaccard threshold where pattern-matching effectiveness drops below statistical significance
- 38% of DocuSign attacks showed high AI-assistance indicators, suggesting automated generation of unique variants
- Document-heavy industries (legal, pharmaceutical, financial services, real estate) face elevated risk due to operational dependence on electronic signatures

EXECUTIVE SUMMARY

The fundamental challenge: legitimate DocuSign emails exhibit every characteristic security teams train users to treat as suspicious—unfamiliar senders, urgent requests, external links, and requests for sensitive actions. Attackers exploit this business logic problem at scale, and AI-generated attack variability makes signature-based detection increasingly ineffective. Organizations face an architectural decision: continue investing in pattern-matching approaches that demonstrably fail against variable attacks, or adopt reasoning-based detection that evaluates intent and business context rather than static signatures.

THE DOCUSIGN PROBLEM: A PERFECT STORM FOR ATTACKERS

DocuSign has achieved something rare in enterprise software: near-universal adoption across industries. Contracts, NDAs, employment agreements, vendor paperwork, regulatory filings—electronic signatures have become embedded in daily business operations. That ubiquity creates opportunity for attackers.

Why DocuSign Impersonation Works

Legitimate DocuSign emails routinely exhibit characteristics that security awareness training teaches users to treat as red flags:

Phishing Indicator	Legitimate DocuSign Behavior
Email from unfamiliar sender	Routine—signers often receive requests from unknown parties
Urgent call to action	Standard—documents have deadlines
External link to click	Required—signatures happen on DocuSign's platform
Request for sensitive action	Expected—signing commits the user legally

Legitimate DocuSign emails routinely exhibit characteristics that security awareness training teaches users to treat as red flags:This creates what security researchers call a "business logic vulnerability." The attack vector cannot be closed without blocking legitimate business activity. Security teams cannot write rules that distinguish malicious DocuSign emails from real ones based on these surface characteristics—because they're identical.

Attackers have recognized this gap and are exploiting it systematically. DocuSign impersonation accounted for 13.8% of all advanced attacks in our Q4 2025 analysis—the single most prevalent attack theme among threats sophisticated enough to evade enterprise email security.

The Attack Pattern

DocuSign impersonation attacks follow a consistent operational model:

1. **Lure construction:** Email arrives with subject lines like "REVIEW DOCUMENT NOW" or "[Client Name] - Urgent Document Signature Required"
2. **Sender spoofing**: Domains are typosquatted or use legitimate compromised infrastructure to pass casual inspection
3. **Credential harvesting:** Links redirect to convincing Microsoft 365 login pages
4. **Account compromise:** Harvested credentials provide access to email, OneDrive, SharePoint, and connected enterprise systems

The ultimate target is rarely the DocuSign interaction itself—it's the Microsoft 365 credentials that unlock far more valuable access. A single compromised account can enable business email compromise, data exfiltration, lateral movement, and ransomware deployment.

Figure 1 - Docusign Phishing Sample

METHODOLOGY: ANALYZING WHAT GETS THROUGH

Data Collection

This analysis examines 2,500+ advanced email attacks detected by StrongestLayer between October and December 2025 across monitored enterprise environments spanning legal services, pharmaceutical, manufacturing, healthcare, and professional services sectors.

What Makes This Dataset Unique

Every attack in this analysis had already bypassed the organization's existing email security stack before StrongestLayer detection. Monitored environments included:

- Microsoft E3/E5 with Defender for Office 365
- Mimecast secure email gateway
- Proofpoint secure email gateway

This is not a sample of "all phishing attempted." It is specifically a sample of attacks sophisticated enough to evade state-of-the-art detection—the attacks that actually reach users and cause breaches. When DocuSign impersonation emerges as the dominant theme in this dataset, it indicates where the real detection gap exists.

Analysis Framework

Each attack was analyzed across multiple dimensions:

- **Attack theme classification:** Brand impersonation, business process exploitation, payload type
- **Jaccard similarity scoring:** Feature overlap between attacks within the same theme
- **AI-assistance indicators:** Contextual sophistication, linguistic patterns, infrastructure characteristics
- **Evasion techniques:** Authentication bypass, URL obfuscation, content variability

WHY SIGNATURES FAIL: THE JACCARD EVIDENCE

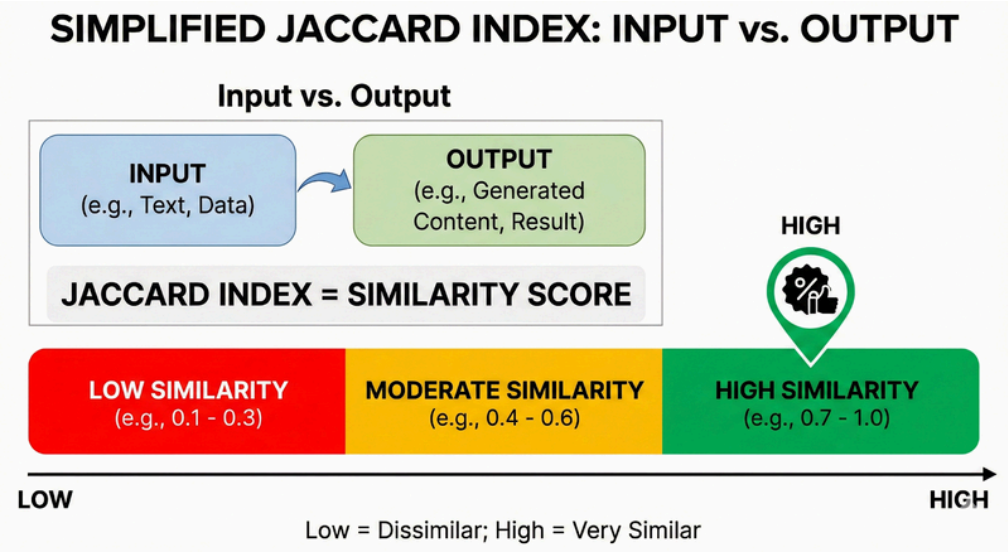
How Signature-Based Detection Works

Traditional email security relies on signatures—rules that identify attacks based on shared characteristics. When a security vendor identifies a phishing campaign, they extract distinguishing features (sender patterns, URL structures, content fingerprints) and create rules to block matching emails.

This approach works when attacks are similar to each other. If a threat actor sends 10,000 phishing emails using the same template, infrastructure, and payload, a single signature can block the entire campaign.

The Similarity Problem

The effectiveness of signature-based detection depends entirely on how similar attacks are to each other. The Jaccard similarity index quantifies this by measuring feature overlap between attacks. A score of 1.0 means two attacks are identical; a score of 0.0 means they share no features.



Jaccard Score	Signature Effectiveness	Implication
0.85 - 1.0	High	One signature catches most variants
0.50 - 0.85	Moderate	Multiple signatures needed
0.30 - 0.50	Low	Signatures catch minority of variants
Below 0.30	Ineffective	Pattern matching breaks down

Traditional template-based phishing campaigns typically exhibit Jaccard scores of 0.85-0.95. Attackers reuse infrastructure, copy templates, and make minimal modifications. Security vendors can identify a campaign, write signatures, and block subsequent attacks effectively.

DocuSign Attacks Break the Model

Analysis of the 281 DocuSign impersonation attacks revealed an average Jaccard similarity of just 0.458—meaning each attack shares less than half its features with other DocuSign attacks in the same period.

At this similarity level, a signature that perfectly matches one attack will fail to match more than half of the other variants. Security teams cannot write rules fast enough to keep pace with attack variability.

The problem extends beyond DocuSign. Across all 2,500+ advanced attacks analyzed, 68% fell below the 0.30 Jaccard threshold—the point we term the "Pattern Matching Cliff," where signature effectiveness drops below statistical significance.

The False Positive Trap

Low similarity scores create a secondary problem beyond missed detections. To catch highly variable attacks, security teams must write broader rules—looser patterns with fewer required conditions.

But broader rules inevitably flag legitimate emails that share some characteristics with attacks. Organizations face an impossible choice:

- **Tight rules:** Low false positives, but miss most attack variants (detection gap)
- **Broad rules:** Catch more variants, but flood security teams with false positives (operational failure)

DocuSign impersonation exemplifies this trap. Any rule broad enough to catch the variants would also flag a significant percentage of legitimate DocuSign traffic—blocking business operations and generating alert fatigue that degrades security team effectiveness.

THE AI ACCELERATION

Evidence of AI-Assisted Attack Generation

Analysis indicates 38% of DocuSign impersonation attacks showed high AI-assistance indicators:

- **Contextually appropriate language:** References to specific transaction types, industry terminology, realistic business scenarios
- **Legitimate-looking infrastructure:** Domains and sending patterns that mimic authentic business communication
- **Sophisticated workflow understanding:** Urgency framing aligned with actual business processes rather than generic "click now" demands

These aren't the crude "Dear Customer, Click Here to Sign Document" templates of traditional phishing. AI-generated attacks reference specific transaction types, use terminology appropriate to the target's industry, and create urgency that mirrors legitimate business pressure.

Why AI Compounds the Detection Problem

The combination of AI generation and low Jaccard similarity creates a compounding problem:

1. Unlimited variant generation: AI can produce thousands of unique attack variants at near-zero marginal cost
2. No consistent signature: Each attack is different enough to evade rules
3. Consistent intent: All variants pursue the same objective (credential harvesting), making them effective despite variability
4. Continuous evolution: AI-generated attacks can adapt faster than signature updates

Current data suggests 35-45% AI-assistance across all advanced attack categories in Q4 2025. Based on observed acceleration, we project this will reach 60-75% by Q4 2026. Organizations relying solely on pattern-matching face a growing blind spot as AI becomes the default attack construction method.

DOCUMENT-HEAVY INDUSTRIES FACE ELEVATED RISK

DocuSign impersonation disproportionately impacts industries where electronic signatures are woven into daily operations:

Legal Services

Attorneys routinely receive legitimate DocuSign requests from clients, courts, opposing counsel, expert witnesses, and co-counsel—often from unfamiliar senders with genuine urgency. Court deadlines, closing timelines, and client demands create time pressure that overrides security skepticism.

The consequences of successful attacks extend beyond credential theft. One prevented incident involved an attorney who nearly sent privileged documents to what appeared to be co-counsel but was actually an impersonation attack—estimated malpractice exposure of $1.5 million.

Pharmaceutical and Healthcare

Clinical trial agreements, vendor contracts, regulatory submissions, and patient consent forms drive constant DocuSign volume. A compromised credential in these environments can expose clinical data, regulatory filings, and intellectual property.

Financial Services and Real Estate

Transaction velocity creates operational pressure to process signature requests quickly. Wire fraud, closing document manipulation, and account takeover all begin with credential compromise.

The Common Thread

Any industry where high-volume document signing is standard faces elevated DocuSign impersonation risk. The attack exploits operational necessity—organizations cannot slow down document workflows without business impact, and that urgency creates the opening attackers exploit.

DETECTION REQUIREMENTS FOR VARIABLE ATTACKS

When attacks exhibit low similarity scores and high variability, effective detection requires a fundamentally different approach than signature matching. Based on analysis of what distinguishes detected threats from false positives, the following capabilities are required:

1. Intent Analysis Over Pattern Matching

Variable attacks share intent even when they share few technical features. Effective detection must evaluate what an email is trying to accomplish:

- Is this email attempting to harvest credentials?
- Does the requested action benefit the sender in ways inconsistent with stated purpose?
- Does the business scenario described match legitimate transaction patterns?

2. Business Context Understanding

The same email can be legitimate or malicious depending on context:

- Has this recipient received DocuSign requests from this sender type before?
- Does the transaction type match the recipient's role and responsibilities?
- Does the urgency level match actual business timelines?

3. Real-Time Reasoning

Pre-computed signatures cannot adapt to novel variants. Detection must reason about each email individually:

- Evaluate sender, content, links, and context together
- Weigh multiple signals rather than matching single patterns
- Adapt conclusions based on specific organizational context

4. Low False Positive Tolerance

Detection that blocks legitimate business email is operationally unacceptable. Effective solutions must maintain low false positive rates even while catching variable attacks—a requirement that signature-broadening approaches cannot meet.

CONCLUSION: THE ARCHITECTURAL DECISION

The DocuSign detection gap illustrates a broader shift in email security. When attackers can generate unlimited unique variants through AI, the only stable signals are malicious intent and business context. Technical signatures that work against template-based campaigns break down against variable attacks.

Organizations face an architectural choice:

Option A: Continue signature-based approaches

- Increasing investment in rule writing and tuning
- Accepting growing detection gaps as attack variability increases
- Managing escalating false positives from rule broadening

Option B: Adopt reasoning-based detection

- Evaluate intent and context rather than matching patterns
- Maintain effectiveness against variable attacks
- Preserve low false positive rates without sacrificing detection

The 2,500+ attacks in this analysis share one characteristic: they were sophisticated enough to evade the best pattern-matching defenses enterprises currently deploy. That's not a sample of "all phishing"—it's a sample of what actually causes breaches. As AI-generated attack variability accelerates, the gap between signature-based detection and actual threat landscape will widen. Organizations that recognize this architectural shift now can adapt before the detection gap becomes a breach.