2

15. Mai 2026
Fritz Knapp Verlag
37. Jahrgang
ISSN 0937-597 X
D 25079

# KARTEN
cards | cartes

ZEITSCHRIFT FÜR ZAHLUNGSVERKEHR UND PAYMENTS

Digitaler Sonderdruck

# PAYMENT-INFRASTRUKTUR IM UMBRUCH

„Agentic Commerce führt zu einem strukturellen Betrugsproblem“

Interview mit Adam Davies

# „Agentic Commerce führt zu einem strukturellen Betrugsproblem“

## Interview mit Adam Davies

Foto: AdobeStock/kyo

Betrug hat sich weit über das hinaus entwickelt, was traditionelle Erkennungsmodelle erfassen, sagt Adam Davies. Dabei haben sowohl KI als auch Echtzeitzahlungen das Spielfeld massiv verändert. Und mit dem agentischen Handel entsteht ein ganz neues strukturelles Betrugsproblem, weil der Mensch als Kontrollinstanz ausfällt. KI hat aber nicht nur das Bedrohungsniveau erhöht, sondern verbessert auch die Verteidigungsfähigkeit der Banken. Die Umsetzung bleibt allerdings einstweilen noch eine Herausforderung. Und weiterhin bleibt der Mensch das schwächste Glied im digitalen Ökosystem. Red.

**KARTEN** **Künstliche Intelligenz hat klassische Betrugsmuster verändert. Welche spezifischen KI-gestützten Angriffe sehen Sie derzeit als die größte Gefahr für Banken in der Euro-Zone?**

Die größte Veränderung besteht darin, dass Betrüger mit KI hyperrealistische Betrugsmaschen entwickeln und diese in großem Maßstab einsetzen können.

Eine große Bedrohung ist der Einsatz von Deepfakes, bei denen Betrüger mit KI Sprach- und Videoinhalte generieren, um sich in Echtzeit als echte Person auszugeben. Mit generativer KI können Betrüger außerdem Inhalte für viele Sprachen und Kontexte erstellen, um Einzelpersonen in größeren geografischen Regionen gezielt anzusprechen. Angreifer produzieren heutzutage mit KI fehlerfreie, lokalisierte Messages oder E-Mails in jeder europäischen Sprache, zugeschnitten auf spezifische Banken oder regulatorische Anforderungen. Mit KI können Betrüger zudem sehr schnell synthetische Identitäten erstellen, indem sie reale und erfundene Daten verwenden, um Onboarding-Kontrollen zu umgehen. Das wirkt sich besonders in Märkten mit hoher Nutzung von Digital Banking aus.

**KARTEN** **Wie verändern KI-gestützte Karten-Scams, SIM Swapping und personalisierte Social-Engineering-Angriffe die Art und Weise, wie Betrug identifiziert werden muss?**

Betrug hat sich weit über das hinaus entwickelt, was traditionelle Erkennungsmodelle erfassen. Mit KI-gestützten Tools können Betrüger heutzutage hochgradig überzeugende Social-Engineering-Angriffe, synthetische Identitätsmodelle und Deepfakes in großem Maßstab ausführen. Laut einer von Fico in Auftrag gegebenen Studie sehen Finanzinstitute erhebliche operative Auswirkungen auf ihre Betrugsprogramme als Folge dieser KI-gestützten Angriffsarten.

Gleichzeitig können Kriminelle mit Taktiken wie SIM-Swapping SMS-basierte Authentifizierungsmechanismen umgehen. Dabei übernehmen sie die Mobilfunknummer eines Opfers, um Einmalpasswörter abzufangen und ein Konto zu übernehmen. Für Standard-Erkennungssysteme erscheint das völlig legitim.

Betrug mit autorisierten Push-Zahlungen (APP) umgeht Identitätsprüfungen vollständig. Hierbei werden legitime Kunden dazu gebracht, betrügerische Transaktionen selbst zu autorisieren – häufig während der Betrüger sie aktiv in einem laufenden Telefongespräch anleitet.

Die zentrale Herausforderung bei der Betrugserkennung besteht darin, dass diese Angriffe oberflächlich nicht wie Betrug aussehen. Der Kunde ist echt, das Gerät ist bekannt, und

Foto: Fico

Adam Davies, Vice President – Fraud & Identity Product Management, FICO Deutschland GmbH, Berlin

die Authentifizierung ist erfolgreich. Um solche Angriffe zu stoppen, ist ein tiefergehender Ansatz erforderlich, der Behavioral Analytics, Echtzeit-Telekommunikationssignale und KI-Modelle kombiniert, die speziell darauf ausgelegt sind, Zwang und Manipulation zu erkennen – nicht nur unbefugten Zugriff. Klassische Fraud Scores erkennen nicht, wenn ein Kunde zwar selbst eine Zahlung ausführt, dies aber unter Druck tut. Dafür ist ein dedizierter Scam Detection Score erforderlich, der die Wahrscheinlichkeit bewertet, dass ein legitimer Kunde unter dem Einfluss eines Betrügers handelt.

»Betrüger testen bereits agentische Systeme.«

**KARTEN** Welche Risiken birgt der agentische Handel? Und wie kann man ihnen begegnen, wenn die KI weitgehend eigenständig bestellt und bezahlt und der Mensch als Kontrollinstanz ausfällt?

Agentic Commerce – also wenn KI-Systeme autonom Produkte oder Dienstleistungen suchen, auswählen und bezahlen, ohne dass ein Mensch jede Transaktion freigibt – führt zu einem strukturellen Betrugsproblem. Die meisten Zahlungskontrollen und rechtlichen Rahmenwerke wurden unter der Annahme entwickelt, dass ein Mensch im Moment der Autorisierung anwesend ist. Wenn Menschen nicht mehr beteiligt sind, entfällt diese Kontrollinstanz.

So sind die Risiken real und unmittelbar. Betrüger testen bereits agentische Systeme, um die Schwellenwerte zu identifizieren, innerhalb derer autonome Agenten Transaktionen durchführen können, ohne Alarme auszulösen. Neue Zahlungsmethoden, die von solchen Agenten genutzt werden könnten (zum Beispiel Kryptowährungen oder Stablecoins), bieten möglicherweise keine Chargeback-Rechte und keinen klaren Haftungsrahmen. Zudem können Agenten fehlkonfiguriert oder manipuliert werden, sodass sie gegen Verbraucherinteressen handeln.

Zur Risikominderung sind mehrschichtige Kontrollen erforderlich: Ausgabenlimits, Transaktionsfrequenzgrenzen, Whitelists für Händler, tokenisierte, agentenspezifische Zugangsdaten sowie eine Eskalation an Menschen bei Transaktionen außerhalb definierter Parameter. Auf der Detection-Ebene benötigen Finanzinstitute Behavioral Analytics-Tools, die in der Lage sind, legitime Agentenaktivitäten von kompromittierten oder bösartigen Agenten zu unterscheiden – eine Fähigkeit, die die meisten derzeit erst aktiv aufbauen.

**KARTEN** Klassische regelbasierte Systeme zur Betrugsprävention stoßen im KI-Umfeld an ihre Grenzen. Ab welchem Punkt wird ein regelbasiertes System zur echten Haftungsfalle für Banken?

Systeme, die ausschließlich auf Regeln basieren, sind in der sich schnell verändernden Betrugslandschaft bereits ein Haftungsrisiko. Im Kartenbereich sind KI-basierte Systeme wie neuronale Netze bereits seit 1992 im Einsatz, als der Falcon Fraud Manager eingeführt wurde. Heute schützt dieses System weltweit rund vier Milliarden Zahlungskarten. Die überlegene Mustererkennung von KI ist unverzichtbar, weil Kriminelle seit Langem in der Lage sind, Regeln zur Betrugserkennung zu identifizieren und zu umgehen.

**KARTEN** Ein „Zauberwort" lautet kontextbasierte Entscheidungen. Was genau ist darunter zu verstehen? Und wie verbindet KI Transaktionsdaten mit Verhaltensanalysen und externen Signalen, um beispielsweise einen Fake Store in Echtzeit zu entlarven?

Contextual Decisioning bedeutet, jede Transaktion nicht isoliert zu bewerten, sondern im Kontext dessen, was für diesen Kunden, dieses Konto, dieses Gerät und diesen Zeitpunkt normal ist. Eine Transaktion, die isoliert betrachtet verdächtig erscheint, kann im Kontext völlig legitim sein – und umgekehrt. Ziel ist es, intelligentere Entscheidungen zu treffen, die zwar Betrug verhindern, aber legitime Kunden nicht beinträchtigen.

In der Praxis bedeutet das, alle Eingaben kontinuierlich und gleichzeitig zu bewerten und zu verstehen. Erstens erstellen Behavioral Analytics-Tools dynamische Profile für jeden Kunden, Händler, jedes Konto und jedes Gerät. Mit jeder Interaktion in Echtzeit werden diese Profile aktualisiert, sodass jede Abweichung von etablierten Mustern sofort als Anomalie erkannt wird. Zweitens liefern Transaktionsdaten das unmittelbare Signal: was gekauft wird, für welchen Betrag, über welchen Kanal und auf welchem Gerät. Drittens ergänzen externe Signale – einschließlich Identitätsprüfung durch Drittanbieter, Device Intelligence, Netzwerkanalysen und in einigen Fällen Echtzeit-Telekommunikationsdaten – den Kontext um Ebenen, die keine einzelne Datenquelle allein geben kann.

Auf das Szenario eines Fake-Online-Shops angewendet, ist diese Kombination besonders wirkungsvoll. Netzwerk-Analytics können verborgene Verbindungen zwischen einem neuen Händler und bekannter Betrugsinfrastruktur aufdecken – gemeinsame Telefonnummern, verknüpfte Konten oder Muster, die auf Betrugsringe hindeuten. Verhaltensprofile können erkennen, dass ein Kunde mit einer Händlerkategorie interagiert, die er zuvor nie genutzt hat, auf einem Gerät, das nicht seiner Historie entspricht, und zu einer für ihn oder sie ungewöhnlichen Zeit. Externe Daten wiederum können die Legitimität des Händlers in Echtzeit bestätigen oder eben nicht. Kein einzelnes dieser Signale löst einen Alarm aus. Es ist das Zusammenspiel der Signale, das zur Betrugserkennung führt.

»Die überlegene Mustererkennung von KI ist unverzichtbar.«

**KARTEN** Wie vertragen sich solche datenbasierten Betrugspräventionssysteme mit europäischen Datenschutzregeln?

Datengetriebene Präventionssysteme sind gut mit Datenschutzvorschriften in ganz Europa vereinbar, wenn sie nach den Prinzipien „Privacy by Design" und „Privacy by Default" entwickelt werden. In der Praxis bedeutet das, dass Finanzdienstleister auf klare Rechtsgrundlagen wie berechtigtes Interesse oder gesetzliche Verpflichtungen zurückgreifen, die Menge der verarbeiteten Kundendaten minimieren und Anonymisierung sowie Verschlüsselung einsetzen. Um die Anforderungen der DSGVO einzuhalten, müssen Finanzinstitute automatisierte Entscheidungsfindung mit menschlicher Expertise kombinieren. So entstehen vertrauenswürdige Entscheidungen, die transparent und erklärbar sind.

**KARTEN** Wie lässt sich verhindern, dass die KI-Modelle zur Betrugsabwehr selbst Opfer von Angriffen und damit quasi ausgehebelt werden?

Die Geschwindigkeit, mit der KI vorangetrieben wird, wirft diese interessante Frage sicherlich auf. Um zunehmend ausgefeilte KI- oder agentenbasierte Betrugsvektoren zu bekämpfen, müssen Banken mehr als traditionelle prädiktive KI, Regeln und Fallmanagement einsetzen. KI und Agenten werden über den gesamten Betrugslebenszyklus hinweg benötigt – von der Prävention, Erkennung und Analyse bis hin zur Nachverfolgung.

Banken tun gut daran, darauf zu achten, dass künstliche Intelligenz schneller auf neue Betrugsbedrohungen reagiert, und die Fähigkeit verbessert, neue Strategien und neue Orchestrierungsabläufe zu implementieren. KI-Agenten müssen Trends analysieren und Gegenmaßnahmen vorschlagen. Und für Analysten und Ermittler ein KI-gestütztes Fallmanagement darstellen, damit sie Fälle schneller und genauer bearbeiten können. Jeder dieser Ansätze wirkt als mehrstufige, durch agentenbasierte KI gesteuerte Pipeline mit menschlicher Kontrolle in Form von Freigabe, Anleitung und Feedback. Nur so sind Banken in der Lage, bestehende Schwachstellen zu adressieren und künftige Betrugsvektoren zu bewältigen, die mit der zunehmenden Nutzung von KI durch Angreifer einhergehen.

Banken können ihre KI-Betrugsmodelle schützen, indem sie adaptive Komponenten integrieren, die das Scoring-Verhalten kontinuierlich verändern. Für die Angreifer wird es dann schwieriger, das Modell genau zu replizieren. Darüber hinaus kann defensive KI Angreifer aktiv täuschen. Dafür erzeugt sie irreführenden Output als Antwort auf Sondierungsaktivitäten und korrumpiert damit die offensive KI der Kriminellen. Und sie platziert sogar detektierbare Muster, die später genutzt werden können, um sie zu stellen. All dem liegt eine „Security-by-Design"-Philosophie zugrunde, bei der Modelle von Grund auf so entwickelt werden, dass sie sich selbst überwachen, Angriffe in Echtzeit erkennen und dynamisch reagieren, anstatt statisch und anfällig zu bleiben.

»Banken können ihre KI-Betrugsmodelle schützen, indem sie adaptive Komponenten integrieren.«

**KARTEN** Der Spagat zwischen Sicherheit und Betrugsbekämpfung war schon immer schwierig. Hat sich das im KI-Umfeld noch einmal verschärft? Oder ermöglicht KI mehr Sicherheit bei gleichzeitig besserer oder zumindest gleicher Nutzererfahrung?

Die ehrliche Antwort ist: beides. KI hat das Bedrohungsniveau erhöht, aber sie erhöht auch die Verteidigungsfähigkeiten von Finanzinstituten, die sie richtig einsetzen. Die Herausforderung besteht darin, dass sich beide nicht mit derselben Geschwindigkeit entwickeln.

Auf der Bedrohungsseite können Betrüger mit KI Angriffe automatisieren, Betrugsmaschen personalisieren und Erkennungssysteme in einem bisher nicht gekannten Maßstab und mit neuer Raffinesse testen. Betrüger unterliegen keinen regulatorischen Anforderungen oder Haftungsfragen; sie können frei testen, iterieren und ihre Methoden optimieren. Legacy-Systeme wurden für diesen Umfang nie entwickelt. Unsere aktuelle Studie zeigt, dass weltweit jedes dritte Finanzinstitut „hohe" oder „sehr hohe" False-Positive-Raten verzeichnet – mit realen Auswirkungen auf Kunden: abgelehnte Transaktionen, unnötige Verifizierungsschritte und Vertrauensverlust in entscheidenden Momenten.

Gleichzeitig bietet KI eine starke Antwort auf dieses Problem, wenn sie richtig eingesetzt wird. Dort, wo Finanzinstitute von fragmentierten Legacy-Systemen zu integrierten KI- und Machine-Learning-Plattformen übergegangen sind, zeigt sich ein deutlich anderes Bild. Moderne KI-Modelle erkennen Betrug präziser, sodass legitime Kunden Sicherheit erleben, die sich eher unsichtbar als invasiv anfühlt. Eine bessere Erkennung reduziert sowohl Betrugsverluste als auch die unnötige Beeinträchtigungen legitimer Kunden.

Die operative Umsetzung bleibt jedoch eine große Herausforderung. Unsere aktuelle Studie zeigt, dass nur 28 Prozent der Banken KI-basierte Betrugserkennung vollständig und im großen Maßstab implementiert haben. Das verdeutlicht die Lücke zwischen den Möglichkeiten von KI und der tatsächlichen Leistungsfähigkeit der Systeme.

»Die operative Umsetzung bleibt eine große Herausforderung.«

**KARTEN** Wie wirken sich Echtzeitzahlungen auf die Betrugsentwicklung aus? Wie gut kann Betrugser-

**kennung im Instant-Payment-Umfeld funktionieren?**

Echtzeitzahlungen haben die Betrugslandschaft grundlegend verändert. Wenn Zahlungen sofort und unwiderruflich abgewickelt werden, schrumpft das Zeitfenster für die Erkennung auf Millisekunden. Betrüger haben dies früh erkannt und betrachten insbesondere APP-Betrug als attraktiv, da Geld sofort transferiert und über mehrere Konten verteilt werden kann, bevor eine Analyse überhaupt beginnt.

»Banken müssen früher im Betrugslebenszyklus ansetzen.«

Die Herausforderung geht aber über die Geschwindigkeit allein hinaus. APP-Betrug sieht zum Zeitpunkt der Transaktion nicht wie Betrug aus, da der Kunde echt ist, das Gerät erkannt wird und die Authentifizierung erfolgreich ist. Klassische Modelle gegen Betrug wurden dafür nie entwickelt. Effektive Erkennung erfordert KI-Modelle mit integriertem Scam Score wie das Consumer-Payments-Modell von Fico. Sie sind darauf ausgelegt, zu erkennen, ob das Verhalten eines legitimen Kunden plötzlich auf eine Betrugsbeeinflussung hindeutet.

Banken, die solche spezialisierten Modelle und Scores einsetzen und in der Lage sind, in Echtzeit zu intervenieren und mit Kunden zu interagieren, können Betrug auch im Instant-Payment-Umfeld wirksam stoppen – allerdings nur, wenn sie über die dafür geeigneten Systeme verfügen.

Auch der regulatorische Druck nimmt zu. Beispielsweise schaffen die Erstattungsregeln für APP-Betrug in Großbritannien eine direkte und wachsende finanzielle Haftung für Institute, die keine geeigneten Kontrollmechanismen einsetzen.

**KARTEN Wie müssen Banken ihre Fraud-Strategie anpassen, um trotz des wachsenden regulatorischen Drucks profitabel zu bleiben? Und lässt die Verlagerung des Risikos auf Banken die Verbraucher nicht auch ein wenig zu sorglos werden?**

Die Verlagerung der Haftung für Betrugsverluste auf Banken bedeutet, dass traditionelle „Detect-and-Recover"-Modelle nicht mehr ausreichen. Um profitabel zu bleiben, müssen Banken früher im Betrugslebenszyklus ansetzen und stark in Echtzeit-Erkennung investieren, die durch Dynamic Profiling gesteuert wird und Entscheidungen auf Basis des aktuellsten und präzisesten Kundenkontexts ermöglicht.

Die Fähigkeit, risikoreiche Interaktionen frühzeitig zu erkennen und proaktiv zu unterbrechen, ist entscheidend, um das Vertrauen der Kunden zu stärken und die Profitabilität vor steigenden Betrugskosten zu schützen.

Ein weiterer Faktor ist Automatisierung: Manuelle Prozesse führen häufig zu hohen operativen Kosten und hemmen die Ausweitung von Erkennungsstrategien, da Ressourcen nicht mit dem Betrugsvolumen Schritt halten können.

**KARTEN Wie kann die branchenübergreifende Zusammenarbeit, etwa von Telekommunikationsunternehmen und Banken, verbessert werden, um Betrugsversuche frühzeitig zu erkennen?**

Branchenübergreifende Zusammenarbeit ist entscheidend zur Bekämpfung von Betrug. Beispielsweise werden konsortiale Daten von Tausenden von Finanzinstituten genutzt, um die neuronalen Netzwerkmodelle im Falcon Fraud Manager zu trainieren. Das macht diese Modelle besonders leistungsfähig.

Im Telekommunikationsbereich haben Mobilfunkanbieter in Großbritannien, in Südafrika und Spanien einen Dienst namens „Scam Signal" für Banken ins Leben gerufen. Diese Lösung nutzt Echtzeit-Telefoniedaten, um APP-Betrug zu erkennen und zu stoppen. Bei dieser Betrugsform werden Opfer am Telefon dazu gebracht, Geld an Kriminelle zu überweisen.

Die Lösung kombiniert Echtzeit-Daten aus Telekommunikationsnetzen mit Kunden- und Zahlungsinformationen während laufender Transaktionen. Wird ein potenzieller Betrug erkannt, nutzt die Bank automatisierte Kommunikationsdienste, um die Manipulation zu unterbrechen und den Kunden zu kontaktieren.

**KARTEN Was kann die europäische digitale Identität in Sachen Sicherheit bringen? Ist sie eine echte Hilfe für die Betrugsprävention – oder kann sie womöglich zu einem Einfallstor werden, das es Betrügern noch leichter macht, wenn es ihnen gelingt, eine solche Identität zu hacken?**

Digitale Identitäten haben immer Vorteile für die Verbesserung des Betrugsmanagements. Jedes Mal, wenn Daten integriert, abgerufen und für Betrugssysteme sowie KI-Modelle/Regeln/Strategien verfügbar gemacht werden, entstehen so mehr digitale Spuren, die für eine präzise Betrugsverfolgung erforderlich sind. Die Schwäche in jedem System liegt jedoch in der Robustheit der Sicherheit und im Faktor Mensch.

Das Risiko von Datenverletzungen beziehungsweise Kompromittierungen besteht immer. Und immer ausgefeiltere Social-Engineering-Methoden, Betrugsmaschen und ihre Ausnutzung bedeuten, dass Menschen stets das schwächste Glied in jedem digitalen Ökosystem darstellen. Glücklicherweise werden Banken besser bei der Erkennung, da sie dieses Verhalten auch auf ihren bestehenden digitalen Kanälen beobachten. ■

»Digitale Identitäten haben immer Vorteile für die Verbesserung des Betrugsmanagements.«