Picture credit: AdobeStock

TAPPING TOWARDS DANGER: THE RISE OF NFC RELAY MALWARE

Krishna Vishnubhotla *examines the rising threat of Near-field communication hackers*

As an old saying in cyber security goes: attackers will always find the path of least resistance. We shouldn't be surprised to see that cyber criminals have found a new way in which to reap profits from unsuspecting victims by exploiting a vector that – despite being a technology used every day by countless people – is often overlooked as a potential risk. We talk, of course, of Near Field Communication (NFC).

Yet this form of attack is exploding in popularity in the cyber criminal underground, with mobile malware variants evolving quickly and a lively industry emerging to serve this new tactic. The mere fact that both mobile devices and NFC readers are so deeply embedded in everyday life yet often ignored as a potential risk, makes NFC relay attacks a significant and stealthy opportunity for attackers and a growing danger for normal users.

Tapping our phones against a payment terminal to make transactions is a relatively new development, but one that has become so common we barely think about it. Tap-to-pay depends on two things: the NFC hardware that handles the tap itself and the secure storage that holds the payment credentials.

Relay attacks work because they exploit a part of the device that traditional mobile defences rarely monitor

NFC works the same on Android and iOS, but the way payment data is stored is different. On Android, wallets can choose where to store and process those credentials – either in the Secure Element or software through Host Card Emulation. That flexibility is why you see Google Pay and many third-party wallets built by banks, fintechs and regional payment providers.

On iOS, payment data can only live in the Secure Enclave and only Apple's system-level wallet can access it. Third-party apps cannot emulate a card or store payment credentials themselves, which is why Apple Pay is the only tap-to-pay wallet on iPhones. Together, these technologies enable digital wallets to work without the physical card and have become central to how we make everyday transactions.

Contact between a mobile device's NFC sensor and a payment card allows the phone to read the card and store its details as a token, which is then added to the digital wallet. This is the same process used when you hold your card near the phone to "add to wallet." A token is a stand-in for your real card number. Instead of storing the actual card details, the wallet stores a unique string of numbers and letters that represents your card only inside that payment system. It looks nothing like the real card number.

Google Pay uses the Secure Element on most modern Android devices, so the token stays in hardware. Many third-party wallets rely on HCE, so the token is stored and managed in software instead. The Android OS allows both approaches and each wallet chooses the model it supports based on whether it has access to the Secure Element or not. On iOS, all wallets store tokens the same way because Apple does not allow any variation. Every token is stored inside the hardware Secure Element and only Apple Pay can access or present it. Third party apps cannot store tokens themselves or emulate cards in software.

When you bring your phone near a payment terminal, the terminal's radio field induces a current that wakes up the phone's NFC chip. Instead of providing actual card details, the phone sends a secure, one-time token to the terminal. The terminal then processes the token as if it were a regular card number and the payment goes through.

On Android, Host Card Emulation (HCE) allows any app to act as a mobile wallet and store payment data in software. When this happens, it shifts the burden of protecting that information entirely to the wallet app, rather than relying on secure hardware. Attackers exploit this by creating fake or repackaged wallet apps that mimic the same HCE logic or by installing malware that can trigger that logic and make the wallet behave as if a tap is happening. This exposes the payment flow in software and sets the stage for more advanced attacks like NFC relay malware.

NFC relay malware is designed to trigger the live payment flow inside Android wallets that use HCE, letting attackers run a transaction on a terminal they control as if the victim were physically tapping it. Relay malware does not steal the victim's card number or a reusable token. Instead, it forwards the real-time NFC command-and-response (APDU) messages that a terminal normally exchanges with the wallet during a tap. These messages include the dynamic cryptogram and other one-time values needed to authorise that specific transaction. They are only valid for a few seconds, but that is enough for an attacker using a modified POS terminal or NFC emulator to run a charge on their side. The attacker never gets a card, they simply borrow the victim's payment capability at that moment.

That tactic has proved remarkably successful for cyber criminals and is now surging in popularity. According to ESET's 2025 Threat report, the growth in NFC-based attacks has spiralled in the last year alone. In the first half of 2024, they would typically see only one NFC scam detection a week. By the first half of 2025, authors note, it had become dozens of times a week. Still, that wasn't even the biggest spike of the last two years. In fact, ESET's report shows a 35-fold increase in NFC scams from the first to the second half of 2025.

EFFORTS TO EXPLOIT NFC TECHNOLOGY ARE BEING DEVELOPED, SCALED AND AUTOMATED

The sophistication of these attacks is actively progressing too – with new NFC-targeting mobile malware variants appearing in quick succession, iterating and evolving to perform stealthier attacks and outwit cyber defences. As ESET's Senior Malware Researcher, Lukáš Štefanko, notes in the report: "Each iteration of NFC fraud demonstrates how attackers adapt to new security measures. Even advanced solutions – like multifactor authentication or real-time transaction monitoring – face challenges when criminals physically relay the card data in seconds."

NGate was one of the first to be recognised as an NFC relay malware. Distributed through phishing websites which compelled users to download an app that – ironically – promised to secure their payment cards, the app would compel victims to tap their cards against their devices' NFC reader and hand off that card info to attacker-controlled devices.

NFC relay malware quickly evolved from there. Relay NFC was discovered in late 2025 and, upon discovery, researchers noted how lightweight the malware was – requiring few permissions in order to avoid detection. It also used a hermes-compiled payload, which made it hard to analyse and obfuscated the underlying logic thus frustrating static inspection. Its anti-detection capabilities were apparently quite effective, and security researchers noted that anti-malware engines could not detect it.

SuperCardX has also emerged around the same time. Upon analysis, researchers noted how difficult it was to detect and mitigate. Firstly, it can often circumvent detection because it merely appears as a harmless NFC app, operating with few permissions while also employing advanced encryption, making it difficult for security researchers to reverse engineer and analyse.

That evolution reveals what has made NFC Relay malware so successful, so quickly. Relay attacks work because they exploit a part of the device that traditional mobile defences rarely monitor. These apps often look harmless, request minimal permissions and use heavy obfuscation, making them difficult to detect. More importantly, the payment flow they target sits deep inside system components and HCE services, far outside what signature-based or network-based protections inspect. Even advanced controls like MFA or transaction monitoring struggle in the face of real-time relayed payments that appear legitimate to the issuer.

NFC-TARGETING MOBILE MALWARE VARIANTS ARE EVOLVING TO PERFORM STEALTHIER ATTACKS

Cyber criminals have spotted an opportunity here and these efforts to exploit NFC technology are actively being developed, scaled and automated. What was once a handful of isolated relay tools has quickly grown into a broader ecosystem. There are multiple reports of 'farms' using dozens of phones with multiple stolen cards loaded onto them, set up to scale the scam along with NFC-enabled POS systems they can use to 'pay' themselves with their victim's stolen card information. Threat groups are now sharing methods, trading modified POS hardware and distributing NFC relay kits through Telegram channels with thousands of members. It's being actively discussed online with security researchers flagging multiple Telegram groups devoted to the practice, with some having close to 6,000 members.

We are seeing dedicated marketplaces that sell NFC Relay malware as a service, complete with user guides, support channels and purpose-built NFC terminals designed for rapid 'cashing out'. This level of organisation shows that NFC exploitation is no longer experimental – it is being operationalised and scaled.

Organisations cannot treat this as a niche or regional issue. Relay malware has shown that attackers can weaponise the tap-to-pay stack, automate it and scale it globally with very little friction. As NFC expands beyond payments into identity verification and access control, the incentives for attackers will only grow. Defending against this new class of threat requires a shift in thinking – from scanning for known malware to monitoring behaviour on-device, understanding how apps interact with sensitive system components and detecting malicious actions as they occur ●

Krishna Vishnubhotla is Vice President of Product Strategy at Zimperium, where he has spent over a decade shaping the company's vision and advancing mobile security solutions.

ESET's report shows a 35-fold increase in NFC scams from the first to the second half of 2025

Picture credit: AdobeStock