
Protection Through Indistinguishability

DISCUSSION DRAFT | JULY 2026 | SENIOR DEFENSE AND NATIONAL-SECURITY LEADERS

Core Architecture

Klök provisions services and network-facing identities within an MNO-independent control layer. The protected handset leaves the cellular plane; a controlled gateway presents transformed identity, RF attachment, posture, and traffic context without carrier integration. The MNO sees the gateway presentation - not the handset or internal control logic.

1. Strategic Context

Modern cybersecurity has become increasingly effective at protecting content and controlling access to it. Encryption reduces the value of intercepted payloads; endpoint security protects devices; and Zero Trust continuously evaluates who or what may reach a resource [1]. These measures make stealing the message itself harder. They do not eliminate the evidence generated as people and systems communicate.

That residual evidence creates a different attack plane. Identity, location, posture, and behavior appear in subscriber, device, network, and session identifiers; RF attachment and mobility; transport fingerprints; timing, volume, and destination patterns; and application activity. An observer may not need to recover content if those signals can answer the more operationally useful questions: who is acting, where, with what system, and according to what routine. Attribution - rather than content recovery - is cybersecurity's blind spot.

Historically, reliable cross-domain attribution against a protected target was expensive. It required durable collection, access to multiple datasets, entity-resolution expertise, custom analytics, significant compute, and teams of analysts. That combination generally limited the most capable campaigns to nation-state intelligence services or similarly resourced organizations. The inference was not impossible; collection, integration, time, and cost imposed the barrier.

AI changes that economic barrier. AI systems can normalize dissimilar records, propose entity matches, detect recurring patterns, rank competing hypotheses, and maintain correlations across far more observations than an analyst team could review manually. Cloud compute and commercially available data reduce the entry cost further. AI does not create metadata an adversary cannot access, and it does not make weak evidence conclusive. It does compress the labor and time needed to fuse accessible, purchased, or compromised datasets - expanding the set of actors able to attempt attribution and increasing the speed at which it can be repeated.

The defensive objective therefore changes from invisibility to statistical indistinguishability: preserve mission activity while degrading the mapping from observable evidence to the protected person, device, system, facility, or operational role. Phase I addresses identity, location, posture, and behavior on the cellular side. Application, broker, cloud, and physical-sensor fusion remain later-stage threat models [4-5]. Residual timing, volume, destination, and mobility patterns motivate active synthetic behavior and the identity twin.

What Current Klök Changes

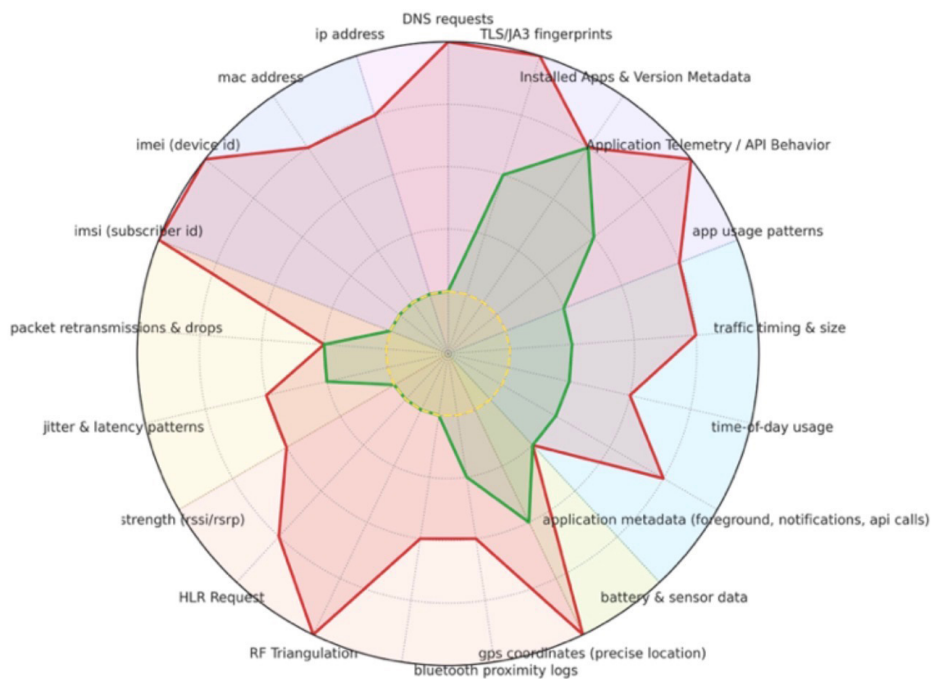
1. **Klök-controlled provisioning:** services and network-facing identities are provisioned and assigned within Klök; the handset reaches its gateway by Wi-Fi in airplane mode. The MNO authenticates access and carries traffic but does not control the user-to-service/identity mapping.
2. **Coordinated rotations:** substitutes or rotates subscriber, equipment, network, path, address, and session presentation so an unchanged feature cannot bridge successive identities.
3. **RF attachment manipulation:** selects among non-dominant but serviceable BTS candidates, disrupting the expected relationship among signal hierarchy, serving cell, location, handover path, and identity.
4. **Posture, traffic, and behavior mediation:** substitutes gateway posture; mediates DNS/IP and TLS/QUIC; and runs real gateway applications at controlled times with remote workflows and synthetic geospatial state.

5G Does Not Remove the Permanent Identity

In 5G, an IMSI-based SUPI is ordinarily concealed as a SUCI across the access link and recovered within authorized authentication functions [2-3]. The relevant protection objective is therefore mediation of the permanent identity and disruption of its linkage to the protected person - not merely concealment on the air interface.

2. Quantifying Attribution Advantage

Attribution is ultimately a probabilistic decision problem. The operational Attribution Index compares an observer's confidence before and after metadata is observed. The current workbook performs a different function: it organizes engineering expectations about which observables remain linkable under ordinary and protected conditions.



Operational Attribution Index

$$\text{Attribution Index} = 10 \times (p1 - p0) / (1 - p0)$$

$p0$ is the threat-model prior and $p1$ is the posterior after observing metadata. Labeled attribution trials supply these probabilities; the ordinal exposure positions do not

Threat Model Precedes Score

One score cannot silently combine different collection authorities. The reviewed model therefore maps every observable to one or more explicit observer classes and requires those views to be reported separately before any fusion case is considered.

1. **Carrier core and RAN:** subscription, session, serving-cell, measurement, and gateway records available through carrier or network penetration.
2. **External RF:** independent sensors observe the transmitter without access to carrier identity mappings.
3. **Network traffic observer without carrier-core access:** DNS, IP, TLS/QUIC, timing, size, destination, and session evidence visible at an ISP, VPN exit, backbone collection point, CDN, or destination service.
4. **Endpoint, provider, or broker:** application, account, precise-location, sensor, and commercial datasets that may bypass network camouflage.

Engineering Exposure Model

The executive spider chart uses a 1-to-10 ordinal scale to compare four operating conditions. It collapses identity, location, posture, and behavior for executive communication; the supporting analysis continues to evaluate each observer separately.

Attribution contribution = visibility x continuity x distinctiveness x person-binding

A feature may remain visible while contributing little to personal attribution when rotations, synthetic behavior, account discipline, or competing tracks break continuity, distinctiveness, or binding.

How to Read the Model

Score each observable through four questions: can the observer see it, follow it over time, distinguish it from peers, and bind it to the protected entity? The four displayed axes summarize that evidence after threat-model filtering; no single shape substitutes for the observer-specific analysis.

3. Engineering Illustration of the Attribution Plane

Figure 1 collapses the architecture into identity, location, posture, and behavior to show the expected progression from ordinary operation through passive controls, gateway-native application synthesis, and a physically separate identity twin.

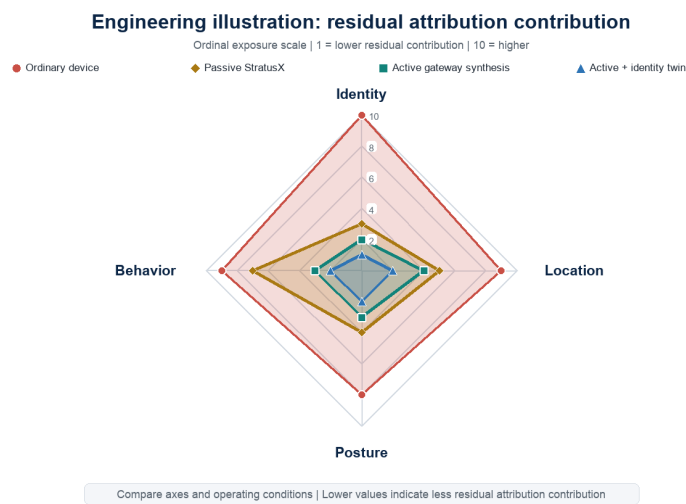


Figure 1. Engineering illustration of residual attribution contribution. Values are ordinal design hypotheses, not measured probabilities or effect sizes; lower is better.

Positions shown in the figure:

Operating condition	Identity	Location	Posture	Behavior
Ordinary device	10	9	8	9
Passive Klök	3	5	4	7
Active gateway synthesis	2	4	3	3
Active + identity twin	1	2	2	2

How Controls Move the Profile

Passive controls reduce identity continuity and substitute gateway posture. Active gateway synthesis uses real applications, controlled timing, remote operation, and synthetic GPS to reduce behavioral distinctiveness and provider-side location binding. Synthetic GPS does not move the RF transmitter; the physically separate identity twin supplies the competing physical track.

4. Protection logic and active deception

Ordinary cellular evidence can be represented as $X = (\text{identity, location, posture, behavior})$. With the handset in airplane mode, the MNO receives no live handset attachment; it instead observes the gateway's network-facing identity, serving cell, modem posture, signaling, and traffic envelope. Provisioning, mappings, rotations, RF policy, and synthetic behavior remain outside MNO operations.

Observed cellular evidence = (rotated gateway identity, altered RF attachment, gateway posture, mediated behavior)

Carrier independence is not radio invisibility: the MNO observes gateway attachment and signaling, but standard network data do not reveal the upstream handset or local policy. Attribution must reconstruct the gateway-to-person mapping.

From Passive Disruption to Active Diversion

Passive controls break direct identity and cross-session continuity. Active gateway synthesis executes real applications inside the gateway at controlled hours, presents synthetic GPS state, and supports remote workflows. It produces genuine TLS/QUIC sessions, API calls, destinations, timing, and provider telemetry - synthetic intent expressed through real traffic.

A separate, authorized Klök router can present the target-correlated identity and behavior - including equivalent IMSI/IMEI where the network arrangement permits. The primary remains more heavily defended; the identity twin supplies the cleaner attribution narrative and a competing physical track.

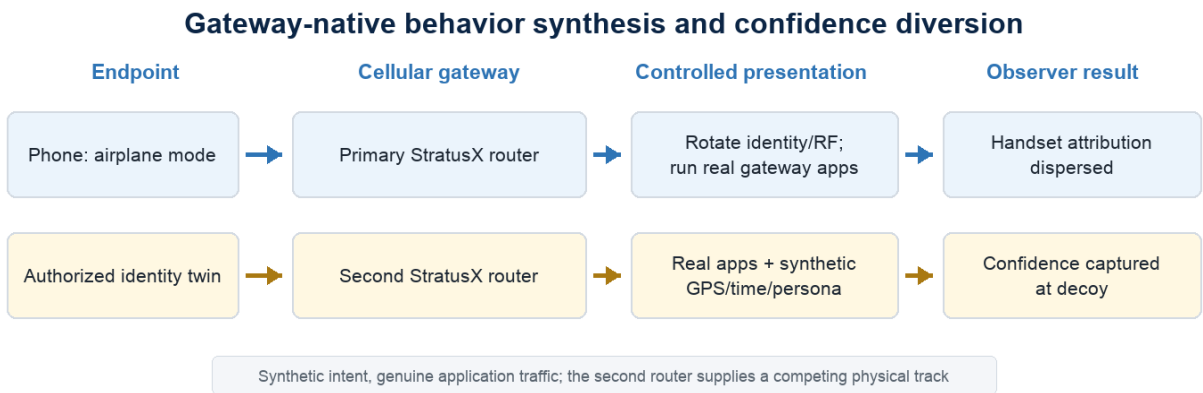


Figure 2. Gateway-native execution creates genuine application traffic under controlled identity, time, location, and persona context; an authorized second gateway adds a competing physical attribution track.

The twin is not a loud beacon. It presents the residual observability expected of a guarded target and leaves enough coherent evidence to sustain collection. The observer is not assumed to know Klök is present; the strongest evaluation case includes mechanism-aware and adaptive observers.

Real-target posterior = real-track likelihood / combined likelihood of real and twin
With equal priors, equally plausible tracks divide confidence approximately 50/50; additional credible tracks divide it further.

Implementation and Safety Boundry

Identity equivalence must be authorized and credentialed; IMSI/SUPI alone does not authenticate. Keep synthetic behavior isolated from mission traffic, control logic, and physical processes [3, 5].

5. Validation Standard and Decision Criteria

The evaluation should be a comparative attribution experiment across independent people, devices, routes, days, facilities, software versions, and mission windows. Each condition must preserve the same mission activity, and packet volume cannot substitute for independent experimental units.

Required Comparison Conditions

- **Ordinary:** the same activity without Klök.
- **Passive:** handset separation, rotations, RF-attachment manipulation, and gateway mediation.
- **Active gateway synthesis:** gateway-resident applications, controlled schedules, remote workflows, synthetic GPS, and disciplined identity separation.
- **Identity twin:** an authorized second gateway with the target-correlated profile and synthetic behavior.

KPIs

Measure	Decision question
True-target attribution	How often was the real target correctly selected?
Attribution Index	How much did metadata improve confidence over the prior assessment?
Identity-twin capture	How often did attribution confidence move from the real target to the twin?
Defense/twin detection	Could the observer or MNO recognize Klök, infer its controls, or detect the twin?
Calibration and cost	Was confidence accurate, and how much additional collection, fusion, computation, and checking were required?
Utility and safety	Did protection remain within mission, network, latency, bandwidth, authentication, and safety limits?

Neutralizing the AI-Enabled Cost Advantage

AI Advantage Neutralization Ratio = protection-imposed cost multiplier / original AI cost advantage

For example, if AI reduced attribution cost by a factor of five, Klök must impose at least a fivefold cost multiplier to neutralize that advantage. A ratio of 1 or greater meets that threshold.

Executive Decision Standard

Proceed when Klök materially reduces true-target attribution and confidence calibration, shifts collection toward the identity twin, and raises adversary cost without unacceptable mission, network, authentication, or safety effects.

Selected Scientific and Standards Basis

[1] NIST SP 800-207, Zero Trust Architecture (2020). [Source](#)

[2] 3GPP TS 23.003, Numbering, addressing and identification. [Source](#)

[3] 3GPP TS 33.501, Security architecture and procedures for 5G. [Source](#)

[4] de Montjoye et al., mobility traces and uniqueness (2013). [Source](#)

[5] NIST SP 800-160 Vol. 2 Rev. 1, cyber-resilient systems (2021). [Source](#)

users, infer identities, and build pattern-of-life profiles. This capability, called observability at scale, enables adversaries to identify high-value targets, predict behavior, and determine when and where to strike; digitally or physically.

This threat isn't theoretical. If an AI system can observe that a certain device ID appears regularly near defense facilities, follows a specific daily movement pattern, or logs into a sensitive cloud portal every Monday from a hotel in a foreign capital, it doesn't need to decrypt anything. It can simply correlate that metadata and escalate the target for exploitation.

Once a target is classified, adversaries can deploy tailored exploits: phishing that mimics known Wi-Fi networks, zero-click malware using baseband vulnerabilities tied to a known device model, or even physical surveillance and interdiction based on location history. Metadata is not just an observation layer, it's a kill chain initializer.

Salt Typhoon: A Case Study of a Metadata-Driven Campaign

Salt Typhoon (formerly TAG-74) is a Chinese state-aligned threat actor that offers a real-world case study. This group didn't rely on malware-laced phishing campaigns or zero-day attacks as its primary tactic. Instead, it focused on compromising telecom providers, cloud service resellers, and network infrastructure operators, and the Treasury, embedding itself into the very systems that manage and process metadata.

Their objective wasn't necessarily to steal documents, it was to observe. They collected call records, IP login histories, subscriber profiles, their locations and behavioral patterns of users inside defense contractors, government agencies, and infrastructure providers. Microsoft confirmed that Salt Typhoon's operations enabled long-term surveillance, pattern recognition, and ultimately target selection at scale.

This is the model for modern AI-powered adversaries: don't bother to break encryption; bypass it. Don't fight the firewall; watch from behind it. In Salt Typhoon's case, metadata allowed them to track movement, prioritize targets, and escalate access gradually, all while flying under the radar of conventional security tools.

The Cybersecurity Paradox

This gives rise to a core paradox in cybersecurity:

The very metadata we use to *secure* systems is now being used to *target* them.

The entire Cybersecurity industry rely (indeed, is built on) on metadata for visibility, access control, and threat detection. TLS fingerprints help detect spoofed clients. MAC addresses enforce NAC. Behavioral telemetry underpins anomaly detection.

But every one of these identifiers, when observed externally or leaked upstream, becomes a beacon. The persistence and consistency required for enforcement are exactly what adversaries exploit for classification and correlation