Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Exemption Without Exposure: Formally Verified Conditional Privacy-Preserving Authentication with Secure-Ranging Binding for Authorised Armed Carriers in AI Weapon Surveillance

Abas Aliu

abasaliu@gmail.com

Edo State University Uzairue

Braimoh Ikharo

Edo State University Uzairue

Ayodeji Samuel Makinde

Edo State University Uzairue

Okoi Michael Obeten

Auchi Polytechnic

Research Article

Keywords:

Posted Date: November 3rd, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11376107/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Exemption Without Exposure: Formally Verified Conditional Privacy-Preserving Authentication with Secure-Ranging Binding for Authorised Armed Carriers in AI Weapon Surveillance

Abas Aliu[1*], A. Braimoh Ikharo[1], Ayodeji Samuel Makinde[1], Okoi Michael Obeten[2]

[1*]Department of Computer Science, Edo State University Uzairue, Uzairue, Edo State, Nigeria.
[2]Department of Artificial Intelligence, Auchi Polytechnic, Auchi, Edo State, Nigeria.

*Corresponding author(s). E-mail(s): abasaliu@gmail.com;
Contributing authors: michealobeten@auchipoly.edu.ng;

Abstract

Automated weapon detectors treat every visible firearm as a threat, yet plain-clothes soldiers, police officers and licensed guards lawfully carry weapons on the roads where kidnappers operate. Exempting them keeps false alarms manageable, but an exemption is also an attack surface: a stolen, relayed or wrongly attributed credential can silence an alert about a genuine threat. This paper presents conditional privacy-preserving authentication (CPPA) for authorised armed carriers in artificial intelligence surveillance. Carriers present short-lived pseudonymous role credentials through a challenge-response exchange, so the verifier learns a role but not an identity, while an oversight authority can resolve a pseudonym only under lawful order. Ultra-wideband secure ranging binds the credential to the vehicle the camera flagged, and a three-outcome policy ensures that no weapon detection is silently discarded. The protocol is verified in the Tamarin prover: unforgeability, injective agreement, identity secrecy and unlinkability hold, and weakening any of four design elements yields an attack. A reference implementation verifies a 225-byte presentation in a median of 253 µs on one 2.1 GHz virtual processor. Monte Carlo simulation shows a binding error of 6.7×10^{-4} for side-by-side vehicles with 0.1 m ranging and 0.5 m camera localisation, against 0.08 with 1 m ranging and localisation. In a representative scenario, binding lowers the probability that a threat is silently exempted from 0.21 to the credential compromise rate of 0.01. A closed-form analysis of multi-agent threat verification shows that half of its accuracy gain is lost once agent errors correlate at about 0.4.

Keywords: Conditional privacy-preserving authentication, Formal verification, Tamarin prover, Secure ranging, Pseudonyms, Weapon detection

1 Introduction

Armed banditry and kidnapping for ransom have become severe internal security threats in Nigeria, with consequences that extend from local economies to regional stability [1, 2]. Computer vision offers a way to detect weapons on the roads where such attacks occur, and detector-driven

alarms have been studied for more than a decade [3]. Real-time weapon detection in surveillance footage nevertheless remains an open problem [4], recent reviews report inconsistent evaluation and limited real-world validation [5], and 2026 studies continue to benchmark new detector generations for this task [6].

A practical difficulty is rarely addressed in this literature. On Nigerian roads, soldiers, police officers, paramilitary units and licensed private guards routinely carry firearms, often out of uniform and in unmarked vehicles. A detector that alerts on every weapon will therefore generate a stream of alarms about lawful carriers, and responders who learn to ignore such alarms will also ignore genuine ones. Some form of authorised-carrier exemption is needed. Any exemption, however, is a mechanism for silencing alerts, and an adversary who can obtain, relay or borrow an exemption gains exactly what a kidnapping gang needs.

Conditional privacy-preserving authentication is well developed for vehicular networks, where it authenticates messages while hiding vehicle identities from other participants and preserving traceability for an authority [7]. Its security is subtle: cryptanalysis has exposed impersonation and repudiation attacks on published schemes [8], and forgery attacks on aggregate-signature schemes were reported as recently as 2026 [9]. Pseudonym certificates are the most widely studied practical instantiation [10]. In all of this work, authentication answers whether a message is genuine. It does not answer whether a weapon seen by a camera belongs to an authorised carrier, which is a physical, cross-modal question with direct safety consequences.

This paper, which extends the authors' earlier work on weapon detection for Nigerian surveillance [11], makes six contributions.

1. A threat model for authorised-carrier exemption with six adversary classes and eight security goals, including the safety goal that no exemption path may silently discard a weapon detection.
2. A CPPA design built from short-lived pseudonymous role credentials, challenge-response presentation, revocation, escrowed conditional traceability and salted audit commitments, with a three-outcome decision policy (escalate, downgrade, review).
3. A machine-checked analysis of authentication, identity secrecy and unlinkability in the Tamarin prover, with an ablation showing each protocol element to be necessary; the analysis also exposed and corrected a linkability flaw in the original credential design.
4. A credential-to-vehicle binding mechanism based on secure ranging, evaluated by Monte Carlo simulation across ranging accuracy, camera localisation accuracy and traffic geometry.
5. A reference implementation with a functional attack test suite and measured computational, communication and storage costs.
6. An exemption policy analysis that quantifies silent-suppression risk, and a closed-form analysis of the collaborative multi-agent threat verification (CMTV) fusion rule that quantifies its dependence on error correlation.

Section 2 reviews related work, Section 3 defines the threat model and Section 4 the design. Section 5 reports the formal verification, Section 6 the security and privacy analysis and Section 7 the evaluation. Section 8 discusses limitations and deployment, and Section 9 concludes.

2 Related work

2.1 Conditional privacy in vehicular networks

CPPA schemes for vehicular ad hoc networks either preload vehicles with large pools of pseudonym certificates under a public-key infrastructure or use group signatures, authenticating messages anonymously while allowing a trusted authority to trace misbehaving vehicles [7]. Kabil et al. showed that one published scheme lets an attacker who learns a user's real identity value impersonate that user, and that the same scheme admits repudiation [8]. Shim and Kwon showed in 2026 that two certificateless aggregate-signature CPPA schemes can be forged by an adversary able to replace a vehicle's public key [9]. These results motivate the machine-checked analysis in Section 5. Pseudonym schemes are surveyed in [10], and changing pseudonyms does not by itself prevent linking, since an observer can correlate

successive pseudonyms [12]. The security credential management system developed for vehicle-to-everything communication issues pseudonym certificates and divides their generation among several organisations to protect privacy against insiders [13]. A 2026 scheme authenticates emergency vehicles with zero-knowledge proofs and a blockchain, avoiding certificates, pseudonym pools and online interaction with a trusted authority [14]; it is the closest work in spirit, but it authenticates a vehicle's messages rather than exempting a camera-detected event. Blockchain-based trust and authentication for vehicular networks are surveyed in [15].

2.2 Zero-knowledge proofs and anonymous credentials

Zero-knowledge proofs allow a prover to convince a verifier that a statement is true without revealing anything beyond its truth [16]. Anonymous credential systems build on this idea: Camenisch and Lysyanskaya proposed a system in which credential shows cannot be linked, anonymity can be revoked for particular transactions, and sharing one credential with another person gives that person all of the owner's credentials, which discourages lending [17]. Such systems offer stronger unlinkability than pseudonym certificates at a higher computational cost.

2.3 Relay attacks and secure ranging

Proving possession of a credential does not prove where the holder is. Distance-bounding protocols, introduced by Brands and Chaum, bound the physical distance to a prover by timing rapid challenge-response rounds [18]. Their importance is practical: Francillon et al. relayed the radio exchange between passive keyless entry systems and their keys to open and start ten car models from eight manufacturers, independently of the cryptography in use [19]. Ultra-wideband (UWB) impulse radio makes secure ranging practical: UWB with pulse reordering resists physical-layer distance-reduction attacks and achieved ranges of up to 93 m with 10 cm precision in line of sight [20].

2.4 Verification of detections

False-alarm control in weapon detection has relied on temporal confirmation, for example an alarm that fires only after several consecutive positive frames [3]. Tracking associates detections across frames using overlap and motion models at very low cost [21], and contrastive vision-language models such as CLIP provide semantic embeddings that can check whether an image region is consistent with a textual class description [22]. Combining several such checks is an ensemble problem: ensemble accuracy depends on how the members' errors are related as well as on their individual accuracies, although measuring this diversity in practice is difficult [23]. Agreement between methods should likewise be measured rather than assumed: in cost-aware industrial fault classification, two widely used explanation methods applied to the same model showed only moderate agreement, which varied across instances and random seeds [24].

2.5 Privacy of audit records

Hashing personal data before publishing it is a weak form of anonymisation, because small or structured input spaces can be enumerated [25]. The European Data Protection Board (EDPB) advises keeping personal data off public ledgers and discusses commitments as a mitigation [26], and the Nigeria Data Protection Act 2023 sets equivalent principles of minimisation and security [27]. Whether blockchain-anchored records are admissible as evidence is a separate legal question that recent work examines for criminal prosecutions [28].

2.6 Formal protocol analysis

Symbolic protocol analysis assumes an adversary who controls the network and can intercept, replay and compose messages but cannot break cryptographic primitives [29]. The Tamarin prover automates such analysis for an unbounded number of sessions and either constructs a proof or produces an attack trace [30].

Table 1 summarises the gap.

Table 1 Positioning of this work

Strand	What it authenticates or verifies	What it does not address
Vehicular CPPA [7, 9, 13, 14]	Messages between vehicles and infrastructure	Exemption of a camera-detected physical event
Anonymous credentials [16, 17]	Possession of attributes without identity	Proximity of the holder to a detected object
Secure ranging [18–20]	Physical distance to a prover	Which detected vehicle the prover is in
Detection verification [3, 21, 22]	Consistency of a detection over time and meaning	Lawful authorisation of the carrier
This work	Role-based exemption, bound to the detected vehicle, with no silent suppression	Field deployment and observational unlinkability proof

3 Problem statement and threat model

3.1 Entities and assumptions

Four entities take part. A credential authority (CA), operated jointly by the security agencies, enrols officers and their assigned vehicles. Each officer carries a device, such as a hardened smartphone or token with a UWB radio, holding a batch of short-lived pseudonymous credentials. A verifier, a roadside unit with at least three UWB anchors and the camera-based detector, issues challenges and makes decisions. An oversight authority holds escrow records that link pseudonyms to identities and operates a ledger for audit commitments. We assume the CA signing key is protected, that devices store keys in secure hardware, and that the verifier's clock is loosely synchronised.

3.2 Adversaries and goals

Table 2 lists the adversary classes and the goals they threaten.

The eight goals are: G1 only CA-issued credentials are accepted, with the role they were issued for; G2 an accepted presentation was produced by the genuine device for this verifier and this challenge, and each challenge is accepted at most once; G3 an accepted credential belongs to the detected vehicle; G4 presentations by the same officer cannot be linked across validity windows; G5 an officer's identity is learned only through lawful resolution by the oversight authority; G6 a lost or compromised credential can be invalidated quickly; G7 no outcome silently discards a weapon detection; and G8 every decision leaves a verifiable record.

4 Design

4.1 Overview

Fig. 1 shows the decision flow. Every weapon detection triggers a presentation attempt. A valid credential leads to a binding check; no credential, or an invalid or revoked one, leads to CMTV. Each path ends in one of three outcomes: ESCALATE, which alerts responders; DOWNGRADE, which routes the event to a sealed oversight channel instead of the public responder channel; or REVIEW, which places the event in an operator queue with its evidence retained.

4.2 Role credentials and the role credential verifier

At enrolment, the CA issues each officer a batch of K credentials, each valid for a short window. A credential contains a random 16-byte pseudonym identifier pid, a fresh Ed25519 public key pk_i, a role class, a vehicle commitment vc, a validity window $[t_s, t_e)$ and the CA signature:

$$\sigma_{\mathrm{CA}} = \mathrm{Sign}_{sk_{\mathrm{CA}}}\big(pid \,\|\, pk_i \,\|\, role \,\|\, vc \,\|\, t_s \,\|\, t_e\big) \tag{1}$$

$$vc_i = H(\mathrm{VID} \,\|\, s_i), \quad s_i \in \{0,1\}^{128} \text{ fresh per credential} \tag{2}$$

Table 2 Adversary classes and security goals

Adversary	Capability	Goals threatened
A1 Outsider forger	Fabricates credentials or responses	G1 unforgeability
A2 Replayer	Records and replays an earlier presentation	G2 freshness and agreement
A3 Relayer	Forwards an officer's device exchange from elsewhere	G3 binding to the detected vehicle
A4 Credential thief or coercer	Obtains an officer's device or forces its use	G3, G6 revocability, G7 no silent suppression
A5 Malicious insider	Abuses enrolment or escrow	G5 conditional traceability, G8 accountability
A6 Curious observer	Reads ledger records or colludes across verifiers	G4 unlinkability, privacy of carriers

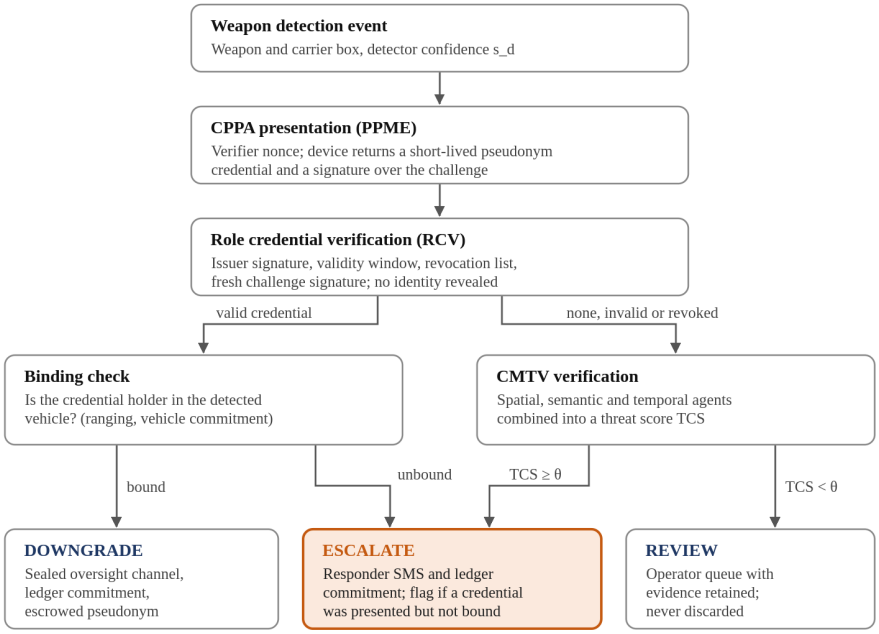


Fig. 1 CPPA and CMTV decision flow. Every branch ends in escalation, a sealed downgrade or operator review

where VID identifies the assigned vehicle and s_i is a 16-byte salt unique to each credential and held by the officer's device. A fresh salt per credential is essential: with one salt per batch, every pseudonym in the batch would carry the same vc, letting any verifier link all of an officer's presentations (Section 5.3). No name, rank or service number appears in the credential. The CA hands each pid with its owner's identity to the oversight authority as a sealed escrow record and deletes its own copy. Table 3 gives the encoded format.

The role credential verifier (RCV) accepts a credential only if the CA signature verifies, the current time lies in the validity window, and pid is not on the revocation list.

4.3 Presentation protocol and the privacy-preserving match engine

Fig. 2 shows the exchange. On a detection, the verifier broadcasts a challenge containing a fresh 16-byte nonce n, its identifier V_{id} and the time

Table 3 Credential fields and sizes

Field	Size (bytes)	Content
pid	16	Random pseudonym identifier
pk_i	32	Ed25519 public key for this pseudonym
role	1	Military, police, paramilitary or licensed private
vc	32	Commitment to the assigned vehicle
t_s, t_e	16	Validity window
σ_{CA}	64	Issuer signature
Total	161	

t. Each device in range that holds a currently valid credential responds with that credential and a signature over the challenge:

$$\rho = \mathrm{Sign}_{sk_i}\big(n \,\|\, V_{id} \,\|\, t \,\|\, role\big) \tag{3}$$

The privacy-preserving match engine (PPME) checks that the challenge is fresh and that the response verifies under pk_i. Because each credential is used for a short window and then discarded, a verifier sees a different pid, key and vehicle commitment in each window, and only the oversight authority can link them. This pseudonym-certificate realisation is deliberately simple; an anonymous-credential realisation [17] would give stronger unlinkability within a window at a higher computational cost.

4.4 Binding by secure ranging

A valid response proves that a credential holder is within radio range, not that the holder is in the vehicle the camera flagged. Binding closes this gap. Each of $J \geq 3$ roadside anchors at known positions $\mathbf{a}_j$ performs secure UWB ranging with the responding device, in the same session as the presentation, giving distance estimates $\hat{d}_j$. Because secure ranging resists distance reduction [20], a relay can only make a device appear farther away, never closer. The verifier estimates the device position by nonlinear least squares:

$$\hat{\mathbf{p}} = \arg\min_{\mathbf{p}} \sum_{j=1}^{J} \big(\|\mathbf{p} - \mathbf{a}_j\| - \hat{d}_j\big)^2 \tag{4}$$

The camera supplies ground-plane estimates $\hat{\mathbf{c}}$ of the centres of all vehicles in view. The credential is bound to the detected vehicle A only if the device lies within a gate γ of A's centre and closer to A than to any other vehicle:

$$\begin{aligned} \mathrm{bind}(A) \Leftrightarrow \|\hat{\mathbf{p}} - \hat{\mathbf{c}}_A\| < \gamma \\ \wedge \ \|\hat{\mathbf{p}} - \hat{\mathbf{c}}_A\| < \min_{B \neq A} \|\hat{\mathbf{p}} - \hat{\mathbf{c}}_B\| \end{aligned} \tag{5}$$

A complementary check compares the vehicle commitment vc with a commitment computed from the plate read by the camera. This requires the device to reveal the salt s_i of the presented credential only, and binds the credential to the registered vehicle rather than to a position, at the cost of revealing to the verifier that the plate belongs to an authorised vehicle; it is therefore optional.

4.5 Confidence adjustment module

The confidence adjustment module (CAM) applies the rules in Table 4. The governing principle is fail-safe behaviour in the dependability sense [31]: when information is missing or uncertain, the system moves towards the safer outcome, which here is escalation or review rather than exemption.

DOWNGRADE does not mean that nothing happens. The event is sent to the oversight authority with the pseudonym, the officer's command can be notified through the escrow, and the decision is committed to the ledger. A coerced officer can therefore be identified after the fact, and a duress credential that appears valid to observers but triggers covert escalation is a natural extension.

4.6 Collaborative multi-agent threat verification

CMTV examines detections that have no valid credential. A spatial agent checks that the detection persists across consecutive frames with sufficient overlap, as in simple online tracking [21]. A semantic agent checks that the image region is consistent with the predicted class using a contrastive vision-language embedding [22]. A temporal agent requires the detection to recur within a short frame window, generalising the consecutive-frame alarm of [3]. Each agent returns a score s_k in $[0, 1]$,

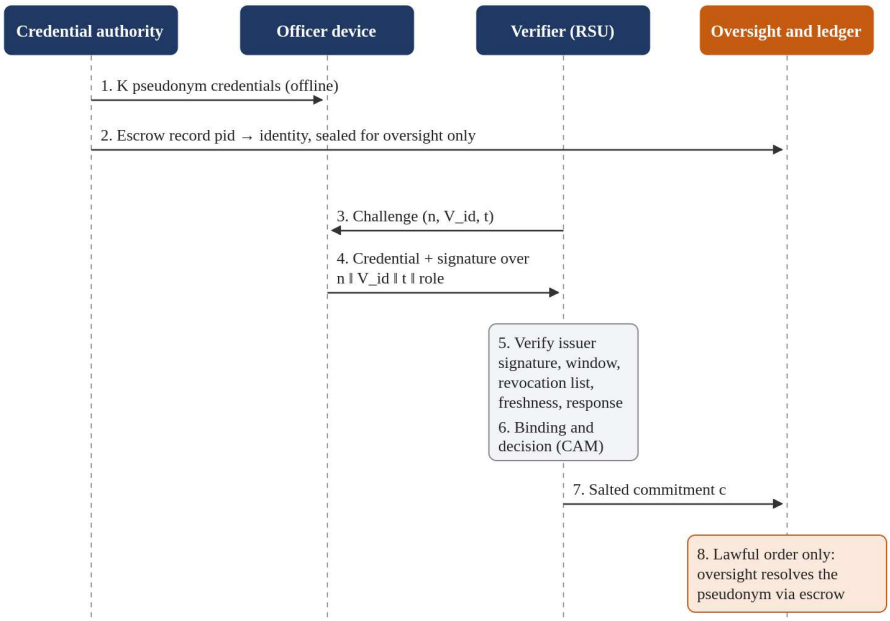


Fig. 2 Enrolment, presentation and audit messages. The verifier learns a role class and a pseudonym, never an identity

Table 4 Decision rules of the confidence adjustment module

Credential status	Binding	CMTV score	Outcome
Valid	Bound to the detected vehicle	Not used	DOWNGRADE to sealed oversight channel
Valid	Not bound	Not used	ESCALATE, flagged as unbound credential
Revoked or expired	Not checked	Not used	ESCALATE, flagged as possible stolen credential
None or invalid	Not applicable	TCS $\geq \theta$	ESCALATE
None or invalid	Not applicable	TCS $< \theta$	REVIEW, evidence retained

combined into a threat confidence score (TCS) as

$$\mathrm{TCS} = \sum_{k=1}^{3} w_k s_k, \qquad \sum_{k=1}^{3} w_k = 1 \tag{6}$$

Low scores lead to REVIEW, not to discard, so CMTV can lower the priority of a false positive but cannot erase a true one.

4.7 Audit record

Each decision d is anchored on the ledger as a salted commitment

$$\kappa = H(r \,\|\, d \,\|\, pid \,\|\, n \,\|\, t), \qquad r \in \{0,1\}^{256} \tag{7}$$

with r stored off-chain by the oversight authority. Plain hashing of such a record would be open to dictionary recovery because its fields are structured [25]; the random r prevents this, consistent with the EDPB guidance on commitments [26]. The ledger thus records that a decision was made at a given time without revealing its content, which supports chain-of-custody practice for digital evidence [32].

5 Formal verification

5.1 Model

The enrolment and presentation protocol of Fig. 2 was modelled in the Tamarin prover, version 1.12.0 with Maude 3.5.1 [30], under a Dolev-Yao adversary who controls all messages [29]. The model includes CA setup, enrolment with a secret officer identity held only in the escrow, device responses to arbitrary challenges, verifier challenge and acceptance, and three adversary capabilities beyond the network: compromise of the CA key, compromise of a device key (theft or coercion, A4), and lawful resolution of a pseudonym by oversight, whose output is made public to be conservative. Signatures follow Tamarin's built-in equational theory. Goals G1, G2 and G5 were stated as trace properties, with an executability lemma to rule out vacuous proofs. Unlinkability (G4) was stated as observational equivalence between two systems: in one, every presentation comes from the same officer and vehicle; in the other, each presentation comes from a different officer and vehicle. If the adversary cannot distinguish the two, presentations cannot be linked. All models are available from the corresponding author on reasonable request.

5.2 Results

All properties were proved for an unbounded number of sessions (Table 5): the trace properties in 1.59 s and the equivalence in 10.42 s. Each trace property holds except where the relevant key has been compromised, which is the strongest statement possible for these goals.

5.3 Ablation

To show that each design element is necessary, four variants were analysed (Table 6). Removing the nonce from the signed response lets the adversary replay an old response; removing the verifier identifier lets the adversary obtain a response intended for one verifier and present it to another; placing the officer identity in the credential instead of the escrow exposes it to every verifier; and using one salt per batch for the vehicle commitment, as in the original design, makes an officer's pseudonyms linkable through their identical commitments. In each case Tamarin produced a concrete attack trace. The last finding led to the per-credential salt in (2).

5.4 Scope of the proofs

The symbolic model does not capture physical distance, so relay attacks on proximity (A3) are outside it and are addressed by secure ranging (Section 4.4). Nor does it capture side channels such as the timing and location of presentations, which can still link pseudonyms [12]. Cryptographic primitives are assumed perfect, so the proofs complement, and do not replace, the computational security of Ed25519 [33].

6 Security and privacy analysis

Table 7 maps each goal to its mechanism, the form of evidence and the residual risk.

Three residual risks deserve emphasis. First, binding is only as good as the camera's localisation and the geometry of traffic; Section 7.4 quantifies both. Second, coercion (A4) cannot be prevented cryptographically, which is why DOWNGRADE keeps the event visible to oversight and why duress credentials are worth adding. Third, a lent credential is a valid credential; pseudonym certificates do not deter lending, whereas the all-or-nothing sharing property of anonymous credentials does [17].

7 Evaluation

7.1 Setup

The reference implementation is written in Python 3.12.3 with the cryptography library 46.0.6, using Ed25519 for all signatures and SHA3-256 as the commitment hash. All timing measurements were taken on a single virtual central processing unit (CPU) of an Intel Xeon processor at 2.10 GHz in five independent runs of 2,000 repetitions per operation, and the verifier's revocation list was populated with 10,000 entries. Because the virtual machine is shared, we report the median across runs together with the range of the run medians. The binding simulation and the policy and fusion analyses use NumPy with fixed random seeds. All code, the Tamarin models and the raw results are available from the corresponding author on reasonable request.

Table 5 Tamarin results for the presentation protocol

Lemma	Property	Result	Steps
executable	An honest presentation can be accepted	Verified (exists-trace)	10
G1 unforgeability	Every accepted credential was issued by the CA with that role, unless the CA key is compromised	Verified	7
G2 agreement	Acceptance implies the genuine device signed this verifier's nonce and role, unless the device or CA key is compromised	Verified	10
G2 injectivity	Each challenge is accepted at most once	Verified	10
G4 unlinkability	Presentations by one officer are indistinguishable from presentations by different officers	Verified (observational equivalence)	695
G5 identity secrecy	An officer's identity becomes known only through lawful resolution	Verified	4

Table 6 Protocol ablation in Tamarin

Variant	Lemma falsified	Attack found
Response signs role only (no nonce, no verifier)	G2 agreement	Replay of an earlier response (9 steps)
Response signs nonce and role, not the verifier identifier	G2 agreement	Cross-verifier misuse of a response (10 steps)
Officer identity placed in the credential	G5 identity secrecy	Identity read from any presentation (4 steps)
One salt per batch for the vehicle commitment	G4 unlinkability	Linking of an officer's pseudonyms (15 steps)

7.2 Functional tests

Table 8 lists the test cases. The valid presentation was accepted, and each of the six attack or fault cases was rejected for the expected reason. Consecutive pseudonyms from the same batch shared no field other than the role class.

7.3 Computational, communication and storage cost

Table 9 reports the measured costs. A full verifier check, including both signature verifications and a revocation lookup, took a median of 253 µs, which is negligible beside detector inference at tens to hundreds of milliseconds per frame [4]. Issuing a week of hourly credentials for one officer took a median of 16.6 ms. A presentation is 225 bytes, small enough for a single short-range radio exchange. A week of hourly credentials with their private keys occupies about 32 kB on the device, and a 10,000-entry revocation list occupies 160 kB at the verifier. Mobile devices will be slower than the server CPU used here, but the device's only per-session cryptographic operation is one signature, which took 38 µs.

7.4 Binding evaluation

The binding mechanism of Section 4.4 was evaluated by Monte Carlo simulation with 200,000 trials per configuration. Three anchors were placed at $(-15, -6)$, $(15, -6)$ and $(0, 10)$ m relative to the threat vehicle. Vehicles measure 4.5 m by 1.8 m, and the officer's device was placed uniformly in the central cabin region of its vehicle. Ranging errors were Gaussian with standard deviation σ_r of 0.1 m, representing UWB secure ranging [20], or 1 m and 3 m, representing coarse ranging; camera centre estimates had isotropic Gaussian error σ_c of 0.5 m or 1 m; and the gate γ was 3 m. Two

Table 7 Goals, mechanisms, evidence and residual risks

Goal	Mechanism	Evidence	Residual risk
G1 Unforgeability	Ed25519 issuer and response signatures [33]	Tamarin proof; functional tests	CA key compromise; mitigated by splitting issuance across agencies [13]
G2 Freshness and agreement	Nonce and verifier identifier in the signed response	Tamarin proof; ablation	Clock skew window of a few seconds
G3 Binding	Secure UWB ranging [20] with position-based association	Monte Carlo simulation (Section 7.4)	Side-by-side vehicles; multipath not modelled
G4 Unlinkability	Short-lived pseudonyms with fresh keys and per-credential salts	Tamarin equivalence proof; ablation	Linking through timing and location [12]; role class is visible
G5 Conditional traceability	Escrow held by oversight only	Tamarin proof; ablation	Insider abuse at oversight; needs logging and dual control
G6 Revocability	Revocation list of pseudonym identifiers	Functional tests	Lists must reach verifiers quickly; short validity limits exposure
G7 No silent suppression	Three-outcome CAM; REVIEW instead of discard	Policy analysis (Section 7.6)	Operator workload; coerced officers present valid, bound credentials
G8 Accountability	Salted commitment of every decision	Construction	Ledger persistence; off-chain salts must be preserved

Table 8 Functional test outcomes

Case	Adversary	Expected	Observed
Valid presentation	None	Accept	Accepted
Expired credential	A4	Reject	Rejected: outside validity
Revoked credential	A4	Reject	Rejected: revoked
Response signed with another key	A1	Reject	Rejected: bad response
Replayed challenge (30 s old)	A2	Reject	Rejected: stale challenge
Role field altered after issue	A1	Reject	Rejected: bad issuer signature
Credential signed by a rogue issuer	A1, A5	Reject	Rejected: bad issuer signature

quantities were measured: the binding error b, the probability that a device in a neighbouring officer vehicle is bound to the threat vehicle, and the miss rate m, the probability that a device in the detected vehicle itself is not bound to it, which causes a false escalation of an authorised carrier.

Fig. 3 shows b for a following vehicle at increasing gaps, and Table 10 adds the hardest case, an officer vehicle level with the threat vehicle in the adjacent lane. With UWB ranging and 0.5 m camera localisation, b was 6×10^{-5} at a 0.5 m gap, below the measurement floor of 5×10^{-6} from 2 m onwards, and 6.7×10^{-4} side by side, with a miss rate of at most 0.09%. Camera localisation matters as much as ranging: with the same UWB ranging but 1 m camera error, the side-by-side binding error rose to 3.9% and the miss rate to about 4%. Coarse 3 m ranging is unusable, missing more than half of authorised carriers.

7.5 CMTV fusion analysis

Equation (6) fuses agent scores by their weighted mean. To quantify what this buys, model each agent's score as $s_k = y\delta + e_k$, where $y = \pm 1$ indicates threat or non-threat and the errors e_k

Table 9 Measured costs (five runs of 2,000 repetitions each)

Operation	Median (µs)	Range of run medians (µs)	95th percentile (µs)
Issue one credential (CA)	92.2	88.5 to 150.0	139.6
Device response (one signature)	38.3	38.2 to 64.5	59.5
Full verifier check (two verifications and revocation lookup)	253.0	248.5 to 254.3	343.8
Salted decision commitment	5.8	5.8 to 5.8	9.3

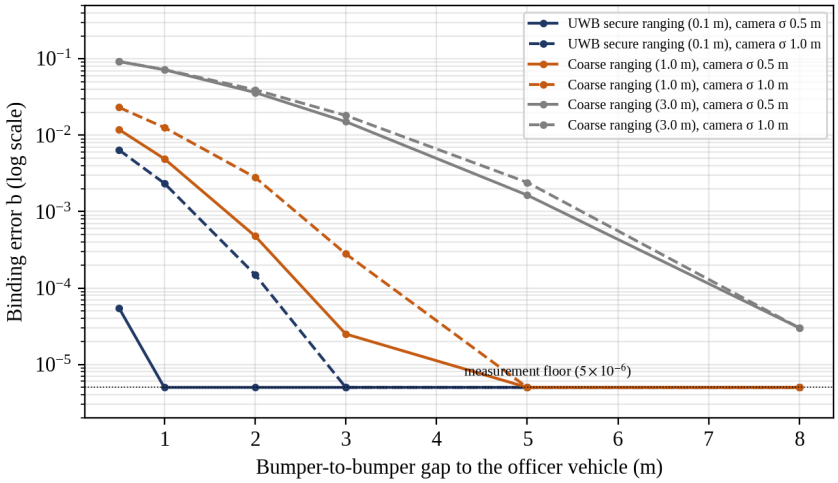


Fig. 3 Binding error for a following officer vehicle as the gap increases, by ranging and camera localisation accuracy. Values below the measurement floor are drawn at the floor

are Gaussian with equal variance and pairwise correlation ρ. A single agent thresholded at the midpoint then has accuracy $p = \Phi(\delta/\sigma)$, and the variance of the mean of K equally weighted errors is $\sigma^2(1 + (K - 1)\rho)/K$, so the fused score has accuracy

$$A_{\mathrm{TCS}} = \Phi\left(\Phi^{-1}(p)\sqrt{\frac{K}{1 + (K-1)\rho}}\right) \qquad (8)$$

A Monte Carlo check with 400,000 trials at twelve combinations of p and ρ agreed with (8) to within 0.0011. Fig. 4 plots (8) for three agents. With independent errors, three agents of accuracy 0.80 reach 0.928, but at $\rho = 0.5$ they reach only 0.849, and half of the gain is lost by $\rho \approx 0.39$; across single-agent accuracies of 0.75 to 0.85 the half-gain point lies between $\rho = 0.36$ and 0.43. Mean-score fusion never falls below a single agent, but its benefit vanishes as ρ approaches 1. Because the three agents inspect the same frames and fail on the same hard cases, such as small, occluded or blurred weapons, their errors are likely to be positively correlated. Any claimed CMTV improvement must therefore be supported by a measured error correlation on real footage, not only by per-agent accuracies [23].

7.6 Exemption policy analysis

Let c be the probability that a threat vehicle carries a valid compromised credential, q the probability that some legitimate credential holder is within radio range of the verifier, and b the binding error from Section 7.4. Three policies are compared: P1 suppresses an alert whenever any valid credential responds; P2 suppresses only when the credential is bound to the detected vehicle; and P3, the policy of Table 4, never suppresses silently. The probabilities that a threat is silently

Table 10 Binding error b and miss rate m (200,000 trials per value)

Ranging σ_r	Camera σ_c	b, 0.5 m gap	b, 2 m gap	b, side by side	m, same lane
0.1 m (UWB)	0.5 m	6×10^{-5}	$< 5 \times 10^{-6}$	6.7×10^{-4}	$\leq 0.02\%$
0.1 m (UWB)	1.0 m	0.64%	0.015%	3.9%	3.8 to 4.2%
1.0 m	0.5 m	1.2%	0.048%	4.0%	3.2 to 3.8%
1.0 m	1.0 m	2.3%	0.28%	8.1%	11.4 to 12.7%
3.0 m	0.5 m	9.2%	3.6%	16.9%	51.6 to 52.7%
3.0 m	1.0 m	9.2%	3.9%	16.9%	55.0 to 56.9%

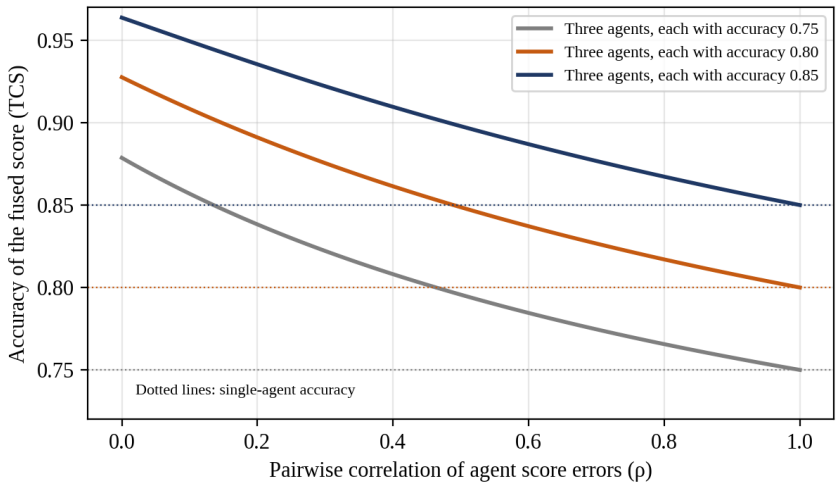


Fig. 4 Accuracy of the fused CMTV score for three agents as a function of the correlation of their errors, from Eq. (8). Dotted lines show single-agent accuracy

suppressed are

$$P_1 = 1 - (1-c)(1-q), \qquad P_2 = c + (1-c)\,q\,b, \qquad P_3 = 0 \tag{9}$$

Fig. 5 plots these for $c = 0.01$, using the side-by-side binding errors from Table 10 as the worst case. At $q = 0.2$, P1 silences 0.208 of threats, almost entirely because a legitimate officer happens to be nearby. Binding with coarse ranging ($b = 0.081$) reduces this to 0.026, and binding with UWB ranging ($b = 6.7 \times 10^{-4}$) to 0.0101, which is essentially the compromise rate c itself. Good binding thus removes the coincidental-proximity risk and leaves credential compromise as the residual, which is exactly the risk that P3 addresses by routing every exemption to oversight instead of silencing it.

8 Discussion

8.1 What the results show

The cryptographic core of CPPA is cheap and, in the symbolic model, correct: its cost is more than two orders of magnitude below detector inference, and its essential properties, including unlinkability, are machine-checked with every design element shown to be necessary. The decisive factor for safe exemption is binding, and the simulation shows that binding is achievable when secure UWB ranging is combined with camera localisation of about half a metre. With weaker localisation, binding errors of several per cent arise in dense traffic, and the three-outcome policy becomes the main safeguard. CMTV can reduce operator workload, but its benefit must be established with real detections and an explicit measure of error correlation.

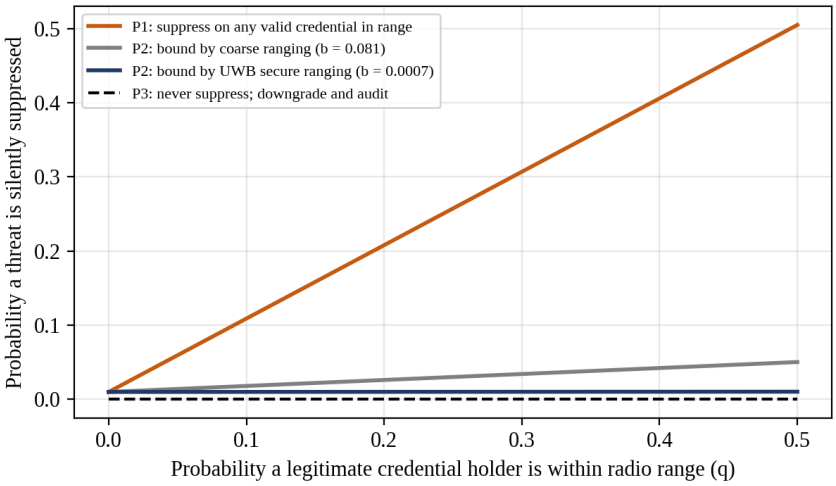


Fig. 5 Probability that a threat is silently suppressed under three policies ($c = 0.01$; side-by-side binding errors from Table 10)

8.2 Deployment considerations

Deployment requires institutional arrangements that technology cannot supply: a joint issuing body for the security agencies, an oversight authority with legal power and audit obligations over the escrow, rules on when officers must carry and present credentials, and data protection impact assessments under the Nigeria Data Protection Act [27]. Roadside sites need at least three UWB anchors with surveyed positions and a camera calibrated for ground-plane localisation. Officers must also accept that downgraded events are logged, which is the price of not being stopped at gunpoint by a responder who believes they are kidnappers. Whether ledger commitments can support prosecutions or disciplinary proceedings depends on procedural law [28].

8.3 Limitations

Timing was measured in software on a server CPU; device-side timing on target hardware, radio latency and battery cost were not measured. The binding evaluation is a simulation with Gaussian errors in line of sight; multipath, occlusion by large vehicles and anchor failures were not modelled and would raise binding errors. The CMTV agents are not implemented, the fusion analysis assumes Gaussian score errors, and the policy parameters c and q are illustrative. The formal model does not capture timing or location side channels. No field trial with officers or vehicles has been carried out.

9 Conclusion

Exempting authorised armed carriers from AI weapon alerts is necessary in settings such as Nigeria's transport corridors, but every exemption is a potential silencer. This paper designed a CPPA layer from short-lived pseudonymous role credentials, bound each credential to the camera-detected vehicle by secure ranging, and adopted a three-outcome decision policy under which no weapon detection is silently discarded. The presentation protocol was verified in Tamarin, including unlinkability, with an ablation showing that each of its elements is necessary and exposing a linkability flaw in the original credential design; a reference implementation showed its cost to be negligible; and simulation showed that UWB binding with half-metre camera localisation reduces the risk of silent exemption to the credential compromise rate. Future work will validate binding with UWB hardware at a roadside site, implement and evaluate the CMTV agents on real footage with measured error correlation, benchmark the device side on target officer hardware, and pilot the scheme with a participating security agency.

Statements and Declarations

Funding. The authors did not receive support from any organization for the submitted work.

Competing interests. The authors have no competing interests to declare that are relevant to the content of this article.

Ethics approval and consent to participate. Not applicable. This study involved no human participants, personal data or field trials. All credentials, identities and vehicles in the implementation and simulations are synthetic.

Consent for publication. Not applicable.

Data availability. The raw results of the timing benchmarks, binding simulation, fusion analysis and policy analysis generated in this study are available from the corresponding author on reasonable request.

Code availability. The reference implementation (`cppa_ref.py`) and repeated benchmark (`bench_repeat.py`), the binding simulation (`binding_sim.py`), the fusion analysis (`fusion_analysis.py`), the policy analysis (`analysis.py`) and the Tamarin models (`cppa.spthy`, the two unlinkability models and three further ablation variants) are available from the corresponding author on reasonable request.

Author contributions. A.A. conceived the study, designed the threat model and the protocol, built the Tamarin models, wrote the reference implementation and the simulations, carried out the analysis, prepared Figs. 1 to 5 and wrote the main manuscript text. A.B.I. and A.S.M. supervised the work and reviewed and edited the manuscript. O.M.O. checked the analysis and the results and reviewed and edited the manuscript. All authors reviewed and approved the final manuscript.

References

[1] Ojo, J.S., Oyewole, S., Aina, F.: Forces of terror: Armed banditry and insecurity in North-west Nigeria. Democracy and Security **19**(4), 319–346 (2023) https://doi.org/10.1080/17419166.2023.2164924

[2] Madueke, O.: Why should we worry about Nigeria's fragile security? The Political Quarterly **97**(1), 108–117 (2026) https://doi.org/10.1111/1467-923X.70010

[3] Olmos, R., Tabik, S., Herrera, F.: Automatic handgun detection alarm in videos using deep learning. Neurocomputing **275**, 66–72 (2018) https://doi.org/10.1016/j.neucom.2017.05.012

[4] Salazar González, J.L., Zaccaro, C., Álvarez-García, J.A., Soria Morillo, L.M., Sancho Caparrini, F.: Real-time gun detection in CCTV: An open problem. Neural Networks **132**, 297–308 (2020) https://doi.org/10.1016/j.neunet.2020.09.013

[5] Murugan, T., Noor Mohamed Badusha, N.A., Semaihi, A.R.O.A., Alkindi, M.M.R., Alnaqbi, E.M.R., Alketbi, G.H.T.: AI-based weapon detection for security surveillance: Recent research advances (2016–2025). Electronics **14**(23), 4609 (2025) https://doi.org/10.3390/electronics14234609

[6] Fierro Silva, C.J., Del-Valle-Soto, C., Bran, C., Varela-Aldás, J.: Comparative analysis of previous YOLO detectors and YOLOv26s for real-time weapon detection in video surveillance. Frontiers in Computer Science **8**, 1789702 (2026) https://doi.org/10.3389/fcomp.2026.1789702

[7] Al-Shareeda, M.A., Anbar, M., Manickam, S., Hasbullah, I.H.: SE-CPPA: A secure and efficient conditional privacy-preserving authentication scheme in vehicular ad-hoc networks. Sensors **21**(24), 8206 (2021) https://doi.org/10.3390/s21248206

[8] Kabil, A.M., Aslan, H., Azer, M.: Cryptanalysis of two conditional privacy preserving authentication schemes for vehicular ad hoc networks. Cryptography **8**(1), 4 (2024) https://doi.org/10.3390/cryptography8010004

[9] Shim, K.-A., Kwon, H.: On the security of aggregate signature-based conditional privacy-preserving authentication schemes for VANETs. ICT Express **12**(3), 566–570 (2026) https://doi.org/10.1016/j.icte.

2026.02.012

[10] Petit, J., Schaub, F., Feiri, M., Kargl, F.: Pseudonym schemes in vehicular networks: A survey. IEEE Communications Surveys & Tutorials **17**(1), 228–255 (2015) https://doi.org/10.1109/COMST.2014.2345420

[11] Aliu, A., Ikharo, A.B., Makinde, A.S., Obeten, O.M.: Real time detection of hand carried weapons for kidnapping mitigation in Nigeria: A YOLOv5-Faster R-CNN hybrid approach. International Journal of Computer Applications **187**(14), 14–21 (2025) https://doi.org/10.5120/ijca2025925295

[12] Boualouache, A., Senouci, S.-M., Moussaoui, S.: A survey on pseudonym changing strategies for vehicular ad-hoc networks. IEEE Communications Surveys & Tutorials **20**(1), 770–790 (2018) https://doi.org/10.1109/COMST.2017.2771522

[13] Brecht, B., Therriault, D., Weimerskirch, A., Whyte, W., Kumar, V., Hehn, T., Goudy, R.: A security credential management system for V2X communications. IEEE Transactions on Intelligent Transportation Systems **19**(12), 3850–3871 (2018) https://doi.org/10.1109/TITS.2018.2797529

[14] Li, H., Oza, D., Yoshida, M., Noguchi, T.: Privacy-preserving emergency vehicle authentication scheme using zero-knowledge proofs and blockchain. IoT **7**(2), 35 (2026) https://doi.org/10.3390/iot7020035

[15] Mollah, M.B., Zhao, J., Niyato, D., Guan, Y.L., Yuen, C., Sun, S., Lam, K.-Y., Koh, L.H.: Blockchain for the Internet of Vehicles towards intelligent transportation systems: A survey. IEEE Internet of Things Journal **8**(6), 4157–4185 (2021) https://doi.org/10.1109/JIOT.2020.3028368

[16] Goldwasser, S., Micali, S., Rackoff, C.: The knowledge complexity of interactive proof systems. SIAM Journal on Computing **18**(1), 186–208 (1989) https://doi.org/10.1137/0218012

[17] Camenisch, J., Lysyanskaya, A.: An efficient system for non-transferable anonymous credentials with optional anonymity revocation. In: Advances in Cryptology, EUROCRYPT 2001. Lecture Notes in Computer Science, vol. 2045, pp. 93–118. Springer, Berlin, Germany (2001). https://doi.org/10.1007/3-540-44987-6_7

[18] Brands, S., Chaum, D.: Distance-bounding protocols. In: Advances in Cryptology, EUROCRYPT '93. Lecture Notes in Computer Science, vol. 765, pp. 344–359. Springer, Berlin, Germany (1994). https://doi.org/10.1007/3-540-48285-7_30

[19] Francillon, A., Danev, B., Capkun, S.: Relay attacks on passive keyless entry and start systems in modern cars. In: Proceedings of the Network and Distributed System Security Symposium (NDSS). Internet Society, San Diego, CA, USA (2011)

[20] Singh, M., Leu, P., Capkun, S.: UWB with pulse reordering: Securing ranging against relay and physical-layer attacks. In: Proceedings of the Network and Distributed System Security Symposium (NDSS), San Diego, CA, USA (2019). https://doi.org/10.14722/ndss.2019.23109

[21] Bewley, A., Ge, Z., Ott, L., Ramos, F., Upcroft, B.: Simple online and realtime tracking. In: Proceedings of the IEEE International Conference on Image Processing (ICIP), pp. 3464–3468 (2016). https://doi.org/10.1109/ICIP.2016.7533003

[22] Radford, A., Kim, J.W., Hallacy, C., Ramesh, A., Goh, G., Agarwal, S., Sastry, G., Askell, A., Mishkin, P., Clark, J., Krueger, G., Sutskever, I.: Learning transferable visual models from natural language supervision. In: Proceedings of the 38th International Conference on Machine Learning (ICML). Proceedings of Machine Learning Research, vol. 139, pp. 8748–8763 (2021)

[23] Kuncheva, L.I., Whitaker, C.J.: Measures of diversity in classifier ensembles and their relationship with the ensemble accuracy. Machine Learning **51**(2), 181–207 (2003) https://doi.

org/10.1023/A:1022859003006

[24] Obeten, O., Jimoh, A.: Physics-informed, cost-aware fault classification with comparative explainability for predictive maintenance: A SHAP–LIME agreement analysis. Scientific Reports (2026) https://doi.org/10.1038/s41598-026-64374-2

[25] Demir, L., Kumar, A., Cunche, M., Lauradoux, C.: The pitfalls of hashing for privacy. IEEE Communications Surveys & Tutorials **20**(1), 551–565 (2018) https://doi.org/10.1109/COMST.2017.2747598

[26] European Data Protection Board: Guidelines 02/2025 on processing of personal data through blockchain technologies, Version 2.0. Adopted 7 July 2026 (2026). https://www.edpb.europa.eu/system/files/2026-07/edpb_guidelines_202502_blockchain_v2_en.pdf

[27] Federal Republic of Nigeria: Nigeria Data Protection Act, 2023. Signed into law 12 June 2023 (2023)

[28] Haji, I.A., Mohammed, S.D., Mousa, K.M.: Blockchain based chain of custody and digital evidence legality in post conflict prosecutions. Frontiers in Blockchain **9**, 1801364 (2026) https://doi.org/10.3389/fbloc.2026.1801364

[29] Dolev, D., Yao, A.C.: On the security of public key protocols. IEEE Transactions on Information Theory **29**(2), 198–208 (1983) https://doi.org/10.1109/TIT.1983.1056650

[30] Meier, S., Schmidt, B., Cremers, C., Basin, D.: The TAMARIN prover for the symbolic analysis of security protocols. In: Computer Aided Verification, CAV 2013. Lecture Notes in Computer Science, vol. 8044, pp. 696–701. Springer, Berlin, Germany (2013). https://doi.org/10.1007/978-3-642-39799-8_48

[31] Avižienis, A., Laprie, J.-C., Randell, B., Landwehr, C.: Basic concepts and taxonomy of dependable and secure computing. IEEE Transactions on Dependable and Secure Computing **1**(1), 11–33 (2004) https://doi.org/10.1109/TDSC.2004.2

[32] Lone, A.H., Mir, R.N.: Forensic-chain: Blockchain based digital forensics chain of custody with PoC in Hyperledger Composer. Digital Investigation **28**, 44–55 (2019) https://doi.org/10.1016/j.diin.2019.01.002

[33] Bernstein, D.J., Duif, N., Lange, T., Schwabe, P., Yang, B.-Y.: High-speed high-security signatures. Journal of Cryptographic Engineering **2**(2), 77–89 (2012) https://doi.org/10.1007/s13389-012-0027-1