applied sciences

MDPI

Article

A Lightweight Zero-Trust Authentication and Key Agreement Scheme for the Industrial Internet of Things

Xun Zhang [1], Zhiying Mu [1,*], Dejun Mu [1] and Xin Liu [2]

[1] School of Cybersecurity, Northwestern Polytechnical University, Xi'an 710072, China; xunzhang0931@outlook.com (X.Z.); mudejun@nwpu.edu.cn (D.M.)
[2] School of Information Science & Engineering, Lanzhou University, Lanzhou 730030, China; xinl@lzu.edu.cn
* Correspondence: muzhiying@mail.nwpu.edu.cn

Abstract

The Industrial Internet of Things (IIoT) demands authentication that protects resource-constrained field devices, supports fine-grained access control, and reduces reliance on implicitly trusted gateways. Existing IIoT authentication and key agreement schemes mainly verify cryptographic identity and establish session keys, but provide limited support for zero-trust access decisions after authentication such as role-specific operation control, gateway-side relation protection, emergency override, and context-aware re-verification. This paper proposes a lightweight zero-trust authentication and key agreement scheme for IIoT. The scheme embeds role-, device-, environment-, and behavior-aware control points into the authentication flow, protects gateway-side authentication relations and operator–asset mappings using secGear-based confidential computing, and supports pseudonym-based identity protection and break-the-glass emergency access. To complement deterministic access control, an optional auxiliary risk evaluation module provides trust evidence from contextual and operator-state signals without becoming part of the cryptographic critical path. This module is used to trigger re-verification, privilege degradation, audit escalation, or session termination before high-risk control privileges are released; it does not replace cryptographic authentication or constitute a formal guarantee of coercion detection, and the cryptographic layer remains fully functional and formally analyzable even when the auxiliary risk evaluation module is disabled. BAN logic and AVISPA analyses verify the cryptographic authentication and key agreement layer, while a comparative overhead analysis under an analytical operation count basis indicates low computation and communication cost relative to representative resource-constrained IoT authentication schemes.

Keywords: Industrial Internet of Things; zero-trust authentication; authentication and key agreement; confidential computing; access control

Check for updates

Academic Editor: George Drosatos

Received: 3 June 2026
Revised: 22 June 2026
Accepted: 29 June 2026
Published: 6 July 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

The Industrial Internet of Things (IIoT) has become the connective tissue of modern industrial operations, with field devices such as programmable logic controllers (PLCs), actuators, and remote terminal units directly governing the state of physical processes and distributed sensors continuously streaming process telemetry [1]. The scale at which IIoT now underpins critical infrastructure—power, water, manufacturing, and process plants—indicates the reach over which security failures could affect physical safety and the public [1]. Unlike general IoT incidents, attacks against industrial control assets translate

directly into physical consequences and threats to human life. The risk is concrete rather than hypothetical: in 2010, the Stuxnet worm reprogrammed Siemens PLCs to physically damage uranium enrichment centrifuges while reporting normal operation to operators [2]; in 2015 and again in 2016, coordinated attacks on the Ukrainian power grid issued unauthorized breaker-open commands and left customers without electricity, demonstrating remote control of safety-relevant actuators [3]; and in 2017, the TRITON/TRISIS malware specifically targeted the Triconex Safety-Instrumented System (SIS) of a petrochemical plant, attempting to disable the last-line safety controller that prevents catastrophic physical events [4]. These incidents establish IIoT security as first and foremost an operational and process safety concern, in addition to the more familiar data confidentiality concerns.

A second less-visible failure mode operates through people rather than firmware. Industry-wide breach statistics consistently identify the human element as the dominant root cause of security incidents: the Verizon 2024 Data Breach Investigations Report found that the human element is involved in 68% of breaches across all sectors, and that insider misuse and human error remain prominent contributors in the manufacturing and industrial verticals [5]. The mechanism connecting human state to physical risk is also well established in safety engineering: latent operator conditions such as fatigue, cognitive overload, and stress are recurrent contributors to control room error and major industrial accidents, with impaired situational awareness identified as a consistent mediator between human state and unsafe control actions [6]. The implication for IIoT authentication is that an authorized operator acting under coercion, severe psychological distress, or impairment is cryptographically indistinguishable from the same operator acting in a normal state, yet the operational safety consequence of an erroneous remote setpoint change or actuator command in those two regimes is very different. Therefore, human-factor risk is the deepest residual uncertainty in IIoT security, and the one that current authentication schemes do not measure.

Existing IIoT authentication schemes have mainly addressed mutual authentication, anonymity, and session key agreement; they remain necessary but no longer sufficient. Many still assume trusted authorities, trusted gateways, or perimeter-based security [7–10], which fails when a privileged gateway administrator, compromised operating system, or implanted malware can read authentication relation tables and operator–asset mappings from within the plant network. Zero-trust architectures attempt to remove this implicit trust by verifying each request against identity, device, environment, and behavioral context [11–13], but conventional zero-trust policy engines are in-line and add latency that is poorly tolerated by time-sensitive industrial control workflows [14,15]. Equally important, the four standard zero-trust dimensions (subject, device, environment, behavior) treat the human as a binary credential holder—authenticated or not—and do not measure the operator's psychological or affective state at the moment of authentication. Therefore, the human-factor dimension is a structural blind spot in existing IIoT authentication.

This motivates a multi-dimensional state-aware design that augments authentication with three coupled axes: *personnel state* (affective and behavioral signals correlated with coercion or distress), *environment state* (process area, terminal, time, network context), and *device state* (target device safety class and operation type). When the joint risk evaluation across these axes is favorable, full read/write device control is granted; when it is ambiguous, only data-observation privileges are released and high-risk remote control or parameter adjustment operations are withheld pending re-verification. This graduated permission model directly enforces the operational safety asymmetry that data observation is reversible whereas remote actuation is not. A second non-negotiable design constraint is that the field endpoints—resource-constrained controllers and sensor nodes operating under hard real-time and functional safety requirements—cannot host the AI inference

required for personnel state evaluation; their sub-MHz to low-MHz compute budgets, deterministic timing obligations, and safety-certified firmware make on-device neural inference both infeasible and operationally hazardous, and many such devices cannot be patched or re-flashed in the field over their multi-year service life. Therefore, all AI inference must reside on the industrial gateway side, with the gateway acting as the fusion point for personnel, environment, and device state evidence and as the issuer of permission decisions to the endpoint.

This paper proposes a lightweight zero-trust authentication and key agreement scheme for IIoT that realizes this multi-dimensional design. The cryptographic core establishes verifiable identity and a fresh session key between an industrial user and a field device with the assistance of the gateway, while the zero-trust access control layer determines whether the authenticated session should receive read-only, re-verification, or control privileges. The scheme embeds role-, device-, environment-, and behavior-aware control points into the authentication flow and protects gateway-side relations inside a secGear-based trusted execution environment. An optional auxiliary risk evaluation module can also be attached to the gateway-side policy engine to supply personnel state evidence from facial, behavioral, or physiological signals. This module is deliberately placed outside the formal cryptographic core: it is used to trigger re-verification, privilege degradation, session termination, or audit escalation rather than to replace identity authentication or serve as a mathematical proof of coerced login detection.

The research gap is threefold: existing IIoT AKE schemes (i) do not protect gateway-side authentication relations from privileged insiders, (ii) do not gate post-authentication access by role, device, environment, and behavior context, and (iii) do not account for operator state at the moment of authentication. The contributions are summarized as follows. First, we design a lightweight zero-trust authentication and key agreement scheme for IIoT that combines mutual authentication and session-key establishment with fine-grained multi-dimensional access control—enforcing role and device safety checks at the smart card, credential and authorization checks inside the gateway enclave, and environment and behavior checks at the access proxy—so that post-authentication access is continuously gated rather than granted unconditionally upon identity verification. Second, we protect gateway-side authentication relations and operator–asset mappings using secGear-based confidential computing and pseudonym-based identity management. Third, we embed environment- and behavior-aware control points into the access release procedure so that high-risk field device operations can be withheld or downgraded after cryptographic authentication when the context is abnormal. Fourth, we support break-the-glass emergency access and post hoc auditability, and evaluate the cryptographic protocol through BAN logic, AVISPA, and comparative overhead analysis; the auxiliary risk evaluation module is discussed as a replaceable zero-trust extension with thresholds that require site-side calibration.

2. Related Work

2.1. Lightweight Authentication for Resource-Constrained IoT

Authentication for resource-constrained IoT devices has been studied from the perspectives of lightweight cryptography, anonymity, and formal verification. Representative schemes for the industrial IoT and adjacent IoT domains share a common three-tier user–gateway–device structure. Li et al. [7] proposed a robust ECC-based authenticated key agreement scheme with privacy preservation for the industrial IoT, achieving mutual authentication and anonymity but evaluating elliptic-curve point multiplications on the resource-constrained sensor side. Xiong et al. [8] designed a privacy-preserving scheme with user revocability for multi-WSN industrial IoT, combining ECC with a one-

way hash chain to provide forward secrecy. Das et al. [9] presented a biometrics-based privacy-preserving user authentication scheme for cloud-based industrial IoT using only hash and XOR operations, reducing device-side cost but still assuming a trusted gateway. Vinoth et al. [10] proposed a multifactor authenticated key agreement scheme that uses secret sharing and the Chinese remainder theorem to let a user establish a session key with multiple sensing devices simultaneously. These schemes improve different aspects of lightweight IoT authentication, yet they do not jointly address privileged gateway insiders, emergency override, role-specific control operations, and coerced logins. Coerced-login detection has not been directly studied in the IIoT authentication context, but related techniques such as facial expression recognition [16,17] and physiological stress monitoring [18,19] provide candidate signals that have not been integrated into the cryptographic authentication pipeline of IIoT systems.

Beyond academic AKE schemes, a line of standardized lightweight security protocols addresses the same constrained regime at the transport or message security layer. EDHOC performs an ephemeral Diffie–Hellman key exchange over COSE to bootstrap an OSCORE security context [20,21]; DTLS 1.3 secures datagram transport against eavesdropping and forgery [22]; and the Noise framework provides a family of Diffie–Hellman handshake patterns [23], with EDHOC/OSCORE using compact CBOR/COSE structures and DTLS/Noise using their own compact native encodings [24,25]. These protocols are mature and interoperable, and their constrained-device cost has been measured on real microcontrollers [26,27]; their key establishment cores have also received deeper formal treatment than the symbolic analysis used in this paper, including Tamarin/ProVerif models and a comprehensive automated analysis of EDHOC [28–30]. They establish a secure channel but do not, on their own, protect gateway-side authentication relations inside a trusted execution environment, enforce role/device/environment/behavior-aware release of control privileges, or provide a break-the-glass workflow. Therefore, the proposed scheme is complementary to this standardized stack rather than a replacement for it. In Section 5.6, we position it against EDHOC, DTLS 1.3, and Noise; we also revisit the formal analysis boundary in Section 5.1.

The common limitation is architectural rather than merely cryptographic. Most protocols establish a session key between authorized entities, but do not ask whether the access context remains trustworthy after cryptographic identity is verified. In an industrial facility, a legitimate credential may be used under abnormal process area, terminal, time, or coercion conditions. Recent comprehensive treatments of IIoT security [1] likewise identify the lack of context- and human state-aware authentication as a persistent gap relative to the threat surface. This motivates integrating zero-trust control points directly into the authentication process.

2.2. Zero-Trust and Context-Aware Risk Evaluation

Zero-trust networking proposes the removal of implicit trust in network location and requirement of continuous verification before access is granted [11,12,31]. Prior work has examined policy enforcement, first-packet authentication, software-defined perimeter design, and domain-oriented zero-trust protection [14,15,32–34]; more recently, Edo et al. [13] have argued that zero-trust is the most credible architectural answer to the persistent rise of insider attacks in safety-critical information systems. Recent commentary has further argued that even “trust” itself can become a vulnerability when implicit assumptions persist within the perimeter [35]. These studies demonstrate the value of identity- and context-aware access control, but a complete lightweight scheme for resource-constrained IIoT devices remains underdeveloped. In addition, an inline policy engine may lengthen

the client-to-server connection time, which weakens suitability for time-critical industrial workflows.

Multimodal learning provides a complementary path for trust evaluation. Facial expression recognition has been studied using datasets such as AffectNet [16,17]; behavioral anomaly detection uses sequential access histories or reconstruction errors to detect deviations [36,37]; physiological stress detection exploits HRV, EDA, and related wearable signals [18,19]; and cross-modal representation learning has been shown to outperform single-modality baselines under noisy or missing-modality conditions in cyber–physical anomaly detection [37], an observation also reflected in broader multimodal machine learning surveys [38]. However, these techniques are usually deployed as standalone monitoring or audit tools. Existing IIoT authentication schemes do not integrate facial, behavioral, and physiological streams into a real-time session-level trust evaluator. In the proposed architecture, such signals are treated only as auxiliary risk evidence in the zero-trust access-control layer: they can trigger re-verification, privilege degradation, or audit escalation without changing the primary key agreement protocol.

Table 1 summarizes how the proposed work differs from the closest lines of research. The point of comparison is not only whether a scheme uses formal verification, but whether it changes the trust boundary of the industrial gateway and whether it supports a login-time response to duress. This distinction is important because an authentication protocol may be cryptographically sound while still allowing an authenticated but coerced or internally compromised session.

Table 1. Position of the proposed scheme relative to prior work.

Line of Work	Representative Techniques	Main Strength	Remaining Gap for IIoT
Lightweight IIoT AKE	ECC-based AKE, hash chains, secret sharing, CRT	Low computation and formal verification	Gateway and authority are usually trusted.
Privacy-preserving monitoring	Pseudonyms, group signatures, anonymous credentials	Pseudonyms and user anonymity	Emergency access and insider resistance are limited.
Zero-trust networking	SDP, first-packet auth, policy engines, micro-segmentation	Context-aware access decisions	Inline policy engines may add latency.
Auxiliary risk evaluation	FER, behavioral anomaly detection, physiological stress monitoring	Additional context and operator-state signals	Not part of conventional authentication protocols.
This work	Hash/symmetric AKE, secGear TEE, late-fusion multimodal scoring, parallel veto	Lightweight AKE plus zero-trust access control	Auxiliary risk evaluation requires site-side calibration.

A detailed property-by-property security comparison is provided later in the evaluation section.

2.3. Human Factors in Industrial IoT Security

The four standard zero-trust dimensions (subject, device, environment, behavior) treat the human as a binary credential holder and do not measure the operator's psychological or affective state at the moment of authentication. This omission matters because human factors dominate the threat landscape rather than constituting a residual class. The Verizon 2024 DBIR reports that the human element is involved in 68% of all breaches, with insider misuse and human error remaining prominent in the manufacturing and industrial verticals [5]. The mechanism connecting human state to physical risk is also well established in

safety engineering: latent operator conditions such as fatigue, cognitive overload, distraction, and stress are recurrent contributors to unsafe control actions and major industrial accidents, where degraded situational awareness mediates the path from human state to error [6]. Coercion is the acute end of this spectrum—an operator forced to authenticate under duress—while psychological distress, financial pressure, and severe fatigue form a continuous background that an authenticated-but-impaired operator acts within. None of the standard zero-trust dimensions resolves this regime.

Existing work has studied facial expression recognition [16,17], behavioral anomaly detection [36,37], and physiological stress detection [18,19] as standalone monitoring or audit tools, not as authentication-time evidence. The proposed scheme treats these signals as an optional fifth soft-trust dimension within the access control layer, consisting of a personnel-state risk indicator fused at the gateway and consumed by the same decision pipeline that already evaluates role, device, environment, and behavior. Therefore, the contribution is architectural rather than algorithmic: not a new affect-recognition model but a bounded way to use human-state evidence for re-verification, privilege degradation, and audit escalation without slowing the cryptographic path.

3. System Model and Preliminaries

3.1. Network Model and Entities

The system contains industrial users (operators) (U_i), smart cards (SC), an industrial gateway node (GWN), field devices (FD_j), an access proxy, and a bypass policy engine. Figure 1 illustrates how these entities are organized and how the cryptographic critical path is separated from the bypass-deployed policy engine. Industrial users include control engineers, field operators, maintenance technicians, and on-call (emergency response) engineers. Each user authenticates with three factors: SC, password, and fingerprint. Each user also has a role attribute $Role_i$ and field-device access list $FDLIST_i$. Each device has an identity ID_j, safety level $SecLevel_j$, and permitted operation list $OPLIST_j$.

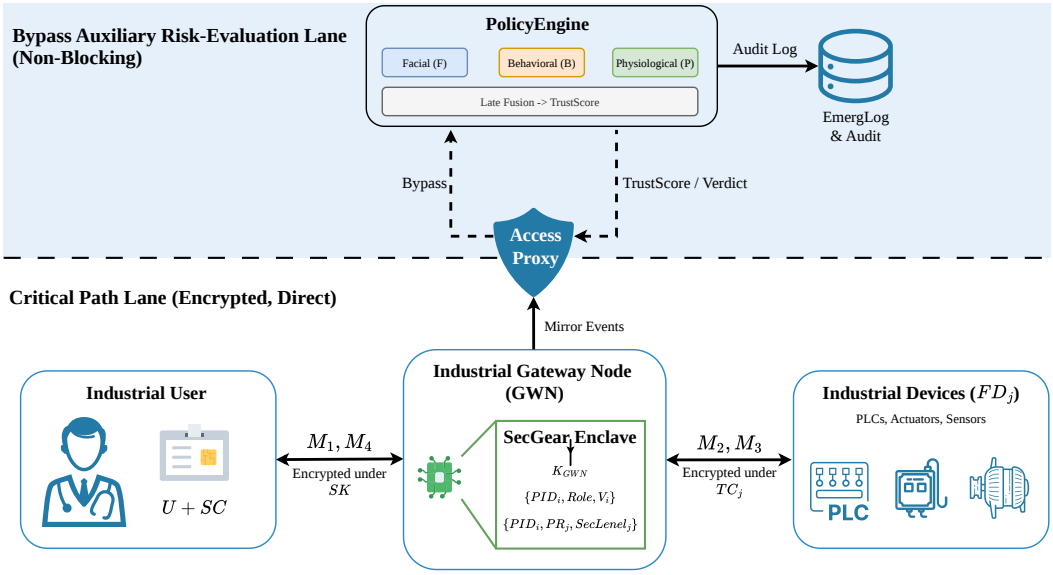

Figure 1. System architecture of the proposed zero-trust IIoT authentication scheme. The cryptographic critical path (solid arrows) flows through U_i, SC, the industrial gateway node (GWN), and field devices (FD_j); the access proxy mirrors authentication events to the bypass policy engine that hosts the auxiliary risk evaluation module (dashed arrows). secGear-protected enclave storage, role/level access control, and break-the-glass emergency channels separate context-aware AI evaluation from the cryptographic core, enabling lightweight authentication on resource-constrained field endpoints.

The GWN is deployed at the plant edge and authenticates both users and field devices. It forwards protocol messages and stores authentication relation tables, role assignments, operator–asset mappings, and device authorization records inside a secGear-based trusted execution environment (TEE), so that the gateway operating system and privileged administrators cannot directly read or modify them [39,40]. secGear is openEuler's confidential computing framework, supporting Intel SGX [41] and ARM TrustZone [42] backends, with AMD SEV [43] as a comparable alternative root. Recent applications of secGear-style confidential computing have demonstrated its viability across domains, including federated learning node authentication [44], blockchain-anchored Internet-of-Vehicles authentication [45], and enclave-protected key–value storage [46], supporting the design choice of placing IIoT-side relation tables inside the same enclave abstraction. The access proxy mirrors login and authentication events to the policy engine, which hosts the auxiliary risk evaluation module and returns a trust decision; if the score is high, the proxy can request secondary verification or terminate the session before application-layer access is released. This separation lets the cryptographic protocol remain executable even when the policy engine is temporarily slow while still allowing the session to be stopped before control commands reach the application layer.

Gateway availability and failure containment. Because the GWN concentrates authentication, relay, and policy functions, a production deployment should provision it for availability: redundant gateways with health-checked failover, replication of the secGear-protected relation store across attested enclaves, and graceful degradation of the bypass policy engine all apply. The impact of a compromised gateway is also bounded. Relation tables and K_{GWN} never leave the enclave, so a host-OS compromise on one gateway yields only pseudonymized or encrypted material rather than operator–asset relationships; device-side pairing material is recoverable only at the SC and FD from PTC_j, so a single gateway compromise does not by itself expose field-device channels; and the key rotation and revocation procedures of Section 4.8 bound the period during which stale material remains usable.

3.2. Threat Model and Assumptions

We adopt the Dolev–Yao model [47]: an adversary can eavesdrop, replay, delete, and modify public-channel messages, and may attempt impersonation, man-in-the-middle, replay, or denial-of-service attacks. We additionally consider three IIoT-specific threats: a privileged insider exfiltrating authentication relation tables or operator–asset mappings from the GWN; malware on the gateway operating system attempting to read keys or tamper with authorization records; and a coerced-login attacker physically forcing an authorized operator to authenticate under duress. We assume that TEE attestation cannot be forged and that enclave memory is isolated from the gateway OS, which places only the secGear enclave, not the host OS, inside the trust boundary at the GWN side. Field devices in a guarded plant are not trivially captured; portable terminals or wearables may be stolen, but since pairing keys are held at the SC and GWN rather than embedded plainly in the device, theft does not directly compromise the device channel. The auxiliary risk evaluation server resides inside the plant security boundary. The SC resists casual extraction but its user may still be coerced; the FD is resource-constrained but not trivially clonable. Attacks against the TEE hardware itself are outside the scope of this paper. These include microarchitectural and transient execution attacks against Intel SGX such as Foreshadow and Plundervolt, and cache side channels against ARM TrustZone [48–50]; they form a separate research area, and a full deployment threat model should add frequent re-attestation and side channel-resistant enclave coding [51]. For attestation we follow the terminology of the IETF Remote ATtestation procedureS (RATS) architecture [52], in which the secGear

enclave is the Attester and the gateway policy logic is a Relying Party; attestation evidence such as an Entity Attestation Token [53] and a concise software inventory [54] can be carried in compact CBOR/COSE form, and constrained-disclosure attestation [55] keeps this to pseudonymized identifiers rather than raw operator–asset relations.

Operator and asset identities are never used as plaintext routing keys in the authentication records. Operator–asset and asset–process mappings are maintained inside the TEE and referenced through pseudonymous identifiers, so a host-OS dump reveals only encrypted or hashed identifiers, not the underlying operator–asset relationship or process topology. Table 2 lists the key symbols used throughout the paper.

Table 2. Key notation.

Symbol	Description
U_i, FD_j, GWN	industrial user (operator), field device, gateway node
ID_i, ID_j, ID_{GWN}	identities of user, device, and gateway
$Role_i$, $SecLevel_j$	user role and device security level
SK, KEY_{ij}	U–GWN symmetric key and U–FD session key
K_{GWN}	GWN master key stored inside the TEE
PID_i, PID_j	pseudonym identities of user and device
$PID_i^{(s)}$	single-session pseudonym derived inside the enclave
T, N_i, N_j	timestamp and session nonces
e, b, p	facial, behavioral, and physiological anomaly scores
$\sigma(\cdot)$	logistic sigmoid function $\sigma(t) = 1/(1 + e^{-t})$
$\mu_u, \hat{s}_u$	per-user mean and standard deviation of LSTM error
$\tau_b, \hat{s}_b$	shift and scale parameters of the behavioral sigmoid
α_k	fusion weight for modality k
$\theta_{low}, \theta_{high}$	re-verification and rejection thresholds
$\Delta\theta$	emergency mode threshold offset added to $\theta_{low}, \theta_{high}$

4. Proposed Scheme

This section presents the proposed scheme in two parts. Sections 4.1–4.3 define the zero-trust access control layer: the four deterministic control-point dimensions, the optional auxiliary risk evaluation module and its fusion logic, and the parallel-veto procedure that governs the interaction between the cryptographic path and the auxiliary verdict. The architecture is driven by three constraints: (i) field endpoints cannot host AI inference due to sub-MHz compute budgets, deterministic timing obligations, and safety-certified firmware; (ii) gateway-side authentication relations must be protected from privileged insiders via a trusted execution environment; and (iii) auxiliary risk evaluation must not delay the cryptographic critical path, motivating the bypass/parallel-veto deployment. Sections 4.4–4.8 then specify the cryptographic protocol itself: registration, login and key agreement, message flow, auxiliary authentication with emergency handling, and the key management lifecycle.

4.1. Control Points

The proposed scheme extracts four zero-trust dimensions and embeds them into scheme execution. The role (subject) dimension checks the user role and access list (control engineers can read and adjust assigned devices, field operators can read monitoring points and acknowledge alarms, maintenance technicians can calibrate designated devices, and on-call engineers obtain temporary access through break-the-glass). The device (object) dimension classifies field devices by safety impact, consistent with functional-safety integrity levels [56]: high-risk devices (safety-instrumented systems and emergency shutdown actuators) require Level-3 privileges and secondary confirmation; medium-risk

devices (regulatory control loops) require Level-2 or higher; low-risk monitoring-only sensors require Level-1 or higher. The environment dimension verifies login time range, process-area range, terminal type, and MAC address. The behavior dimension compares the current access event with historical login time, process area, terminal type, and operation sequence; it combines rule-based fallbacks with an LSTM-based anomaly score (e.g., bulk access to multiple field devices or sudden parameter adjustment requests by a user whose history contains only read operations will raise the behavioral score).

The four dimensions are intentionally checked at different locations. The SC performs the earliest role/device screening because it already stores $FDLIST_i$ and $Role_i$; the GWN performs stronger checks against the TEE-protected relation tables, process-area lists, terminal lists, MAC whitelist, and gateway-managed device list; the access proxy and policy engine evaluate softer signals that do not fit naturally into deterministic cryptographic verification. A login request is rejected immediately on a hard-constraint failure (invalid timestamp, unmanaged device, role level below device requirement), while ambiguous context changes are forwarded to the policy engine—for example, access from a different process area during a cross-area intervention, or a night-shift login during an upset—and treated as re-verification events. Emergency access is handled by a dedicated *EmergToken* on a dedicated emergency response SC, activated by an *EmergCode* entered by the on-call engineer (Section 4.7); environment and behavior control points are bypassed while subject identity and device-operation privileges remain active, and all emergency operations are logged for post hoc audit.

4.2. Auxiliary Risk Evaluation

Beyond the four classical zero-trust dimensions (role, device, environment, behavior), the deepest residual uncertainty in safety-critical industrial authentication is human-factor risk: the possibility that a cryptographically valid and contextually plausible login is nevertheless being executed by an operator under coercion, severe psychological distress, or significant cognitive impairment. As Section 2.3 argued, industry-wide breach data [5] and safety engineering evidence linking impaired operator state to unsafe control actions [6] both point to human factors as a dominant unaddressed risk in industrial-sector security. The framework presented in this subsection introduces facial, behavioral, and physiological signals as a fifth soft trust dimension—a personnel-state evaluator—that augments rather than replaces the cryptographic authentication path.

The auxiliary risk evaluation module treats a login event as a configurable trust evaluation problem. It uses three independent modality modules and a late-fusion layer, summarized in Figure 2. The facial expression modality F captures a short camera frame window during login and outputs a risk score $e \in [0, 1]$. We adopt a MobileNetV2-based FER model pretrained on AffectNet-8 [16,57]; MobileNetV2 is chosen for its lightweight inverted-residual architecture, which reduces inference cost to approximately 12 ms on a gateway-side CPU without GPU—an order of magnitude below ResNet or VGG alternatives—while providing well-documented AffectNet-8 baselines for reproducible comparison. The model outputs emotion probabilities, and the risk score is computed as

$$e = w_1 p_{fear} + w_2 p_{anger} + w_3 \max([0, 1] - p_{neutral} - p_{happy}),$$

where $w_1 = 0.45$, $w_2 = 0.35$, and $w_3 = 0.20$. This formulation directly models coercion-associated affect (fear and anger) rather than inferring it from the absence of neutral or happy probabilities, which is more robust when a coerced user attempts to suppress emotional expression by feigning calm. A larger value indicates stronger evidence of coercion-associated distress. A five-frame sliding window is used to reduce sensitivity to transient facial changes.

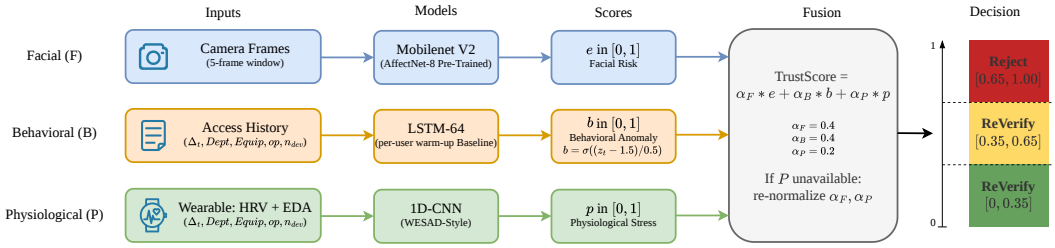

Figure 2. Auxiliary risk evaluation module. Three modality modules—facial expression (F, MobileNetV2 on AffectNet-8), behavioral sequence (B, LSTM on per-user access history), and physiological stress (P, 1D-CNN on WESAD-style wearable signals)—produce independent risk scores e, b, and p. A late-fusion layer combines them into a single *TrustScore*, which is mapped to one of three decisions (Accept/Re-verify/Reject) by configurable thresholds θ_{low} and θ_{high}. When the physiological channel is unavailable, the remaining weights α_F and α_B are renormalized. The asterisk symbol in the figure denotes scalar multiplication in the weighted late-fusion operation.

The behavioral modality B models sequential access histories. Each login event is encoded as

$$x_t = [\Delta t, dept_t, equip_t, op_t, n_{dev,t}],$$

where Δt is the elapsed time since the previous login, $dept_t$ is the process area, $equip_t$ is the terminal type, op_t is the requested operation, and $n_{dev,t}$ is the number of devices accessed in a recent window. Categorical fields ($dept_t$, $equip_t$, op_t) are mapped to learned embeddings of dimension 8 before concatenation with the continuous fields (Δt, $n_{dev,t}$), so that the L2 distance $\|x_t - \hat{x}_t\|_2$ in (1) is well-defined in the joint embedding space. A lightweight LSTM with 64 hidden units (LSTM-64) [58]—selected because the low-dimensional access history feature vector (five fields per event) does not require the capacity of deeper models, and LSTM-64 offers deterministic per-step latency suitable for the authentication window—predicts $\hat{x}_t$ from the user's history. Let μ_u and $\hat{s}_u$ denote the user's running mean and standard deviation of the reconstruction error $\|x_t - \hat{x}_t\|_2$, and let $z_t = (\|x_t - \hat{x}_t\|_2 - \mu_u)/\hat{s}_u$ be the corresponding z-score. Throughout this paper, $\sigma(\cdot)$ denotes the logistic sigmoid function $\sigma(t) = 1/(1 + e^{-t})$; the standard-deviation symbol always carries a subscript ($\hat{s}_u$ here, $\hat{s}_b$ below) to avoid conflicting notation. Direct sigmoid evaluation $\sigma(z_t)$ would map a typical (well-modeled) login to $\sigma(0) = 0.5$, which would inflate the score for normal events; therefore, we use a shifted and scaled sigmoid as the behavioral anomaly score:

$$b = \sigma\left(\frac{z_t - \tau_b}{\hat{s}_b}\right) \tag{1}$$

with default $\tau_b = 1.5$ and $\hat{s}_b = 0.5$, so that $b \approx 0$ for $z_t \lesssim 1$, the transition region $b \in [0.3, 0.7]$ corresponds to $z_t \in [1.0, 2.0]$, and $b \to 1$ for $z_t \gtrsim 2.5$ (i.e., events at least $2.5\hat{s}_u$ from the user's typical reconstruction error). Therefore, the pair $(\tau_b, \hat{s}_b)$ is an interpretable per-user calibration knob: τ_b sets the z-score at which the score crosses 0.5, and $\hat{s}_b$ controls how sharply the score saturates. The module falls back to rule-based checks when insufficient user history is available.

The physiological modality P uses wearable-derived heart-rate and electrodermal-activity windows, when available, and outputs a stress probability $p \in [0, 1]$. Following the WESAD-style setting [18], a compact 1D-CNN with convolutional layers and global pooling is sufficient for CPU-side inference; this architecture is the standard lightweight choice for time-domain HRV/EDA classification in the WESAD benchmark family, providing adequate binary accuracy ($F_1 \approx 0.95$) at minimal parameter count and latency. This modality is optional. If no wearable is available, its weight is set to zero and the remaining weights are re-normalized.

The three scores are fused as

$$TrustScore = \alpha_F e + \alpha_B b + \alpha_P p, \quad \sum_k \alpha_k = 1.$$

A higher *TrustScore* means a more anomalous login context. The decision rule is

$$Decision = \begin{cases} Accept, & TrustScore < \theta_{low}, \\ Re\text{-}verify, & \theta_{low} \leq TrustScore < \theta_{high}, \\ Reject, & TrustScore \geq \theta_{high}. \end{cases}$$

Weight and threshold methodology. We present the fusion weights and thresholds as documented initial settings inside an explicit admissibility region, not as empirically optimal or uniquely derived deployment parameters. The facial sub-weights (w_1, w_2, w_3) encode a configurable affect prior that ranks coercion-associated classes (fear, anger) above a residual-distress term. The fusion-weight region is constrained by three design requirements: (i) a single-modality break-even constraint, so that with the thresholds below no single modality can by itself push *TrustScore* past θ_{high}; (ii) a clean two-modality fallback, so that removing the physiological channel renormalizes the remaining facial and behavioral weights to 0.5/0.5; and (iii) treating physiological evidence as supporting, because its deployment availability is lower (wearable pairing, contact quality, and motion artefacts vary across operators and shifts, so over-weighting it risks alert flapping), while weighting facial and behavioral signals equally because they fail in complementary regimes (facial cues degrade under masks or suppression, behavioral cues when an attacker mimics the user's pattern). The default $\alpha_F = \alpha_B = 0.4$, $\alpha_P = 0.2$ is one illustrative point satisfying these constraints. The thresholds $\theta_{low} = 0.35$ and $\theta_{high} = 0.65$ then give the break-even property (per-modality maxima 0.4, 0.4, 0.2) and an intentionally asymmetric reject coverage: $F + B$ saturated jointly contribute $0.8 > \theta_{high}$ and can trigger rejection, whereas saturated $F + P$ or $B + P$ contribute only $0.6 < \theta_{high}$ and lift the score into re-verification. These values are documented initial settings, not universal deployment optima. We do not claim that site-side calibration has been carried out here; a reproducible sensitivity analysis over these weights and thresholds is reported in Section 5.5. The bypass deployment ensures that computing the auxiliary risk evaluation does not block key agreement: upon a login request, the access proxy mirrors the facial frame window and access features to the policy engine, the auxiliary risk evaluation module produces the trust score within the authentication window, and high-risk device control release is held or terminated before application-layer control privileges are issued if the score exceeds θ_{high}.

Site-side calibration of fusion weights and thresholds. In deployment, the default weights and thresholds should be re-estimated from site-side calibration data collected under the plant's own lighting, PPE, terminal, shift, and operator population conditions. A practical calibration procedure has five steps. First, collect benign login sessions and controlled high-stress or reverification-triggering sessions under an approved safety and privacy protocol. Second, compute per-modality scores e, b, and p using the same scalar interfaces used by the auxiliary risk evaluation module. Third, estimate candidate fusion weights α_F, α_B, and α_P by constrained grid search, AHP-assisted expert weighting, entropy weighting, or ROC/F_β-based optimization [59,60]. Fourth, choose θ_{low} and θ_{high} to satisfy site-defined operating constraints, such as an upper bound on benign false rejection and a minimum re-verification rate for high-risk sessions. Fifth, validate the resulting operating point on

a held-out calibration split and periodically re-estimate it when devices, PPE, lighting, or operating policies change. One deployment objective can be written as

$$\begin{aligned} \min_{\alpha,\theta} \quad & \lambda_1 FPR_{benign} + \lambda_2 FN_{high\text{-}risk} + \lambda_3 ReverifyRate_{benign} \\ \text{s.t.} \quad & \alpha_F + \alpha_B + \alpha_P = 1, \quad \alpha_k \geq 0, \\ & FPR_{benign} \leq \rho_{max}, \quad FN_{high\text{-}risk} \leq \eta_{max}. \end{aligned}$$

In the absence of such site-side calibration data, the default configuration is used only to illustrate sensitivity behavior, and should not be interpreted as a plant-ready coercion detector.

Bypass with parallel veto. The auxiliary risk evaluation module is logically off the cryptographic critical path: the key agreement of Section 4.5 completes without waiting for the auxiliary-risk verdict, so cryptographic completion is governed by the protocol path rather than by $t_{crypto} + t_{aux}$. However, release of high-risk device control privileges remains gated at the access proxy. We refer to this combination as a *parallel veto*: the auxiliary risk evaluation module runs concurrently with the protocol, while the access proxy withholds, downgrades, or terminates access according to the verdict and the device safety level. Section 4.3 specifies its commit point, timing, and race behavior.

Training and deployment. The three modality models are trained offline and updated on a maintenance cadence separate from the cryptographic protocol. The FER backbone is pretrained on a public benchmark and may be fine-tuned on site-collected facial frames under explicit operator consent. The behavioral LSTM uses a per-user warm-up period during which the rule-based fallback dominates. The physiological 1D-CNN is trained on the WESAD-style benchmark and accepts wearable streams only when device pairing and signal quality checks pass. Federated personalization across plant sites is identified as future work to avoid centralizing sensitive multimodal data. Per-module inference latency is approximately 12 ms (FER, MobileNetV2 on AffectNet-8), 3 ms (Behavior, LSTM-64 on user history), and 2 ms (Physio., 1D-CNN on WESAD); these are representative benchmark-level estimates for the architectures rather than measured per-platform figures (see Section 5.4), and with the sub-millisecond fusion they give a total auxiliary risk evaluation latency below 18 ms.

4.3. Parallel Veto Procedure: Timing, Commit Point, and Race Freedom

The parallel veto introduced above needs a precise operational model for the interaction between the deterministic and formally analyzed key agreement and the probabilistic auxiliary layer. Three properties define this interaction.

Monotonicity-invariant. The auxiliary risk evaluation module can return only one of $\{\mathsf{maintain}, \mathsf{re\text{-}verify}, \mathsf{degrade}, \mathsf{terminate}\}$, and has no output that grants a privilege that the cryptographic or role/device check has denied. The single mechanism that relaxes a check is the break-the-glass emergency path, which is itself gated by three-factor authentication and non-repudiable logging (Section 4.7). Therefore, the probabilistic layer acts only after a successful deterministic authentication and can only tighten the resulting access; it cannot weaken any property established by the BAN/AVISPA analysis of Section 5. For this reason, the module is kept outside the symbolic proof and a probabilistic decision cannot affect the cryptographic guarantees.

Commit point and timing. The release of high-risk device control privileges to the application layer is the single commit point and the access proxy gates that release on the auxiliary verdict, as specified in Algorithm 1. Under the benchmark-level latency estimates of Section 5.4 ($t_{aux} \approx 18$ ms against $t_{crypto} \approx 87.3$ ms), the verdict is normally available before the commit point. However, race freedom does not rely on this timing margin: when

the verdict is late, the procedure applies a default (below) that is dependent on the safety level, so it is the commit-point gate together with this fallback that prevents races, not the latency margin.

Algorithm 1 Parallel veto procedure at the access proxy (per login)

Require: login event ℓ; thresholds $\theta_{low}, \theta_{high}$; auxiliary window W; device safety level $SecLevel_j$

start the cryptographic path (Section 4.5); **in parallel**, mirror ℓ (read-only) to the policy engine
$v \leftarrow \perp$ ▷ auxiliary verdict, initially unavailable
on auxiliary result, even after W: $v \leftarrow TrustScore(\ell)$
wait until the cryptographic path reaches the commit point (release of high-risk device-control privileges)
if $v \neq \perp$ **then** ▷ verdict available: normal case, since $t_{aux} \ll t_{crypto}$
    if $v < \theta_{low}$ **then commit**: release full control ▷ Accept
    else if $v < \theta_{high}$ **then hold**: require step-up re-verification ▷ Re-verify
    else veto: terminate or restrict release ▷ Reject/degrade
    end if
else ▷ verdict late (degraded conditions): safety-level-dependent default
    if $SecLevel_j$ is high **then veto by default**: degrade/deny pending v
    else read-only monitored release; if a later $v \geq \theta_{high}$ arrives, terminate or downgrade
    end if
end if
append (modality availability, v, decision) to the audit log

Late-verdict fallback. When the verdict is not ready at the commit point, the policy depends on the target device's safety level: high-*SecLevel* operations are degraded or denied until the verdict arrives, while low-*SecLevel* operations may receive read-only monitored access and are terminated or downgraded if a later verdict exceeds θ_{high}. Thus, a late verdict can affect an already authenticated session, but high-risk control privileges are not released before the gate is satisfied.

Race freedom. The auxiliary risk evaluation module reads a copy of the authentication events and shares no mutable cryptographic state with the key agreement path, so there is no data race on the cryptographic computation. The only synchronization point is the commit-point gate at the access proxy (Algorithm 1), where the module can withhold or downgrade access but can never create it.

The preceding subsections define the zero-trust access control policy and the auxiliary risk evaluation mechanism; the remainder of this section specifies the cryptographic authentication protocol that implements them.

4.4. Initialization and Registration

The system administrator assigns ID_j and $SecLevel_j$ to each field device. The GWN generates an SM2 public/private key pair (K_p, K_v) (SM2 is a Chinese national standard elliptic-curve cryptosystem (GB/T 32918 [61], ISO/IEC 14888-3 [62]) that provides digital signatures and key exchange at a 128-bit security level, equivalent in strength to NIST P-256 [63]) and initializes K_{GWN} inside the secGear enclave. The GWN also stores $GWDLIST$, authorized time ranges, process-area lists, terminal lists, MAC address lists, and historical behavior records. Each smart card receives $FDLIST_i$, a list of $(ID_j, SecLevel_j)$ pairs that the user is authorized to access, together with the user's $Role_i$, and each field device stores $OPLIST_j$. Embedding $SecLevel_j$ alongside ID_j in $FDLIST_i$ enables the SC to complete role-level matching locally without an extra round-trip to the GWN.

During user registration, U_i submits the ID_i, password PW_i, and role information through a secure out-of-band U–GWN channel (a controlled plant IT enrollment terminal that is independently authenticated to the GWN; the session key SK of Section 4.5 is established only at login time, and is not yet available at registration). The user also enrolls a fingerprint template at the SC personalization step; the template is written to the SC's tamper-resistant store and used only for on-card match-on-card verification, so that no biometric data ever leaves the SC or reaches the GWN. The terminal computes $UID_i = h(ID_i \| PW_i)$ and transmits the registration parameters, including UID_i, over this enrollment channel. Inside the enclave, the GWN computes $PID_i = h(UID_i \| K_{GWN})$ and stores the tuple $\{PID_i, Role_i, V_i\}$, where $V_i = \mathrm{HMAC}_{K_{GWN}}(PID_i \| Role_i)$ is a record integrity tag (instantiated as a standard keyed-hash MAC, e.g., HMAC-SHA-256, rather than as a length-extendable plain hash) that is verified on every enclave lookup so that an attacker who somehow obtains write access to the encrypted store cannot tamper with the role binding without invalidating V_i. This design avoids storing plaintext identities or passwords at the GWN and confines biometric material to the SC. For on-call engineers, an *EmergToken* is also written to a dedicated emergency smart card.

During device registration, FD_j sends ID_j within a secure environment. The GWN generates R_j, then computes $TC_j = h(ID_j \| R_j)$, $UID_j = h(ID_j \| ID_{GWN})$, $PTC_j = TC_j \oplus UID_j$, and $PID_j = h(ID_j \| K_{GWN})$ inside the enclave. The tuple $\{PID_j, PR_j, SecLevel_j\}$ is stored in the enclave, where $PR_j = R_j \oplus ID_{GWN}$. The device stores PTC_j and $OPLIST_j$. Since ID_{GWN} is a public gateway identifier that is pre-loaded at device manufacture, FD_j can locally recompute UID_j and recover $TC_j = PTC_j \oplus UID_j$ at session time without exposing TC_j in non-volatile storage; therefore, UID_j is a device-local helper and is never transmitted on the wire.

4.5. Login and Key Agreement

For channel establishment, the terminal sends r_1 along with a request flag to the GWN over the public channel. The GWN replies with r_2 together with an SM2 signature on (r_1, r_2) generated by its private key K_v. The terminal verifies this signature against the GWN's public key certificate, then encrypts r_3 under the GWN public key K_p and returns it. Both parties derive

$$SK = h(r_1 \| r_2 \| r_3 \| ID_{GWN}).$$

The GWN signature provides authenticated key establishment, preventing a man-in-the-middle from substituting r_2, while the certificate binding ensures the terminal trusts only the institution-registered GWN. The resulting SK protects login messages between U_i and the GWN; enrollment-time registration uses the separate secure out-of-band channel described in Section 4.4.

During login, the user inserts the smart card, enters ID_i and PW_i, supplies the fingerprint, and specifies the target device ID_j. The SC first performs on-card fingerprint matching against the enrolled template; only on a successful local match does the SC unlock the cryptographic credentials and proceed (the match-on-card cost is specific to the hardware vendor—typically tens of milliseconds on modern smart cards—and is not absorbed into the operation-count basis used in the overhead analysis because no biometric reproduction takes place at the host terminal). The SC then verifies whether $ID_j \in FDLIST_i$ and whether $Role_i$ satisfies the required $SecLevel_j$. It generates a nonce N_i, obtains the terminal MAC address MAC_t, and sends the encrypted login parameters to the GWN under SK. The GWN checks the timestamp freshness, authorized time range, MAC address, gateway device list, and enclave-stored credentials. It retrieves PID_j, PR_j, and $SecLevel_j$ from the enclave, reconstructs $R_j = PR_j \oplus ID_{GWN}$ and the device authentication key $TC_j = h(ID_j \| R_j)$, and forwards a protected challenge to FD_j. The device verifies the message, generates N_j,

checks whether the requested operation is in $OPLIST_j$, and returns its response. Finally, U_i and FD_j derive

$$KEY_{ij} = h(N_i \| N_j),$$

which is used to protect subsequent process data and control commands.

The four messages for login and key agreement can be abstracted as follows:

$$M_1 : U_i \rightarrow GWN : E_{SK}(UID_i, ID_j, N_i, MAC_t, op, T_1), \quad (2)$$

$$M_2 : GWN \rightarrow FD_j : E_{TC_j}(PID_i^{(s)}, N_i, op, T_2, q_1), \quad (3)$$

$$M_3 : FD_j \rightarrow GWN : E_{TC_j}(PID_i^{(s)}, N_j, T_3, q_2), \quad (4)$$

$$M_4 : GWN \rightarrow U_i : E_{SK}(N_j, T_4), \quad (5)$$

where $E_K(\cdot)$ denotes authenticated symmetric encryption (AES–GCM in our reference implementation), $PID_i^{(s)} = \mathrm{HMAC}_{K_{GWN}}(PID_i \| N_i \| T_2)$ is a single-session pseudonym derived by the GWN inside the enclave to bind the session to the user without exposing PID_i to the device (binding to N_i and T_2 ensures that $PID_i^{(s)}$ is per-session fresh and unlinkable across sessions for a network observer who never sees K_{GWN}), $q_1 = \mathrm{MAC}_{TC_j}(PID_i^{(s)} \| N_i \| op \| T_2)$ and $q_2 = \mathrm{MAC}_{TC_j}(PID_i^{(s)} \| N_j \| T_3)$ are explicit MAC tags, and $T_1, \ldots, T_4$ are timestamps used for freshness checks. Because TC_j is held by both the GWN (after enclave reconstruction) and FD_j (after local recovery from PTC_j), the GWN acts as an authenticated relay that forwards N_i from U_i to FD_j under TC_j; the network attacker never observes N_i or N_j in the clear, while the GWN is permitted to learn them because it is the only entity that already holds both SK and TC_j. This formal abstraction is reused in the BAN-logic analysis of Section 5.1, where M_2 corresponds to $\{U_i \overset{N_i}{\leftrightarrow} FD_j, T_2\}_{TC_j}$ and M_3 to $\{U_i \overset{N_i}{\leftrightarrow} FD_j, T_3\}_{TC_j}$.

After key agreement, auxiliary authentication verifies the user's department, terminal type, and behavior history. Abnormal results are sent to the access proxy and policy engine. In emergency mode, the GWN verifies *EmergToken*, bypasses environment and behavior checks, enforces identity and role constraints, establishes the session key, records *EmergLog*, and invalidates the emergency token after use or expiration. Password updates are handled locally by the SC after verifying the old password and biometric factor.

4.6. Message Flow Details

The login and key agreement phase is the critical path. Its message flow is summarized in Table 3; the GWN mediates authentication and device authorization but does not learn application-layer commands.

Table 3. Login and key agreement message flow.

Step	Direction	Main Fields	Check Performed
1	$U_i \rightarrow GWN$	$E_{SK}(UID_i, ID_j, N_i, MAC_t, op, T_1)$.	SC checks $ID_j \in FDLIST_i$ and role sufficiency.
2	GWN local	TEE lookup of PID_i, role, PID_j, device level, and gateway device list.	Timestamp, MAC, time range, credential, and device authorization.
3	$GWN \rightarrow FD_j$	$E_{TC_j}(PID_i^{(s)}, N_i, op, T_2, q_1)$ where TC_j is reconstructed inside the enclave.	Device locally recovers $TC_j = PTC_j \oplus UID_j$, decrypts the payload, and verifies q_1 and T_2.
4	$FD_j \rightarrow GWN$	$E_{TC_j}(PID_i^{(s)}, N_j, T_3, q_2)$ after FD_j checks $op \in OPLIST_j$.	GWN verifies q_2, T_3, and forwards N_j under SK.
5	$GWN \rightarrow U_i$	$E_{SK}(N_j, T_4)$.	U_i verifies freshness; both endpoints derive $KEY_{ij} = h(N_i \| N_j)$.
Bypass	Proxy/policy	Mirrored event (parallel to steps 1–5): context features and auxiliary risk score.	Re-verification or session termination enforced before application-layer access is released.

4.7. Auxiliary Authentication and Emergency Handling

Auxiliary authentication is executed after the main key exchange but before sensitive application-layer operations are released. The smart card encrypts the current department, terminal type, and timestamp under KEY_{ij} and sends them to the GWN, which checks $DeptLIST$, $EquipLIST$, $MacLIST$, and consistency with historical records. These checks are placed after session key derivation because they require protected channel semantics and may involve private contextual attributes. Emergency access follows a stricter audit model rather than a stricter blocking model: on *EmergCode* entry, the emergency SC activates *EmergToken*; the GWN then verifies token validity and expiration, enforces three-factor authentication, and checks operation consistency with emergency privileges, while environment and behavior control points are bypassed because emergency care may legitimately occur outside normal shift, department, or terminal patterns.

To make the interaction between emergency mode and auxiliary risk evaluation explicit, four control layers operate during *EmergToken* activation, in increasing strictness: (i) environment and behavior rule checks are *bypassed*; (ii) the behavioral and physiological risk evaluation channels are downweighted (operators legitimately exhibit stress during time-critical interventions) and the facial channel uses an institution-configurable threshold offset $\Delta\theta$ added to both θ_{low} and θ_{high}; (iii) the facial channel remains *active*, with an extreme score above $\theta_{high} + \Delta\theta$ triggering a high-priority *EmergLog* entry plus an out-of-band notification to the on-duty security officer rather than a session block; and (iv) identity, role, and device-operation privilege checks remain *fully enforced*. Therefore, three residual barriers remain against *EmergToken* misuse: three-factor authentication on the dedicated emergency card, non-repudiable logging of the activation, and security officer escalation upon extreme facial scores. The decision to alert rather than block is consistent with break-the-glass practice, where the cost of false denial during an emergency shutdown or runaway process is typically worse than deferred forensic review; $\Delta\theta$ and the review SLA should be set together with the plant safety and compliance review.

All emergency operations are sent to the policy engine for post hoc audit, supporting regulatory reporting and detection of repeated abnormal *EmergToken* patterns. Password changes are local to the SC (after old password and biometric verification), avoiding network round-trips and frequent gateway-side updates. This scheme separates enrollment-time secrets from session-time nonces: N_i, N_j are freshly generated at each login, supporting per-session key freshness and session key independence under one-way function assumptions, while replayed messages fail timestamp and MAC checks.

4.8. Key Management and Revocation

The scheme uses a small key hierarchy. The gateway holds an SM2 key pair (K_p, K_v) and a symmetric master key K_{GWN}, both of which are generated and kept inside the secGear enclave and never exported. The per-session U–GWN channel key SK is established during the login channel establishment step of Section 4.5; long-term gateway-side derivations from K_{GWN} include the pseudonyms PID_i, PID_j, and the single-session $PID_i^{(s)}$. The device authentication key $TC_j = h(ID_j \| R_j)$ is provisioned during device registration. Session confidentiality rests on the fresh key $KEY_{ij} = h(N_i \| N_j)$, which is recomputed from fresh nonces at every login and is never stored.

Key rotation is performed at the gateway; K_{GWN} and the SM2 key pair are rotated on a fixed schedule or on demand, entirely inside the enclave. Since PID_i, PID_j, record integrity tags, and session pseudonyms are all derived from K_{GWN}, rotating this key requires a controlled enclave migration: the old-key enclave exports only authenticated plaintext enrollment records to a freshly attested new-key enclave; the new enclave recomputes pseudonyms and integrity tags under the new K_{GWN}, and old key versions are retained only

for a bounded grace window needed to complete in-flight sessions. Device key material is refreshed by recomputing PTC_j during scheduled maintenance windows when device credentials are rotated.

Revocation is a local enclave operation and does not require contacting the field devices. To revoke a user, the gateway deletes the user's entries from the enclave-stored relation tables and role assignments and removes the associated terminal from the authorized lists; any later login then fails at the gateway credential check. To revoke a device, the gateway deletes its record from $GWDLIST$ together with its $(PID_j, PR_j, SecLevel_j)$ tuple, after which the gateway no longer reconstructs TC_j for that device, so the protected exchange M_2/M_3 can no longer be completed for it. Emergency credentials follow the same lifecycle: an $EmergToken$ is single-use and time-bounded, and is invalidated after use or expiry (Section 4.7).

Because the session key depends on fresh per-login nonces, compromising one derived session key does not by itself expose past or future session keys; this is the session key independence property noted in Section 4.5. Recovery from a gateway compromise is bounded by the same rotation and revocation steps: the operator re-keys K_{GWN} and the SM2 pair inside a freshly attested enclave, migrates and re-derives the pseudonymized relation store as described above, and refreshes PTC_j at any device where it is suspected that the key material is exposed. The period during which stale key material remains usable is bounded by the rotation interval and the configured grace window for key version rollover.

5. Evaluation

This section evaluates the proposed scheme from two perspectives: Sections 5.1 and 5.2 establish the security properties of the cryptographic authentication layer through BAN-logic derivation and AVISPA model checking, supplemented by informal arguments for the IIoT-specific extensions; Sections 5.3–5.6 then assess computation, communication, and storage overhead, bypass deployment latency, synthetic sensitivity of the auxiliary risk evaluation decision rule, and a positioning comparison against standardized lightweight protocols.

To avoid overgeneralizing the evaluation results, Table 4 explicitly separates the claims supported by this paper from those that require future deployment-stage validation.

Table 4. Evaluation claims, metrics, evidence, and claim boundaries.

Claim	Metric/Evidence	Evaluation Type	Supported Statement	Not Claimed
Cryptographic authentication and key agreement	BAN logic and AVISPA/HLPSL	Symbolic verification	Mutual authentication and replay/MITM resistance under the stated Dolev–Yao model	Computational proof, side-channel resistance, implementation-level security
Lightweight endpoint overhead	Operation counts; hash/XOR/sym. timing basis	Analytical overhead evaluation	Low endpoint cost under the selected operation-count basis	Measured PLC/RTU latency, measured energy, real network throughput
Auxiliary risk-evaluation behavior	Public benchmark metrics, synthetic Monte Carlo sensitivity, threshold/weight sweep	Architectural sensitivity analysis	Parameter trends and non-brittleness under stated synthetic distributions	Real OT coercion-detection accuracy, plant-scale FPR/FNR
Standardized protocol positioning	Published EDHOC/DTLS/Noise metrics and functional comparison	Positioning comparison	Functional scope and qualitative deployment trade-offs	Same-hardware benchmark or superiority ranking
Scalability discussion	Analytical service-time estimate and architecture-level bottleneck analysis	Analytical discussion	Gateway-side bottlenecks and deployment considerations	High-load measured concurrency under OT traffic

Accordingly, the evaluation should be read as a protocol-and-architecture assessment with reproducible analytical and sensitivity components, not as a full hardware-in-the-loop OT deployment study.

5.1. BAN Logic and AVISPA Verification

We first analyze the proposed scheme using BAN logic to verify that mutual authentication and session key agreement hold between U_i and FD_j.

Notation. We use standard BAN: $P|\equiv X$ "P believes X"; $P \triangleleft X$ "P sees X"; $P|\sim X$ "P once said X"; $\#(X)$ "X is fresh"; $\{X\}_K$ "X encrypted under K"; $P \overset{K}{\leftrightarrow} Q$ "K shared between P and Q"; and $P \Rightarrow X$ "P has jurisdiction over X". The six rules used below are receiving (R1), freshness propagation (R2), message meaning (R3), nonce verification (R4), session key (R5), and jurisdiction (R6: $P|\equiv Q \Rightarrow X$ and $P|\equiv Q|\equiv X \vdash P|\equiv X$).

Security goals. The scheme should achieve

$$
\begin{aligned}
&G_1 : FD_j|\equiv U_i \overset{KEY_{ij}}{\leftrightarrow} FD_j,\\
&G_2 : U_i|\equiv U_i \overset{KEY_{ij}}{\leftrightarrow} FD_j,\\
&G_3 : FD_j|\equiv U_i|\equiv U_i \overset{KEY_{ij}}{\leftrightarrow} FD_j,\\
&G_4 : U_i|\equiv FD_j|\equiv U_i \overset{KEY_{ij}}{\leftrightarrow} FD_j.
\end{aligned}
$$

Idealized protocol. The login-and-key-agreement messages relevant to BAN analysis are

$$
\begin{aligned}
&M_1 : U_i \rightarrow GWN : \{U_i \overset{N_i}{\leftrightarrow} FD_j, T_1\}_{SK},\\
&M_2 : GWN \rightarrow FD_j : \{U_i \overset{N_i}{\leftrightarrow} FD_j, T_2\}_{TC_j},\\
&M_3 : FD_j \rightarrow GWN : \{U_i \overset{N_i}{\leftrightarrow} FD_j, T_3\}_{TC_j},\\
&M_4 : GWN \rightarrow U_i : \{U_i \overset{N_i}{\leftrightarrow} FD_j, T_4\}_{SK}.
\end{aligned}
$$

Initial assumptions. The analysis adopts

$$
\begin{aligned}
&A_1 : U_i|\equiv \#(T_1, T_4),\\
&A_2 : U_i|\equiv U_i \overset{SK}{\leftrightarrow} GWN,\\
&A_3 : GWN|\equiv \#(T_1, T_2, T_3, T_4),\\
&A_4 : GWN|\equiv GWN \overset{SK}{\leftrightarrow} U_i,\\
&A_5 : GWN|\equiv GWN \overset{TC_j}{\leftrightarrow} FD_j,\\
&A_6 : FD_j|\equiv \#(T_2, T_3),\\
&A_7 : FD_j|\equiv FD_j \overset{TC_j}{\leftrightarrow} GWN.
\end{aligned}
$$

A_1, A_3, and A_6 follow from timestamp checks; A_2 and A_4 from the authenticated SK establishment of Section 4.5; and A_5 and A_7 from device registration that binds TC_j to the enclave-protected $\{PID_j, PR_j, SecLevel_j\}$ tuple. We additionally adopt two jurisdiction assumptions reflecting the trusted relay role of the GWN inside its TEE-protected enclave:

$$
\begin{aligned}
&A_8 : FD_j|\equiv \big(GWN \implies U_i \overset{N_i}{\leftrightarrow} FD_j\big),\\
&A_9 : U_i|\equiv \big(GWN \implies U_i \overset{N_i}{\leftrightarrow} FD_j\big),
\end{aligned}
$$

where FD_j trusts the GWN's authority over N_i (the GWN is the only entity that holds both SK and TC_j inside its enclave and can faithfully relay N_i) and symmetrically for U_i over N_j. These assumptions are limited to the protocol relay role and collapse if enclave attestation fails, which is excluded by the threat model of Section 3.2.

Derivation. From M_1 via R1, R3 (with A_4), R2 (with A_3), and R4:

$$S_1 : GWN \triangleleft \{U_i \overset{N_i}{\leftrightarrow} FD_j, T_1\}_{SK},$$
$$S_2 : GWN|\equiv U_i|\sim U_i \overset{N_i}{\leftrightarrow} FD_j,$$
$$S_3 : GWN|\equiv \#(U_i \overset{N_i}{\leftrightarrow} FD_j, T_1),$$
$$S_4 : GWN|\equiv U_i|\equiv U_i \overset{N_i}{\leftrightarrow} FD_j.$$

Symmetrically, M_2 with A_7 (then A_6, R3, R2, R4) yields

$$S_5 : FD_j \triangleleft \{U_i \overset{N_i}{\leftrightarrow} FD_j, T_2\}_{TC_j},$$
$$S_6 : FD_j|\equiv GWN|\sim U_i \overset{N_i}{\leftrightarrow} FD_j,$$
$$S_7 : FD_j|\equiv \#(U_i \overset{N_i}{\leftrightarrow} FD_j, T_2),$$
$$S_8 : FD_j|\equiv GWN|\equiv U_i \overset{N_i}{\leftrightarrow} FD_j.$$

For the device contribution, M_3 with A_5 and M_4 with A_2 deliver via the same R2/R3/R4 chain:

$$S_9 : GWN|\equiv FD_j|\sim U_i \overset{N_i}{\leftrightarrow} FD_j,$$
$$S_{10} : GWN|\equiv FD_j|\equiv U_i \overset{N_i}{\leftrightarrow} FD_j,$$
$$S_{11} : U_i|\equiv GWN|\equiv U_i \overset{N_i}{\leftrightarrow} FD_j.$$

Applying the jurisdiction rule R6 to S_8 with A_8 and to S_{11} with A_9 yields the direct beliefs

$$S_{8'} : FD_j|\equiv U_i \overset{N_i}{\leftrightarrow} FD_j,$$
$$S_{11'} : U_i|\equiv U_i \overset{N_i}{\leftrightarrow} FD_j.$$

Since $KEY_{ij} = h(N_i \| N_j)$, FD_j holds N_i via $S_{8'}$ and generates N_j locally, while U_i generates N_i locally and holds N_j via $S_{11'}$. Assuming that h is a random oracle, meaning that KEY_{ij} inherits the freshness and mutual authentication properties of (N_i, N_j), R5 under freshness A_1 and A_6 yields

$$G_1 : FD_j|\equiv U_i \overset{KEY_{ij}}{\leftrightarrow} FD_j,$$
$$G_2 : U_i|\equiv U_i \overset{KEY_{ij}}{\leftrightarrow} FD_j,$$
$$G_3 : FD_j|\equiv U_i|\equiv U_i \overset{KEY_{ij}}{\leftrightarrow} FD_j,$$
$$G_4 : U_i|\equiv FD_j|\equiv U_i \overset{KEY_{ij}}{\leftrightarrow} FD_j.$$

Therefore, all four security goals are established.

The BAN proof is intentionally limited to the cryptographic authentication path. It does not prove that the auxiliary risk evaluation module detects every coerced login; instead, the risk-evaluation module is analyzed as an auxiliary trust signal that raises the cost of using valid credentials under abnormal conditions. This separation avoids conflating symbolic protocol correctness with statistical anomaly detection performance.

The cryptographic protocol skeleton was further specified in HLPSL and checked with AVISPA [64]. Since SATMC and TA4SP do not support XOR operations in this setting, the OFMC and CL-AtSe backends were used. Both backends returned SAFE under

the Dolev–Yao model (Figure 3), supporting resistance to replay and man-in-the-middle attacks at the cryptographic protocol layer. The IIoT-specific extensions, namely, zero-trust contextual checks, auxiliary bypass risk evaluation, and break-the-glass emergency access, operate above the verified cryptographic skeleton and are analyzed informally in Section 5.2.

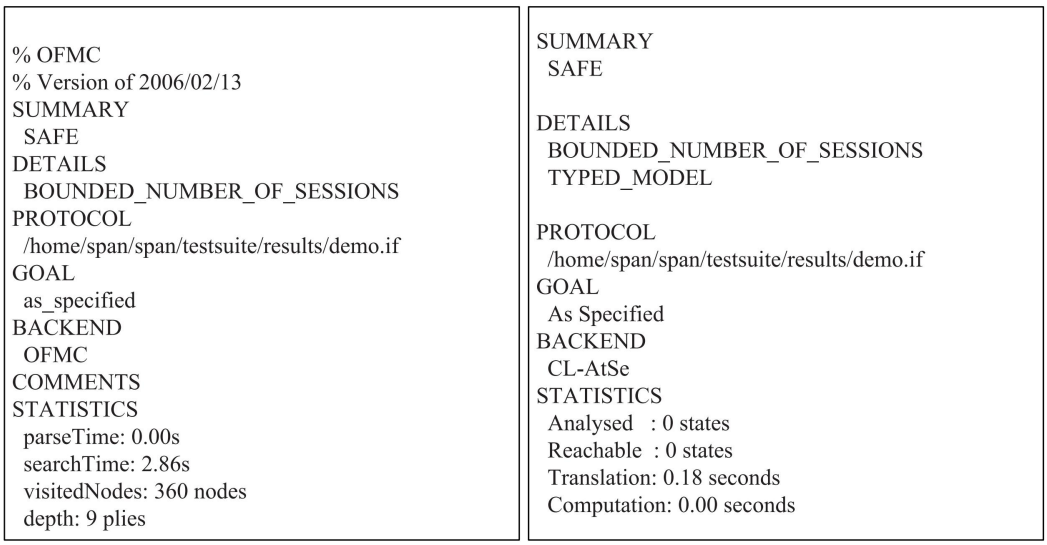

Figure 3. OFMC simulation and CL-AtSe simulation results.

The AVISPA model abstracts implementation-level details such as camera capture, physiological sensing, and TEE memory layout. This abstraction is appropriate because AVISPA is used to verify the message exchange against network-level attacks, not to validate the statistical detector. Therefore, the informal analysis below covers the additional insider, emergency, and coercion cases that are outside the symbolic model.

5.2. Informal Security Arguments

Insider and malware resistance. Authentication relation tables and operator–asset mappings are stored as pseudonymized tuples inside the secGear enclave; enclave memory and K_{GWN} are not exposed even to a gateway administrator with OS privileges, and plaintext ID_i and PW_i are never stored at the GWN. The user identity is transformed as $UID_i = h(ID_i \| PW_i)$ and $PID_i = h(UID_i \| K_{GWN})$, so an eavesdropper cannot derive the real identity or link sessions through stable plaintext identifiers.

Replay, tampering, and freshness. Control commands are encrypted under $KEY_{ij} = h(N_i \| N_j)$ with session-specific N_i, N_j that the network attacker cannot compute. Receivers check timestamp freshness prior to expensive processing, so stale messages are rejected before enclave lookup or device-side computation; modified timestamps fail after decryption or MAC verification. Each session uses fresh nonces, giving per-session key freshness and session key independence under standard one-way function assumptions. These checks limit but do not eliminate flooding attacks.

Coercion and emergency abuse. A coerced-login attacker may obtain valid factors by forcing the operator's participation. The proposed scheme does not treat anomaly detection as a cryptographic defense against such attacks; instead, the auxiliary risk evaluation module provides an additional zero-trust signal that can trigger re-verification, privilege degradation, audit escalation, or session termination before high-risk control privileges are released. *EmergToken* usage is gated by three-factor verification on a dedicated emergency SC; identity and operation privilege checks remain enforced even in emergency mode, and all emergency actions are logged for audit. We state the inheritance precisely: the cryp-

tographic authentication used by the emergency path inherits the Layer-1 BAN/AVISPA analysis, whereas abuse resistance of the emergency *workflow* is established through the structured abuse-case analysis of Table 5, not through a mechanized proof. Break-the-glass bypasses only the environment and behavior checks; identity, role, and device operation privilege checks remain fully enforced, so the only relaxation is contextual.

Table 5. Break-the-glass abuse cases and containment controls.

Abuse Case	Risk	Containment Mechanism	Residual Limitation
Stolen emergency smart card	Unauthorized emergency attempt	Three-factor authentication on the dedicated card; *EmergToken* expiry and one-time invalidation	Attacker must also defeat the remaining factors, but a narrow exposure window remains until expiry
Shared emergency code	Collusive or careless credential sharing	Emergency credentials are bound to an identified operator and logged in *EmergLog*	The scheme deters and audits sharing; it cannot technically prevent pre-use disclosure of a code
Repeated emergency use	Normalization of emergency mode	Every activation is non-repudiably logged and repeated patterns are flagged for security-officer review	Controls are strongest as post-event deterrence, not as a formal proof against first misuse
Emergency use outside shift	Suspicious off-duty activation	Time, role, and device-operation checks remain available for alerting and audit escalation	A true emergency may still require off-shift access, so the design avoids automatic denial solely on time
Low-risk non-urgent operation	Overuse of emergency path for convenience	Break-the-glass relaxes selected environment/behavior checks only; identity, role, and operation privileges remain enforced	Misclassification of urgency is handled by audit review rather than mechanized proof
Post-event log deletion attempt	Covering tracks after misuse	*EmergLog* is enclave-protected and exported to audit channels	Availability of external audit storage is a deployment assumption
Activation under coercion	Forced use of valid factors	Facial channel can remain active and raise high-priority notification when $\theta_{high}+\Delta\theta$ is exceeded	Alert-not-block by design; coercion detection is not guaranteed

Break-the-glass does not bypass identity verification, device operation authorization, or audit logging; it relaxes selected environment and behavior checks under emergency conditions only, meaning that it is an emergency-controlled degradation mode rather than an unauthenticated bypass. The cryptographic authentication of the emergency path inherits the Layer-1 BAN/AVISPA analysis; a fully mechanized model of the emergency workflow remains as future work.

Limits of auxiliary risk evaluation. The auxiliary risk evaluation module provides anomaly evidence, not a mathematical guarantee. A highly controlled coercion scenario could in principle reduce all modality scores below the rejection threshold; therefore, the design relies on a re-verification region, privilege degradation, and logging of modality availability, scores, and emergency mode use. The security claim is confined to *increasing* the opportunity to detect or audit abnormal access relative to credential-only authentication, not to proving complete coercion prevention. Even a fully suppressed auxiliary verdict cannot grant access: by the monotonicity invariant of Section 4.3, the module can only withhold or downgrade privileges, so its worst case degrades to credential-plus-deterministic-control security, the same baseline as a scheme without it, and does not open up a new attack path. The same property bounds the impact of presentation/spoofing attacks against the sensing front-end (Section 6).

Table 6 summarizes the comparison between the proposed scheme and representative resource-constrained IoT authentication schemes. To avoid conflating formally verifiable protocol properties with deployment-oriented zero-trust capabilities, the comparison is

divided into two groups: the first reports cryptographic and protocol-layer security properties, whereas the second reports architectural and operational properties that are outside the design scope of conventional AKE schemes.

Table 6. Comparison of cryptographic and zero-trust operational properties.

Property	[7]	[8]	[9]	[10]	Ours
Cryptographic and protocol-layer security properties					
Anonymity/untraceability	Y	Y	Y	Y	Y
Replay attack resistance	Y	Y	Y	Y	Y
Man-in-the-middle attack resistance	Y	Y	Y	Y	Y
Computation-oriented DoS mitigation	–	Y	–	P	Y
Formal verification	Y	Y	Y	Y	Y
Zero-trust and operational access-control properties					
Privileged-insider mitigation	N	N	N	N	Y
Gateway-side malware mitigation	N	N	N	N	Y
Asset mapping confidentiality	P	P	P	P	Y
Role-based access control	N	N	N	N	Y
Device-level authorization	N	N	N	N	Y
Environment-aware access control	N	N	N	N	Y
Behavior-aware access control	N	N	N	N	Y
Hardware-based relation protection	N	N	N	N	Y
Break-the-glass emergency access	N	N	N	N	Y
Auxiliary risk evaluation *	N	N	N	N	Y
Coerced-login risk-response support *	N	N	N	N	Y

Y: supported; N: not supported; P: partial; –: not discussed. The first group concerns cryptographic or protocol-layer properties. The second group concerns functional zero-trust capabilities, and as such is a deployment-oriented comparison rather than a purely cryptographic one. * The auxiliary risk evaluation module provides optional anomaly evidence for re-verification, privilege degradation, or audit escalation; it is not part of the formal cryptographic security proof and does not constitute a mathematical guarantee of coercion detection.

The preceding security analysis establishes the intended authentication and key agreement properties; the remainder of this section evaluates overhead and positions the scheme against existing protocols. The evaluation has four parts, each of which establishes a different kind of claim. The BAN logic and AVISPA analyses (Sections 5.1 and 5.2) establish symbolic authentication and key agreement correctness and resistance against replay and man-in-the-middle attacks; the overhead analysis below establishes computation, communication, and storage *complexity*; the sensitivity analysis of Section 5.5 is an architectural sanity check of the auxiliary decision behavior, not an operational detection performance claim; and the comparison with standardized lightweight protocols (Section 5.6) is a positioning comparison rather than a same-hardware benchmark. The overhead figures are analytical operation counts, and are consequently deterministic; therefore, a sampling-based confidence interval does not apply to them, and the table legends state this explicitly.

5.3. Computation, Communication, and Storage

Following prior resource-constrained IoT AKE evaluations [65,66], we assume mobile operator terminals and MICAz-class constrained devices (Table 7; basis details in Appendix A). To provide a like-for-like comparison, all totals in Table 8 are recomputed under this single basis, and each compared scheme is also evaluated under a uniform mobile basis in Appendix A. The proposed scheme totals 87.3 ms, the lowest among the compared schemes, and its constrained device-side cost is also the lowest at 24 ms, against 31–4932 ms for the others; the device-side figure is the one that matters most for battery-powered field endpoints. This difference comes from where public key operations are placed. A single T_{ecm} costs several orders of magnitude more on a MICAz-class processor than on a mobile

CPU, so the ECC-based schemes of Li et al. [7] and Xiong et al. [8], which evaluate point multiplications on the sensor or gateway side, reach thousands of milliseconds once costed at constrained-side speed. This cost is hidden when the source papers report totals under a uniform mobile basis, but it dominates in IIoT deployments with battery-powered sensor nodes and field devices. The hash-and-XOR schemes of Das et al. [9] and Vinoth et al. [10] avoid this, and are the closest competitors; our scheme still attains a lower device-side cost, and additionally supplies the zero-trust control points, confidential computing protection, and personnel-state evaluation that none of the compared schemes provide. It places no public key operation on the constrained endpoints, trading a slightly higher hash count on the mains-powered gateway for lower cost on the field device. The MICAz figure is a conservative and dated baseline for the constrained side; on a modern Cortex-M-class microcontroller with a hardware ECC accelerator, a point multiplication costs on the order of tens of milliseconds rather than 2450 ms, which narrows the gap for the ECC-based schemes. The architectural conclusion does not depend on the platform, since the scheme performs *no* public key operation on the field endpoint; Appendix A confirms that the ordering holds under a uniform-mobile basis and under a representative Cortex-M4 with hardware ECC basis.

Table 7. Computation time of basic operations.

Notation	Operation	U (ms)	GWN/FD (ms)
T_H	Hash	0.03	8
T_S	Symmetric encryption	0.12	3.5
T_{SM2}^{sv}	SM2 sign or verify	≈2.5	≈0.6
T_{SM2}^{ed}	SM2 encrypt or decrypt	≈2.5	≈0.6
T_{ecm}	ECC point multiplication	20.23	2450
T_{fe}	Fuzzy extractor	20.23	–

Per-entity hardware assumptions and the sense in which the GWN/FD column is conservative are detailed in Appendix A.

Table 8. Computation overhead in authentication, recomputed under the Table 7 basis.

Scheme	U	GWN	FD	Total
[10]	$T_{fe} + 8T_H + T_S$	$6T_H + 3T_S$	$3T_H + 2T_S$	110.09 ms
[9]	$T_{fe} + 14T_H$	$9T_H$	$7T_H$	148.65 ms
[7]	$T_{fe} + 8T_H + 3T_{ecm}$	$7T_H + T_{ecm}$	$4T_H + 2T_{ecm}$	7519.16 ms
[8]	$5T_H + 3T_{ecm}$	$11T_H + 6T_{ecm}$	$4T_H$	14,880.84 ms
Ours	$2T_H + 2T_S$	$7T_H + 2T_S$	$3T_H$	87.30 ms

All totals are recomputed under the operation-time basis of Table 7 for direct comparability across schemes. Per-scheme operation counts, evaluation under a uniform mobile basis, and entity-mapping conventions are detailed in Appendix A.

The communication overhead is 1216 bits, the lowest among the compared schemes (Das et al. [9] 1536 bits, Xiong et al. [8] 2208 bits, Li et al. [7] 2720 bits, and Vinoth et al. [10] 3040 bits; Table 9).

The proposed scheme achieves the lowest total because it places no public key operation on the constrained endpoint; its field device cost is $3T_H = 24$ ms, which is dominated by inexpensive hash operations. By contrast, schemes that evaluate ECC point multiplications on the sensor side (Li et al. [7], two per session, $2 \times 2450 = 4900$ ms device-side) see their totals rise by orders of magnitude under the MICAz basis. The hash-and-XOR schemes (Das [9], Vinoth [10]) avoid this penalty and are the closest competitors, but their totals remain higher because they evaluate more hash operations at the constrained side.

Table 9. Communication and storage overhead.

Scheme	Communication (bits)	Storage (bits)
Das [9]	1536	– [a]
Xiong [8]	2208	≈1376 [b]
Li [7]	2720	– [a]
Vinoth [10]	3040	≈1952 [b]
Ours	1216	1152

[a] Source paper not available to the authors; the persistent credential set could not be enumerated, so storage is not reported. [b] Estimated by enumerating the persistent smart card and device credentials under the reference field-size convention of Appendix A; to be confirmed against the source.

The scheme storage overhead is 1152 bits across U, GWN, and FD; where the compared schemes' storage could be determined from their published credential sets, it is higher (Table 9). The auxiliary risk evaluation model parameters reside on the policy engine server, and are not stored on field devices or sensor nodes; therefore, the auxiliary risk evaluation module does not increase the storage burden on resource-constrained field devices.

The 87.3 ms in Table 8 is the per-login cost of login and key agreement. The SM2-based channel establishment of Section 4.4 is a one-time setup at the start of an operating shift, contributing one SM2 verify-and-encrypt on the terminal side (≈5 ms) and one SM2 sign-and-decrypt on the GWN side (≈1.2 ms); when amortized over a multi-hour shift, the per-login share is well under 1 ms; therefore, it is reported separately rather than added into the per-login total.

5.4. Bypass Deployment Latency Analysis

The auxiliary risk evaluation module is evaluated as a bypass component on the policy engine server. The FER, LSTM, and 1D-CNN modules require approximately 12 ms, 3 ms, and 2 ms per inference, respectively (Section 4.2), and the late-fusion layer requires less than 1 ms. Therefore, the total auxiliary risk evaluation latency is below 18 ms.

Since the primary authentication path requires approximately 87.3 ms (Table 8), an auxiliary verdict that completes within the benchmark-level 18 ms estimate is hidden behind the cryptographic exchange for the normal case: $\max(87.3, 18) = 87.3$ ms. A hypothetical inline deployment serializing detection after the protocol would require $87.3 + 18 = 105.3$ ms. The dual-modality bypass case (no wearable, FER + behavior $\approx$ 15 ms) likewise yields $\max(87.3, 15) = 87.3$ ms. Therefore, the neural model is not made free, but its inference is no longer serialized after the cryptographic protocol. Release of high-risk control privileges is still governed by the commit-point gate of Algorithm 1: if the auxiliary verdict is late, high-*SecLevel* operations are held or denied pending the verdict, while low-*SecLevel* operations may receive only monitored read-only access according to plant policy. The reported auxiliary risk evaluation latency is based on benchmark-level inference estimates rather than a full plant deployment; end-to-end measurement on calibrated edge hardware together with empirical evaluation under realistic concurrent load is identified as deployment-stage future work.

5.5. Synthetic Sensitivity Analysis of Auxiliary Risk Evaluation

Scope of the analysis. The following sensitivity study is an *architectural design-space check*: it characterizes how the auxiliary decision behaves as its parameters vary under stated synthetic distributions. It is not an operational detection performance claim, and does not replace the site-specific empirical calibration identified as future work.

Table 10 lists literature-level reference metrics for related modality tasks. These values provide context for model selection only; they were obtained on different datasets, tasks, metrics, and operating points, and as such cannot be converted into false negative rates for

this architecture or multiplied into a joint system miss probability. For the facial modality, AffectNet provides the source benchmark for the MobileNetV2-based FER choice [16,57]. For the physiological modality, WESAD provides the wearable stress-sensing task context [18]. For the behavioral modality, the reconstruction error detector in [36] provides an algorithmic analogue rather than a login history validation dataset. Appendix A states these transfer limits explicitly.

Table 10. Per-modality literature reference metrics used for model selection context; these are not system-level FPR/FNR estimates.

Modality	Source Benchmark	Metric	Value
FER (F)	AffectNet-8 [16]	Binary F1	≈0.72
Behavior (B)	wearable AE [36]	AUC	≈0.89
Physio. (P)	WESAD-style [18]	Binary F1	≈0.95

The rows are not directly comparable: F_1 and AUC are different quantities, and none of the rows supplies a calibrated false negative rate for the proposed deployment. They are retained only as literature context for the replaceable modality interfaces.

Sensitivity analysis of weights, thresholds, and cross-modal dependence. We run a scripted Monte-Carlo study over documented synthetic per-modality score distributions (benign $e, b, p \sim \text{Beta}(1.5, 6)$, mean ≈ 0.20; high-risk synthetic population $e, b, p \sim \text{Beta}(4, 3)$, mean ≈ 0.57; $N = 4 \times 10^5$ draws; fixed seed). The weight panels use independent modality draws. To avoid treating independence as a deployment assumption, the threshold analysis also uses a common-factor stress model $s_k^{(\gamma)} = (1-\gamma)s_k + \gamma s_{shared}$ for $\gamma \in \{0, 0.3, 0.6, 0.9\}$, where larger γ makes the three modality scores move together. This mixing coefficient is a sensitivity device, not an empirically estimated correlation parameter. Full assumptions, sweep ranges, metric definitions, and the script are provided in the Supplementary Materials and summarized in Appendix A.

Figure 4a sweeps the physiological weight $\alpha_P \in [0, 0.5]$ with $\alpha_F = \alpha_B = (1-\alpha_P)/2$ at the default thresholds under independent synthetic scores; the decision rates vary smoothly around the illustrative default $\alpha_P = 0.2$. Figure 4b maps $P(\text{high-risk synthetic login not rejected})$ over the weight simplex, locating the default split within a smooth region rather than at a cliff. Figure 5a sweeps θ_{high} under independent and common-factor stress scenarios. Figure 5b reports the Accept/Re-verify/Reject composition at the default thresholds as γ increases. Under independent synthetic draws, approximately 2% of the high-risk synthetic population is fully Accepted, the majority falls into Re-verify, and the rest is Rejected; stronger common-factor mixing changes this composition, demonstrating that the result depends on the stated dependence scenario. Monte Carlo 95% half-widths are below approximately 0.1 percentage points for the reported default proportions, but these intervals quantify simulation error only, and must not be interpreted as confidence intervals for plant operation.

The weight sweep surface in Figure 4 is smooth around the illustrative default, while Figure 5 shows that the decision mix changes as cross-modal dependence is strengthened. Together these results support only a bounded design conclusion: the default is not located at a local weight cliff under the stated synthetic model, but its numerical decision rates are assumption-dependent and require site-side calibration.

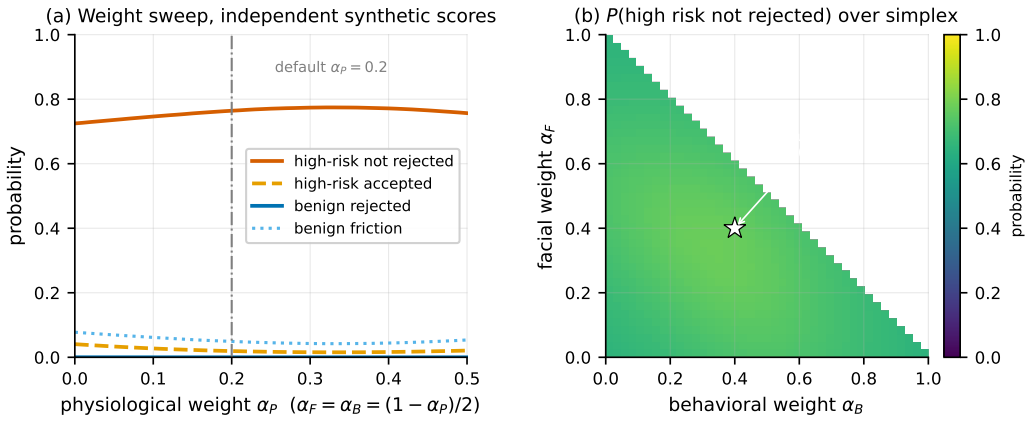

Figure 4. Weight sensitivity of the auxiliary decision under the independent synthetic model of Appendix A. (**a**) Decision rates as the physiological weight α_P is swept with $\alpha_F = \alpha_B = (1 - \alpha_P)/2$ at fixed thresholds $\theta_{low} = 0.35$, $\theta_{high} = 0.65$. (**b**) P(high-risk synthetic login not rejected) over the weight simplex $(\alpha_F, \alpha_B, \alpha_P = 1 - \alpha_F - \alpha_B)$; the illustrative default $(0.4, 0.4, 0.2)$ is marked by a star. This is an architectural sensitivity check, not an operational detection claim.

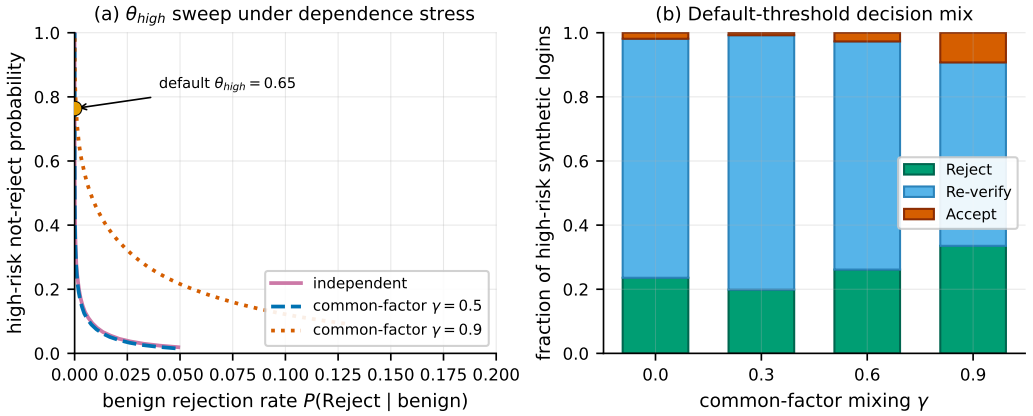

Figure 5. Threshold and dependence sensitivity under the synthetic model of Appendix A. (**a**) Operating characteristics as θ_{high} is swept ($\theta_{low} = 0.35$ fixed) under independent scores and two common-factor stress scenarios. (**b**) Decision composition of the high-risk synthetic population at the default thresholds for common-factor mixing $\gamma \in \{0, 0.3, 0.6, 0.9\}$. The mixing scenarios test dependence sensitivity and are not calibrated plant correlations.

5.6. Comparison with Standardized Lightweight Protocols

To position the proposed scheme within the standardized landscape, Table 11 compares it with EDHOC, DTLS 1.3, and Noise along two axes. This is a *positioning* comparison, not a same-hardware, same-implementation, same-network head-to-head benchmark: the AKE sublayer figures for the reference protocols are drawn from published constrained device measurements [26,27], each under its own platform and configuration, whereas our figure is the analytical operation-count basis of Table 8. The first group (AKE sublayer) shows that all three reference protocols perform public key operations on the constrained endpoint, while the proposed scheme performs none there; the second group (functional scope) shows that the reference protocols are channel establishment mechanisms that do not, on their own, provide TEE-protected gateway relations, role/device/environment/behavior-aware release, personnel-state evidence, or a break-the-glass workflow. Therefore, the comparison is not a speed ranking: EDHOC, DTLS 1.3, and Noise bring mature implementations, standardized authentication semantics, and interoperability that the proposed scheme does not claim to supersede, and the zero-trust layer could in principle be layered over any of them.

Table 11. Positioning comparison with standardized lightweight secure-channel protocols. AKE sublayer costs for the reference protocols are from published constrained-device measurements, and are not a same-hardware benchmark; our cost is the analytical basis of Table 8.

Property	EDHOC	DTLS 1.3	Noise	Ours
AKE/handshake sub-layer				
Key agreement primitive	ECDH/COSE	TLS 1.3 KE	DH patterns	SM2 setup + hash/sym.
Public-key op. on constrained endpoint	Yes	Yes	Yes	**No**
Constrained device cost source	[26]	[27]	pattern-dep.	Table 8 (analyt.)
Functional scope (deployment-oriented)				
TEE-protected gateway relations	N	N	N	Y
Role/device-aware control points	N	N	N	Y
Env./behavior-aware release	N	N	N	Y
Personnel-state auxiliary risk evidence	N	N	N	Y
Break-the-glass emergency access	N	N	N	Y
Pseudonymous op.–asset mapping	N	N	N	Y

Y: provided; N: not provided by the protocol itself. EDHOC commonly bootstraps an OSCORE context [21]; EDHOC/OSCORE uses compact CBOR/COSE structures, whereas DTLS 1.3 and Noise use their own native encodings [24,25]. Reported costs are on heterogeneous platforms and are not directly commensurable; the table positions the schemes, it does not rank their speed.

6. Discussion

Industrial workflow integration. The same scheme degrades gracefully across operating regimes. In routine control room monitoring, a control engineer's SC verifies $ID_j \in FDLIST_i$ at the appropriate $SecLevel_j$ and the auxiliary risk evaluation module runs in bypass mode against a stable behavioral baseline; user-perceived latency is dominated by the 87.3 ms cryptographic path. When a field operator reaches a high-risk safety-instrumented actuator requiring $SecLevel_j = 3$, the SC's local role check rejects the request before any network round-trip, consistent with the least-privilege principle of IEC 62443 [67]. During an emergency shutdown, an on-call engineer obtains time-bounded access via the break-the-glass channel; environment and behavior checks are bypassed, role and operation privileges remain enforced, *EmergLog* records the activation, and the security officer is notified if the facial channel returns an extreme score.

Scalability and concurrency. Under a serial analytical service-time approximation, the 87.3 ms per-login cryptographic cost corresponds to an upper-bound service rate of approximately 11.5 authentications/s before network I/O, enclave contention, queuing delay, and policy engine load are considered. Therefore, this number should be interpreted as an analytical capacity indicator rather than a measured concurrent throughput result. On the constrained device side, cost is independent of the number of registered users or devices because no public key operation resides at the field endpoint; therefore, the deployment bottleneck is expected at the gateway and access-proxy layer. Empirical characterization of this bottleneck under realistic OT traffic remains deployment-stage future work.

Compliance with industrial security regulation. The pseudonymization layer (PID_i, PID_j, $PID_i^{(s)}$) plus enclave-bound mappings align with the access control and auditability requirements of IEC 62443 [67] and the OT security guidance of NIST SP 800-82 [68]: a host-OS dump never reveals which operator interacts with which asset, and each *EmergLog* entry provides the non-repudiable audit trail expected for privileged actions in industrial control systems. Facial and physiological processing may constitute special category data (GDPR Article 9 [69]), and must rest on a locally valid legal basis rather than on employee consent alone because consent in an employment relationship may not be freely given. They also require explicit transparency, worker consultation where applicable, a data protection impact assessment [70], data minimization, retention control, and an opt-out or non-biometric

fallback path where required. The fallback is consistent with the auxiliary risk evaluation module being optional; an operator who opts out is evaluated on the deterministic control points alone. Guidance specific to camera-based processing further constrains transparency and minimization [71]. Data minimization is already enforced: facial frames are confined to a sliding five-frame window processed locally and discarded once the trust score is emitted, and selective disclosure of attestation claims (Section 3) limits what leaves the enclave. Table 12 summarizes the corresponding privacy-preserving design choices.

Table 12. Privacy-preserving design choices for auxiliary risk evaluation.

Privacy Concern	Design Constraint	Implementation Guidance
Facial data	Minimize raw biometric retention	Process short local windows and discard frames after score emission unless a site-approved incident policy requires preservation
Physiological data	Avoid unnecessary health inference	Use only stress risk features needed for login-time scoring; record modality availability and score by default, not raw wearable streams
Operator profiling	Prevent open-ended surveillance	Restrict models to authentication-time risk evaluation and document the site-side calibration protocol
Audit log	Balance accountability and minimization	Store decision, score breakdown, operator pseudonym, and emergency use metadata inside protected audit channels
Opt-out/unavailable modality	Preserve access without biometrics where required	Disable the auxiliary channel or use non-biometric fallback checks without affecting the cryptographic AKE layer
Data retention	Limit long-term exposure	Define retention windows, deletion procedures, and access controls in the plant's data protection impact assessment

The auxiliary risk evaluation module can be disabled without affecting the cryptographic AKE layer. When enabled, it should process only the minimum features necessary for risk scoring, avoid raw biometric retention unless explicitly approved, and support non-biometric fallback modes where required by local regulation.

Interpretability and sustainability. The auxiliary risk evaluation module's late-fusion architecture is interpretable by construction: the breakdown $(\alpha_F e, \alpha_B b, \alpha_P p)$ accompanies every *EmergLog* alert, supporting operator-facing explanation and forensic review without exposing raw physiological signals. Architecturally, lightweight hash and symmetric operations reduce per-session energy on battery-powered field devices and sensor nodes; the modular risk evaluation interface allows modality models to be retrained or replaced without redesigning the cryptographic core or TEE-protected storage; and login-time risk evidence may reduce the downstream cost of incident response and recovery in industrial control systems [68].

Comprehensive risk evaluation. The combination of role/device/environment/behavior rules with auxiliary personnel-state signals yields a graduated permission model rather than a single accept/reject verdict. When the joint risk evaluation is favorable ($TrustScore < \theta_{low}$ and all hard constraints satisfied), the session is granted full read/write device control privileges. When the evaluation is ambiguous ($\theta_{low} \leq TrustScore < \theta_{high}$, or environment/behavior anomalies in the absence of acute personnel-state signals), only data-observation privileges are released and high-risk operations such as remote setpoint changes on a critical control loop, emergency shutdown overrides, or PLC logic downloads are withheld pending re-verification. When the evaluation is unfavorable, the session is terminated. This graduation directly enforces the operational safety asymmetry that data observation is reversible

whereas remote actuation is not, and operationalizes the multi-dimensional state-aware design articulated in Section 1. Table 13 summarizes the scope and mitigation boundaries for attacks on auxiliary risk evaluation inputs.

Table 13. Scope and mitigation boundaries for attacks on auxiliary risk evaluation inputs.

Attack/Failure Mode	Impact	Existing Boundary	Required Deployment Mitigation
Facial photo replay	Lowers apparent facial risk score	Cannot create cryptographic authentication or grant denied privileges	Camera liveness detection and supervised enrollment controls
Deepfake/adversarial face	Suppresses or manipulates facial evidence	Worst case degrades to deterministic authentication and access control	Presentation-attack detection, camera hardening, adversarial testing
Wearable spoofing	Manipulates physiological score or signal availability	Physiological channel is optional and lower-weighted by design	Device pairing, signal-quality checks, and tamper-resistant wearable management
Sensor unavailable	Removes one modality from fusion	Remaining modalities are renormalized; deterministic checks remain active	Site policy for fallback, re-verification, or conservative privilege release
PPE/mask/lighting degradation	Reduces facial reliability in plant conditions	Module is auxiliary and site-calibrated, not a hard gate	Calibration under site PPE/lighting and non-biometric fallback modes
Cross-modal suppression	Lowers several risk scores simultaneously	Common-factor sensitivity analysis tests dependence stress scenarios	Operator-in-the-loop validation and conservative thresholds for high-risk operations
Behavioral cold start	Insufficient history for anomaly baseline	Rule-based fallback dominates until warm-up is complete	Enrollment period, per-role priors, and explicit cold-start policy

Because the auxiliary risk evaluation module is monotonic and can only restrict privileges, successful spoofing of its inputs cannot create a cryptographic authentication result. In the worst case, the system degrades to credential-based authentication plus deterministic zero-trust control points.

Limitations. End-to-end empirical validation of auxiliary risk evaluation requires site-specific coercion datasets that are not publicly available; calibration of θ_{low} and θ_{high} via controlled stress induction protocols and operator-in-the-loop studies is the primary direction for future work. The modalities also degrade under realistic plant conditions: facial cues under personal protective equipment, masks, or deliberate suppression; behavioral cues when an attacker mimics the routine; and physiological capture under motion and poor sensor contact. For this reason, the module is auxiliary and site-calibrated, not a hard gate. Presentation and spoofing attacks (photographs, replay, deepfakes, coordinated cross-modal suppression) are likewise outside the module's claimed scope; liveness/anti-spoofing at the sensing front-end is future work, and by the monotonicity invariant (Section 4.3), a spoofed "benign" reading degrades to credential-plus-deterministic control security and does not create a new attack path. On the cryptographic side, the BAN/AVISPA analysis is symbolic and does not by itself establish computational- or implementation-level security; the TEE is treated as a trusted hardware root, and documented side-channel attacks against SGX and TrustZone variants [48–50] imply that frequent re-attestation and side channel-resistant coding belong in a full deployment threat model. Network-level evaluation under realistic OT traffic (loss, jitter, small MTU, clock drift, restart behavior, and brownfield gateways) and a hardware-in-the-loop campaign across multiple MCU/TEE platforms are deferred to deployment-stage validation. In terms of priority, site-side calibration of the auxiliary risk evaluation thresholds via controlled stress-induction protocols and operator-in-the-loop studies is the most pressing next step, followed by edge-side model compression for lower-power gateway platforms, a multi-platform hardware-in-the-loop measurement campaign, and longitudinal studies on the energy footprint of the scheme in operational OT networks.

7. Conclusions

This paper has presented a lightweight zero-trust authentication and key agreement scheme for IIoT. The scheme embeds role-, device-, environment-, and behavior-aware control points into the authentication process, protects gateway-side relations with secGear-based confidential computing, and uses an optional auxiliary risk evaluation module to provide zero-trust evidence for reverification, privilege degradation, and audit escalation without inflating the cryptographic critical-path latency. BAN-logic and AVISPA analyses support the intended security properties, and an overhead analysis under the analytical operation-count basis of this paper indicates that the field-endpoint cost remains dominated by hash and symmetric operations and lightweight relative to the selected representative IoT AKE schemes. In this way, the proposed scheme addresses the threefold gap identified in the introduction: gateway-side relations are protected from privileged insiders by TEE-confined storage, post-authentication access is gated by four deterministic zero-trust dimensions, and an optional auxiliary risk evaluation module provides a bounded response to the human-factor uncertainty that credential-only authentication leaves unresolved. Future work will focus on site-side empirical calibration of risk evaluation thresholds, edge-side model compression, integration with multimodal industrial large language models for audit explanation, and longitudinal studies on energy footprint in real OT networks.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/app16136765/s1, The supplementary material contains (i) the synthetic sensitivity-analysis scripts, fixed configuration, expected console summaries, and generated sensitivity figures; (ii) the analytical-overhead calculation script; (iii) the same-host Windows x64 benchmark source, scripts, raw data, summary tables, environment snapshot, third-party version metadata, and reproduction report; and (iv) the AVISPA HLPSL model source and README used in this revision.

Author Contributions: Conceptualization, X.Z. and Z.M.; methodology, X.Z. and Z.M.; software, X.Z.; validation, Z.M., D.M., and X.L.; formal analysis, X.Z.; investigation, X.Z.; writing—original draft preparation, X.Z.; writing—review and editing, Z.M., D.M., and X.L.; supervision, D.M. and X.L.; funding acquisition, Z.M. and D.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China, grant numbers 62372069 and 62272389. The APC was funded by the authors' institution.

Institutional Review Board Statement: Not applicable. This work proposes and analyzes an authentication architecture; the reported per-modality reference figures are drawn from previously published public benchmarks for model-selection context, and no new study involving humans or animals was conducted.

Informed Consent Statement: Not applicable.

Data Availability Statement: No new experimental datasets involving human subjects, industrial sites, or field-device measurements were created or analyzed. The data used for the sensitivity analysis are synthetic samples generated by the supplied script with a fixed random seed; the analytical overhead calculations and generated sensitivity figures are provided in the Supplementary Materials. Source code and supporting files for the reproducible analytical, sensitivity, same-host benchmark, and AVISPA-model components are included with the supplementary package.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

IIoT	Industrial Internet of Things
AKE	Authenticated Key Exchange
TEE	Trusted Execution Environment
SC	Smart Card
GWN	Gateway Node
FD	Field Device
RBAC	Role-Based Access Control
FER	Facial Expression Recognition
LSTM	Long Short-Term Memory
CNN	Convolutional Neural Network
ECC	Elliptic Curve Cryptography
HRV	Heart Rate Variability
EDA	Electrodermal Activity
PLC	Programmable Logic Controller
SIS	Safety-Instrumented System

Appendix A. Measurement Bases and Metric Conventions

Operation costs (Table 7). The U column is smartphone-class. The GWN/FD column lists the cost at whichever entity actually performs the operation: hash, symmetric encryption, and ECC point multiplication use MICAz-class speed (the conservative endpoint), bounding the total for Ours from above. The T_{fe} row is unfilled because no scheme invokes biometric reproduction at the constrained side. SM2 only appears at the GWN, so the 0.6 ms is the GWN's server-class cost.

Computation cost basis (Table 8). The totals reported in Table 8 are recomputed under the Table 7 basis to enable apples-to-apples comparison across schemes. Because each compared scheme originally reports its total under a different platform basis, this re-computation step is documented in full below.

Per-scheme operation counts. Each scheme's symbolic operation count, by entity, is: Li [7], $8T_H + 3T_{ecm} + T_{fe}$ (U), $7T_H + T_{ecm}$ (GWN), $4T_H + 2T_{ecm}$ (FD); Xiong [8], $5T_H + 3T_{ecm}$ (U), $11T_H + 6T_{ecm}$ (GWN), $4T_H$ (FD); Das [9], $14T_H + T_{fe}$ (U), $9T_H$ (GWN), $7T_H$ (FD); Vinoth [10], $8T_H + T_{fe} + T_S$ (U), $6T_H + 3T_S$ (GWN), $3T_H + 2T_S$ (FD). Table A1 evaluates each count under both a uniform mobile basis ($T_H \approx 0.32$, $T_{ecm} \approx T_{fe} \approx 17.1$, $T_S \approx 0.12$ ms) and the Table 7 MICAz basis.

Table A1. Computation totals of each scheme under three operation-time bases.

Scheme	Uniform Mobile Basis	Table 7 (MICAz) Basis	Cortex-M4 + HW-ECC †
Vinoth [10]	≈23.26 ms	110.09 ms	≈23.79 ms
Das [9]	≈26.70 ms	148.65 ms	≈25.45 ms
Li [7]	≈125.78 ms	7519.16 ms	≈174.46 ms
Xiong [8]	≈160.30 ms	14,880.84 ms	≈245.34 ms
Ours	≈4.32 ms	87.30 ms	≈3.50 ms

All three columns evaluate the same symbolic operation count of each scheme. The uniform mobile basis uses $T_H \approx 0.32$, $T_{ecm} \approx T_{fe} \approx 17.1$, $T_S \approx 0.12$ ms (smartphone-class, as commonly assumed in the source papers); the Table 7 basis places constrained-side operations at MICAz-class speed. † The third basis is a representative 80–100 MHz Cortex-M4 with a hardware ECC accelerator: constrained-side $T_H \approx 0.3$, $T_S \approx 0.1$, $T_{ecm} \approx 30$ ms, with the user side at the Table 7 U-column; it is included only to show that the ordering is robust to the platform choice, and is not a measured figure. Across all three bases, the proposed scheme has the lowest total and performs no point multiplication on the endpoint, while the ECC-bearing schemes (Li, Xiong) remain the most expensive; the spread is dominated by ECC point multiplications, which are inexpensive on a mobile CPU, moderate with a hardware accelerator, and dominant on a software-only constrained endpoint.

Recomputation under Table 7. The Table 7 basis places hash, symmetric encryption, and ECC point multiplication at MICAz-class speed when executing on the field-device (FD) or gateway (GWN) side, better reflecting the IIoT reality of resource-constrained sensor and field endpoints. Under this basis, the per-row computation is as follows: *Vinoth:* U = $T_{fe} + 8T_H + T_S = 20.23 + 0.24 + 0.12 = 20.59$ ms; GWN = $6T_H + 3T_S = 48 + 10.5 = 58.50$ ms; FD = $3T_H + 2T_S = 24 + 7 = 31$ ms; total 110.09 ms. *Das:* U = $T_{fe} + 14T_H = 20.23 + 0.42 = 20.65$ ms; GWN = $9T_H = 72$ ms; FD = $7T_H = 56$ ms; total 148.65 ms. *Li:* U = $T_{fe} + 8T_H + 3T_{ecm} = 20.23 + 0.24 + 60.69 = 81.16$ ms; GWN = $7T_H + T_{ecm} = 56 + 2450 = 2506$ ms; FD = $4T_H + 2T_{ecm} = 32 + 4900 = 4932$ ms; total 7519.16 ms. *Xiong:* U = $5T_H + 3T_{ecm} = 0.15 + 60.69 = 60.84$ ms; GWN = $11T_H + 6T_{ecm} = 88 + 14{,}700 = 14{,}788$ ms; FD = $4T_H = 32$ ms; total 14,880.84 ms. *Ours:* U = $2T_H + 2T_S = 0.30$ ms; GWN = $7T_H + 2T_S = 63$ ms; FD = $3T_H = 24$ ms; total 87.30 ms. The ECC-bearing schemes (Li, Xiong) rise from tens of milliseconds under a uniform mobile basis to thousands once each point multiplication is costed at 2450 ms on the constrained side (Li: 125.78 → 7519.16 ms; Xiong: 160.30 → 14,880.84 ms), whereas the hash-and-symmetric schemes (Das: 26.70 → 148.65 ms; Vinoth: 23.26 → 110.09 ms) and Ours (4.32 → 87.30 ms) change comparatively little because their constrained-side cost is dominated by inexpensive hash operations. Vinoth additionally performs a small number of $\mathbb{Z}_q$ modular multiplications for CRT-based secret sharing; these scalar operations are not represented in the Table 7 basis and are omitted, which is conservative against Ours because including them would only raise Vinoth's total.

Entity-mapping conventions. The four compared schemes share the three-tier user–gateway–device structure and map directly onto the U/GWN/FD columns of Table 8: Li [7] and Xiong [8] (user/gateway/sensor), Das [9] and Vinoth [10] (user/gateway/sensing-device). For Xiong, the registration center participates only in the periodic temporary certificate issuance, not in per-access authentication; consistent with our treatment of the one-time SM2 channel establishment, only the per-access authentication phase (user/gateway/sensor) is counted. ECC point multiplications are placed at whichever entity actually performs them, so a scheme that evaluates them on the sensor side (Li, two per session) incurs the largest constrained-side cost.

Communication conventions. Reported communication overheads (Table 9) are taken from each source under its own field-size convention; these conventions are closely comparable (identities 32–128 b, nonces and hash outputs 128–256 b, ECC points 160–320 b, timestamps 32 b), so the ordering is robust to normalization. For reference, the proposed scheme is sized with identity 128 b, nonce 128 b, timestamp 32 b, hash 256 b, symmetric block 128 b, ECC point 320 b, and MAC tag 256 b, yielding 1216 b of communication and 1152 b of storage.

Important caveat. Cross-scheme totals in Table 8 characterize relative algorithmic complexity on a single hardware envelope rather than identical-platform measurements. Each total is the same symbolic operation count evaluated under the bases of Table A1; both values are correct for their respective basis, and the difference reflects the choice of platform model rather than an error. We recompute under the Table 7 basis because it more faithfully models the IIoT deployment context (battery-powered field devices and sensor nodes on the device side, mains-powered gateways on the GWN side).

Per-modality metrics (Table 10). The FER row is estimated from fear + anger vs. neutral on AffectNet-8. The Physio. row uses a representative 0.95 binary F1 for WESAD-style 1D-CNN classifiers from follow-up literature; ref. [18] itself reports shallow baselines, not 1D-CNN. The behavioral row is an algorithmic proxy from a reconstruction-error time series detector rather than a validation result for IIoT login histories. These values are retained as model selection context only, and are not converted into false negative rates

or a joint miss probability. A production deployment must replace them with measured operating-point metrics from institutional calibration data.

Sensitivity analysis method (Figures 4 and 5). The sensitivity study of Section 5.5 is a scripted Monte Carlo evaluation of the late-fusion decision under documented synthetic score distributions. Benign logins draw per-modality scores $e, b, p \sim \mathrm{Beta}(1.5, 6)$ (mean ≈ 0.20) and high-risk synthetic logins draw $e, b, p \sim \mathrm{Beta}(4, 3)$ (mean ≈ 0.57); $N = 4 \times 10^5$ logins are drawn per population with a fixed seed. The weight sweep varies $\alpha_P \in [0, 0.5]$ with $\alpha_F = \alpha_B = (1 - \alpha_P)/2$; in the simplex panel, all $(\alpha_F, \alpha_B, \alpha_P = 1 - \alpha_F - \alpha_B) \geq 0$ on a 41×41 grid. The threshold analysis varies $\theta_{high} \in [0.35, 0.95]$ at fixed $\theta_{low} = 0.35$ and repeats the calculation under common-factor mixing $s_k^{(\gamma)} = (1 - \gamma)s_k + \gamma s_{shared}$ for $\gamma \in \{0, 0.3, 0.6, 0.9\}$. The reported quantities are $P(\text{Accept} \mid \text{high-risk synthetic})$, $P(TrustScore < \theta_{high} \mid \text{high-risk synthetic})$, $P(\text{Reject} \mid \text{benign synthetic})$, and $P(TrustScore \geq \theta_{low} \mid \text{benign synthetic})$. At the default independent configuration the script reports $P(\text{Accept} \mid \text{high-risk synthetic}) \approx 0.02$, $P(\text{Reject} \mid \text{benign synthetic}) \approx 0$, and a majority of high-risk synthetic logins in the re-verification region. These figures are an architectural sensitivity check under stated assumptions, not a claim of operational detection performance.

Appendix B. Same-Host Reference Benchmark (Supplementary)

Purpose. The analytical operation count evaluation in Section 5 compares schemes under three operation-time bases (uniform mobile, MICAz, Cortex-M4 with hardware ECC). To complement that analytical evidence with concrete reference implementation numbers, this appendix documents a reproducible same-host benchmark that measures the proposed protocol's online M1–M4 phase alongside three standardized lightweight AKE protocols. The benchmark is a supplementary artefact and is explicitly **not** a substitute for the multi-MCU, multi-TEE, real network, or energy measurement campaigns, which remain as future work.

Claim boundary. The benchmark is a native Windows x64 reference implementation. It is not a low-power MCU measurement; it is not a TEE-backed measurement, not a real OT network packet loss/jitter/small MTU measurement, and not an energy measurement. The transport is in-process for EDHOC and Noise, and is an in-memory lossless datagram pair for DTLS 1.3. Library initialization and static credential loading are placed before the per-iteration timer; the timed region covers per-session allocation, ephemeral key generation, all handshake computations, message serialization and peer processing, key derivation or export, and session cleanup. The proposed protocol is measured in its designed pre-shared key online/steady mode (a pre-established per-session Ui–GWN SK and a provisioned GWN–FD$_j$ TC_j), whereas the three reference protocols are measured as full handshakes that perform ECDHE, credential exchange, and key derivation on top of the symmetric primitives. Therefore, the two comparison strata are labeled separately in the summary table, and no like-for-like cold start speedup ratio is reported.

Protocols and dependencies. The four adapters use unmodified upstream state machines. EDHOC uses libedhoc 1.14.0 at commit `ed2b7ebeac28ab98963713ad3d146a2f61b9c7dc` (MIT) with P-256, AES-CCM, and SHA-256 (Method 3, cipher suite 2). DTLS 1.3 uses wolfSSL 5.9.1 (GPL-3.0; commercial licensing is also offered upstream) with mutual ECDSA-P256 certificates, ECDHE-P256, and `TLS_AES_128_GCM_SHA256`. Noise uses Noise-C 0.0.1 (MIT) at pattern `Noise_XX_25519_AESGCM_SHA256`. The proposed protocol adapter uses AES-128-GCM, HMAC-SHA-256, and SHA-256 via the Windows CNG provider (BCrypt). Exact experiment identifiers (commit SHAs, archive SHAs, binary SHAs, MSVC toolchain, Windows SDK) are recorded in the environment snapshot shipped with the benchmark.

Measurement protocol. Each protocol runs 20 randomized blocks of 50 measured iterations (1000 measured iterations per protocol) with 100 warm-up iterations in every new block process. The runner shuffles the per-iteration blocks across protocols using a fixed seed (`run_order_seed` = 20260622) to spread scheduler and cache effects. The measurement thread is pinned to logical CPU 0. Wall time is taken from `QueryPerformanceCounter` and CPU work from `QueryThreadCycleTime`; the report does not use the Windows process CPU time because its resolution is too coarse for this microbenchmark. The timed region begins immediately before the first handshake computation and ends when matching exporter output or traffic-key material is available to the peer. Library initialization, static credential parsing, and long-term key loading are outside the per-iteration timer; per-session ephemeral key generation for the full-handshake protocols remains inside the timed region.

Statistical analysis. Medians are reported with 95 % cluster bootstrap intervals that resample the 20 randomized blocks rather than individual iterations independently. Pairwise comparisons of the three full-handshake rows use two-sided Mann–Whitney U tests with the normal approximation on the 20 block medians, Holm correction for multiple testing, and Cliff's delta as a non-parametric effect size. Wire bytes and packet counts are taken from the actual adapter boundary; for DTLS, each wolfSSL send callback is one in-memory datagram. No analytical message-size estimate is substituted for the serialized output.

Byte layout convention. The 1216 b analytical communication figure of Table 9 counts the bare ciphertext payload (the per-message field sizes summed over the four M1–M4 messages), and is the cost basis used for the analytical cross-scheme comparison. The 376B figure below, reported by the proposed protocol adapter, counts the actual AES-GCM packets at the adapter boundary and includes the 12-byte IV and 16-byte tag framing that AES-GCM adds per packet. The two figures are not contradictory; they report the protocol's analytical payload cost and the protocol's on-the-wire packet size, respectively.

Table A2 reports the same-host Windows x64 reference benchmark.

Table A2. Same-host Windows x64 reference benchmark. Each row reports 1000 measured iterations (20 blocks × 50 iterations) with 100 warm-up iterations per block. The proposed protocol row is the pre-shared-key online phase, and is reported as a separate comparison stratum; the three reference rows are full handshakes. 95 % CIs are cluster bootstrap intervals on block medians. The DTLS wire size spread comes from variable-length ECDSA signature encoding. The measured iterations totaled $N = 4 \times 10^3$, of which 4000/4000 passed their protocol correctness gates. The benchmark harness, environment snapshot, raw data, randomized schedule, summary tables, and the full `REPRODUCTION_REPORT.md` are shipped with the supplementary material; the benchmark claim boundary is summarized in the final paragraph of this appendix.

Protocol/Mode	Stratum	Packets	Wire Bytes	Median (µs)	95 % CI (µs)
Proposed/`steady`	pre-shared-key online	4	376	8.4	8.3–8.4
EDHOC/`full`	full handshake	4	110	5995.6	5967.7–6020.0
Noise/`full`	full handshake	3	192	1153.5	1140.7–1169.6
DTLS 1.3/`full`	full handshake	13	2924 (2922–2926)	10,002.9	9939.1–10,063.1

Plain timing of in-process protocol cores on Windows x64. The measurement thread is pinned to logical CPU 0. Wall time uses `QueryPerformanceCounter`; CPU work uses `QueryThreadCycleTime`. Library initialization, static credential loading, sockets, network delay/loss/jitter, application data, energy, and TEE calls are not included in the timed region. The proposed protocol row omits the cold-start SM2 channel establishment of Section 4, and as such is reported as a separate comparison stratum that is not directly comparable to the three full-handshake rows.

Pairwise full-handshake tests. The three full-handshake rows are mutually separated on this host and configuration. Mann–Whitney U on the 20 block medians per protocol with Holm correction yields $p < 2.04 \times 10^{-7}$ and Cliff's $\delta = 1.0$ for every pair (DTLS/EDHOC

median ratio ≈ 1.67, DTLS/Noise ≈ 8.69, EDHOC/Noise ≈ 5.22). These tests establish separation on this host and configuration only; they do not establish protocol superiority across platforms or network conditions.

Benchmark scope and claim boundary. The new result is a reproducible same-host simulation added to provide quantitative reference evidence where low-power hardware was not available. The manuscript explicitly acknowledges that the requested low-power MCU/TEE, real-network loss/jitter/MTU, energy, and real PLC/RTU validation remain as future work. The manuscript does not call this a low-power hardware benchmark, and does not use the `steady` row to claim a like-for-like cold-start speedup over the standardized reference protocols.

References

1. Sisinni, E.; Saifullah, A.; Han, S.; Jennehag, U.; Gidlund, M. Industrial Internet of Things: Challenges, opportunities, and directions. *IEEE Trans. Ind. Inform.* **2018**, *14*, 4724–4734. [CrossRef]
2. Langner, R. Stuxnet: Dissecting a cyberwarfare weapon. *IEEE Secur. Priv.* **2011**, *9*, 49–51. [CrossRef]
3. Electricity Information Sharing and Analysis Center (E-ISAC) and SANS ICS. Analysis of the Cyber Attack on the Ukrainian Power Grid. Washington, DC, USA. 2016. Available online: https://web.archive.org/web/20180503055227/https://www.nerc.com/pa/CI/ESISAC/Documents/E-ISAC_SANS_Ukraine_DUC_18Mar2016.pdf (accessed on 28 June 2026).
4. Di Pinto, A.; Dragoni, Y.; Carcano, A. TRITON: The first ICS cyber attack on safety instrument systems. In Proceedings of the Black Hat USA, Las Vegas, NV, USA, 4–9 August 2018; pp. 1–26.
5. Verizon. 2024 Data Breach Investigations Report. Verizon Business. 2024. Available online: https://www.verizon.com/business/resources/reports/dbir/ (accessed on 28 June 2026).
6. Reason, J. *Human Error*; Cambridge University Press: Cambridge, UK, 1990.
7. Li, X.; Niu, J.; Bhuiyan, M.Z.A.; Wu, F.; Karuppiah, M.; Kumari, S. A robust ECC-based provable secure authentication protocol with privacy preserving for industrial Internet of Things. *IEEE Trans. Ind. Inform.* **2018**, *14*, 3599–3609.
8. Xiong, L.; Peng, T.; Li, F.; Zeng, S.; Wu, H. Privacy-preserving authentication scheme with revocability for multi-WSN in industrial IoT. *IEEE Syst. J.* **2023**, *17*, 38–49.
9. Das, A.K.; Wazid, M.; Kumar, N.; Vasilakos, A.V.; Rodrigues, J.J.P.C. Biometrics-based privacy-preserving user authentication scheme for cloud-based industrial Internet of Things deployment. *IEEE Internet Things J.* **2018**, *5*, 4900–4913.
10. Vinoth, R.; Deborah, L.J.; Vijayakumar, P.; Kumar, N. Secure multifactor authenticated key agreement scheme for industrial IoT. *IEEE Internet Things J.* **2021**, *8*, 3801–3811.
11. Kindervag, J. *Build Security into Your Network's DNA: The Zero Trust Network Architecture*; Forrester Research Inc.: Cambridge, MA, USA, 2010.
12. Ward, R.; Beyer, B. BeyondCorp: A new approach to enterprise security. *Login* **2014**, *39*, 6–11.
13. Edo, O.C.; Ang, D.; Billakota, P.; Ho, J.C. A zero trust architecture for health information systems. *Health Technol.* **2024**, *14*, 189–199. [CrossRef]
14. Moubayed, A.; Refaey, A.; Shami, A. Software-defined perimeter (SDP): State of the art secure solution for modern networks. *IEEE Netw.* **2019**, *33*, 226–233. [CrossRef]
15. Chen, B.; Qiao, S.; Zhao, J.; Liu, D.; Shi, X.; Lyu, M.; Chen, H.; Lu, H.; Zhai, Y. A security awareness and protection system for 5G smart healthcare based on zero-trust architecture. *IEEE Internet Things J.* **2021**, *8*, 10248–10263. [PubMed]
16. Mollahosseini, A.; Hasani, B.; Mahoor, M.H. AffectNet: A database for facial expression, valence, and arousal computing in the wild. *IEEE Trans. Affect. Comput.* **2019**, *10*, 18–31. [CrossRef]
17. Li, S.; Deng, W. Reliable crowdsourcing and deep locality-preserving learning for unconstrained facial expression recognition. *IEEE Trans. Image Process.* **2019**, *28*, 356–370. [PubMed]
18. Schmidt, P.; Reiss, A.; Duerichen, R.; Marberger, C.; Van Laerhoven, K. Introducing WESAD, a multimodal dataset for wearable stress and affect detection. In *Proceedings of the 20th ACM International Conference on Multimodal Interaction (ICMI), Boulder, CO, USA*; ACM: New York, NY, USA, 2018; pp. 400–408.
19. Koldijk, S.; Sappelli, M.; Verberne, S.; Neerincx, M.A.; Kraaij, W. The SWELL knowledge work dataset for stress and user modeling research. In *Proceedings of the 16th ACM International Conference on Multimodal Interaction (ICMI), Istanbul, Turkey*; ACM: New York, NY, USA, 2014; pp. 291–298.
20. Selander, G.; Preuß Mattsson, J.; Palombini, F. *Ephemeral Diffie–Hellman over COSE (EDHOC)*; RFC 9528; IETF: Wilmington, DE, USA, 2024. Available online: https://www.rfc-editor.org/info/rfc9528 (accessed on 28 June 2026). [CrossRef]
21. Selander, G.; Mattsson, J.; Palombini, F.; Seitz, L. *Object Security for Constrained RESTful Environments (OSCORE)*; RFC 8613; IETF: Wilmington, DE, USA, 2019. Available online: https://www.rfc-editor.org/info/rfc8613 (accessed on 28 June 2026). [CrossRef]

22. Rescorla, E.; Tschofenig, H.; Modadugu, N. *The Datagram Transport Layer Security (DTLS) Protocol Version 1.3*; RFC 9147; IETF: Wilmington, DE, USA, 2022. Available online: https://www.rfc-editor.org/info/rfc9147 (accessed on 28 June 2026). [CrossRef]
23. Perrin, T. The Noise Protocol Framework. Revision 34. 2018. Available online: https://noiseprotocol.org/noise.html (accessed on 28 June 2026).
24. Bormann, C.; Hoffman, P. *Concise Binary Object Representation (CBOR)*; RFC 8949; IETF: Wilmington, DE, USA, 2020. Available online: https://www.rfc-editor.org/info/rfc8949 (accessed on 28 June 2026). [CrossRef]
25. Schaad, J. *CBOR Object Signing and Encryption (COSE): Structures and Process*; RFC 9052; IETF: Wilmington, DE, USA, 2022. Available online: https://www.rfc-editor.org/info/rfc9052 (accessed on 28 June 2026). [CrossRef]
26. Hristozov, S.; Huber, M.; Xu, L.; Fietz, J.; Liess, M.; Sigl, G. The cost of OSCORE and EDHOC for constrained devices. In *Proceedings of the 11th ACM Conference on Data and Application Security and Privacy (CODASPY)*; ACM: New York, NY, USA, 2021; pp. 245–250. [CrossRef]
27. Restuccia, G.; Tschofenig, H.; Baccelli, E. Low-power IoT communication security: On the performance of DTLS and TLS 1.3. *arXiv* **2020**, arXiv:2011.12035. [CrossRef]
28. Meier, S.; Schmidt, B.; Cremers, C.; Basin, D. The TAMARIN prover for the symbolic analysis of security protocols. In *Computer Aided Verification (CAV)*; LNCS 8044; Springer: Berlin/Heidelberg, Germany, 2013; pp. 696–701. [CrossRef]
29. Blanchet, B. Modeling and verifying security protocols with the applied pi calculus and ProVerif. *Found. Trends Priv. Secur.* **2016**, *1*, 1–135. [CrossRef]
30. Jacomme, C.; Klein, E.; Kremer, S.; Racouchot, M. A comprehensive, formal and automated analysis of the EDHOC protocol. In *Proceedings of the 32nd USENIX Security Symposium*; USENIX Association: Berkeley, CA, USA, 2023; pp. 4203–4220.
31. Rose, S.; Borchert, O.; Mitchell, S.; Connelly, S. *Zero Trust Architecture*; NIST Special Publication 800-207; National Institute of Standards and Technology: Gaithersburg, MD, USA, 2020. [CrossRef]
32. Vanickis, R.; Jacob, P.; Dehghanzadeh, S.; Lee, B. Access control policy enforcement for zero-trust-networking. In Proceedings of the 29th Irish Signals and Systems Conference (ISSC), Belfast, UK, 21–22 June 2018; pp. 1–6.
33. Gutmann, A.; Renaud, K.; Maguire, J.; Mayer, P.; Volkamer, M.; Matsuura, K.; Müller-Quade, J. Zeta—Zero trust authentication: Relying on innate human ability, not technology. In *Proceedings of the IEEE European Symposium on Security and Privacy (EuroS&P), Saarbrücken, Germany*; IEEE: New York, NY, USA, 2016; pp. 357–371.
34. DeCusatis, C.; Liengtiraphan, P.; Pinelli, M. Implementing zero trust cloud networks with transport access control and first packet authentication. In *Proceedings of the IEEE International Conference on Smart Cloud, New York, NY, USA*; IEEE: New York, NY, USA, 2016; pp. 5–10.
35. Campbell, M. Beyond zero trust: Trust is a vulnerability. *Computer* **2020**, *53*, 110–113. [CrossRef]
36. Li, Z.; Zhang, X. An enhanced autoencoder-based anomaly detection model for time series data from wearable medical devices. *IEEE J. Biomed. Health Inform.* **2025**, *29*, 7222–7233. [CrossRef] [PubMed]
37. Luo, Y.; Xiao, X.; Cheng, L.; Peng, G.; Yao, D. Deep learning-based anomaly detection in cyber-physical systems: Progress and opportunities. *ACM Comput. Surv.* **2022**, *54*, 106.
38. Baltrušaitis, T.; Ahuja, C.; Morency, L.-P. Multimodal machine learning: A survey and taxonomy. *IEEE Trans. Pattern Anal. Mach. Intell.* **2019**, *41*, 423–443. [PubMed]
39. Maodong, C. src-openEuler/secGear. 2021. Available online: https://gitee.com/src-openeuler/secGear (accessed on 28 June 2026).
40. Demigha, O.; Larguet, R. Hardware-based solutions for trusted cloud computing. *Comput. Secur.* **2021**, *103*, 102117. [CrossRef]
41. Costan, V.; Devadas, S. Intel SGX explained. *IACR Cryptol. ePrint Arch.* **2016**, *86*, 1–118.
42. ARM Ltd. TrustZone Technology for Armv8-A. Available online: https://developer.arm.com/ip-products/security-ip/trustzone (accessed on 28 June 2026).
43. AMD. AMD Secure Encrypted Virtualization (SEV). Available online: https://developer.amd.com/sev/ (accessed on 28 June 2026).
44. Liu, X.; Li, Y.; Wang, M. A lightweight authentication protocol based on confidential computing for federated learning nodes. *Netinfo Secur.* **2022**, *22*, 37–45.
45. Liu, X.; Wang, J.; Yang, H.; Zhang, R. An Internet of Vehicles authentication protocol based on blockchain and secGear framework. *Netinfo Secur.* **2022**, *22*, 27–36.
46. Ren, F.; Zhang, X.; Lei, Z.; Liu, X. Research on key-value data storage and operation optimization method based on confidential computing. *Netinfo Secur.* **2025**, *25*, 1707–1717.
47. Dolev, D.; Yao, A.C. On the security of public key protocols. *IEEE Trans. Inf. Theory* **1983**, *29*, 198–208. [CrossRef]
48. Van Bulck, J.; Minkin, M.; Weisse, O.; Genkin, D.; Kasikci, B.; Piessens, F.; Silberstein, M.; Wenisch, T.F.; Yarom, Y.; Strackx, R. Foreshadow: Extracting the keys to the Intel SGX kingdom with transient out-of-order execution. In *Proceedings of the 27th USENIX Conference on Security Symposium*; USENIX Association: Berkeley, CA, USA, 2018; pp. 991–1008.

49. Murdock, K.; Oswald, D.; Garcia, F.D.; Van Bulck, J.; Gruss, D.; Piessens, F. Plundervolt: Software-based fault injection attacks against Intel SGX. In *Proceedings of the IEEE Symposium on Security and Privacy (S&P)*; IEEE: New York, NY, USA, 2020; pp. 1466–1482. [CrossRef]
50. Lapid, B.; Wool, A. Cache-Attacks on the ARM TrustZone Implementations of AES-256 and AES-256-GCM via GPU-Based Analysis. IACR Cryptology ePrint Archive, Paper 2018/621. 2018. Available online: https://eprint.iacr.org/2018/621 (accessed on 28 June 2026).
51. Ménétrey, J.; Göttel, C.; Khurshid, A.; Pasin, M.; Felber, P.; Schiavoni, V.; Raza, S. Attestation mechanisms for trusted execution environments demystified. In *Distributed Applications and Interoperable Systems (DAIS)*; Springer: Berlin/Heidelberg, Germany, 2022. [CrossRef]
52. Birkholz, H.; Thaler, D.; Richardson, M.; Smith, N.; Pan, W. *Remote ATtestation ProcedureS (RATS) Architecture*; RFC 9334; IETF: Wilmington, DE, USA, 2023. Available online: https://www.rfc-editor.org/info/rfc9334 (accessed on 28 June 2026). [CrossRef]
53. Lundblade, L.; Birkholz, H.; Fossati, T.; Wallace, C. *The Entity Attestation Token (EAT)*; RFC 9711; IETF: Wilmington, DE, USA, 2025. Available online: https://www.rfc-editor.org/info/rfc9711 (accessed on 28 June 2026). [CrossRef]
54. Birkholz, H.; Fitzgerald-McKay, J.; Schmidt, C.; Waltermire, D. *Concise Software Identification Tags*; RFC 9393; IETF: Wilmington, DE, USA, 2023. Available online: https://www.rfc-editor.org/info/rfc9393 (accessed on 28 June 2026). [CrossRef]
55. Eckel, M.; George, D.R.; Grohmann, B.; Krauß, C. Remote attestation with constrained disclosure. In *Proceedings of the 39th Annual Computer Security Applications Conference (ACSAC)*; ACM: New York, NY, USA, 2023; pp. 718–731. [CrossRef]
56. *IEC 61511*; Functional Safety—Safety Instrumented Systems for the Process Industry Sector. International Electrotechnical Commission (IEC): Geneva, Switzerland, 2016.
57. Sandler, M.; Howard, A.; Zhu, M.; Zhmoginov, A.; Chen, L.-C. MobileNetV2: Inverted residuals and linear bottlenecks. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Salt Lake City, UT, USA*; IEEE: New York, NY, USA, 2018; pp. 4510–4520.
58. Hochreiter, S.; Schmidhuber, J. Long short-term memory. *Neural Comput.* **1997**, *9*, 1735–1780. [CrossRef] [PubMed]
59. Saaty, T.L. *The Analytic Hierarchy Process: Planning, Priority Setting, Resource Allocation*; McGraw-Hill: New York, NY, USA, 1980.
60. Fawcett, T. An introduction to ROC analysis. *Pattern Recognit. Lett.* **2006**, *27*, 861–874. [CrossRef]
61. *GB/T 32918*; Information Security Technology—Public Key Cryptographic Algorithm SM2 Based on Elliptic Curves—Part 5: Parameter Definition. Standardization Administration of the People's Republic of China (SAC): Beijing, China, 2017.
62. *ISO/IEC 14888-3*; IT Security Techniques—Digital Signatures with Appendix—Part 3: Discrete Logarithm Based Mechanisms. ISO: Geneva, Switzerland, 2018.
63. *NIST SP 800-186*; Recommendations for Discrete Logarithm-Based Cryptography: Elliptic Curve Domain Parameters. National Institute of Standards and Technology: Gaithersburg, MD, USA, 2023.
64. Armando, A.; Basin, D.; Boichut, Y.; Chevalier, Y.; Compagna, L.; Cuéllar, J.; Hankes Drielsma, P.; Heám, P.C.; Kouchnarenko, O.; Mantovani, J.; et al. The AVISPA tool for the automated validation of internet security protocols and applications. In *Proceedings of the 17th International Conference on Computer Aided Verification (CAV), Edinburgh, UK*; Springer: Berlin/Heidelberg, Germany, 2005; Volume 3576, pp. 281–285.
65. Chen, C.-M.; Xiang, B.; Wu, T.-Y.; Wang, K.-H. An anonymous mutual authenticated key agreement scheme for wearable sensors in wireless body area networks. *Appl. Sci.* **2018**, *8*, 1074. [CrossRef]
66. Shim, K.-A. BASIS: A practical multi-user broadcast authentication scheme in wireless sensor networks. *IEEE Trans. Inf. Forensics Secur.* **2017**, *12*, 1545–1554. [CrossRef]
67. *IEC 62443*; Security for Industrial Automation and Control Systems. IEC: Geneva, Switzerland, 2022.
68. *NIST Special Publication 800-82 Rev. 3*; Guide to Operational Technology (OT) Security. National Institute of Standards and Technology: Gaithersburg, MD, USA, 2023. [CrossRef]
69. European Parliament and Council of the European Union. Regulation (EU) 2016/679 (GDPR), Article 9: Processing of Special Categories of Personal Data. 2016. Available online: https://gdpr-info.eu/art-9-gdpr/ (accessed on 28 June 2026).
70. European Parliament and Council of the European Union. Regulation (EU) 2016/679 (GDPR), Article 35: Data Protection Impact Assessment. 2016. Available online: https://gdpr-info.eu/art-35-gdpr/ (accessed on 28 June 2026).
71. European Data Protection Board. Guidelines 3/2019 on Processing of Personal Data Through Video Devices. Version 2.0, Adopted 29 January 2020. Available online: https://www.edpb.europa.eu/our-work-tools/our-documents/guidelines/guidelines-32019-processing-personal-data-through-video_en (accessed on 28 June 2026).

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.