[CAPABILITY PROFILE · AUTHREX-01]

SATA · SENSOR AND ACTOR TRUST ASSESSMENT

Produce a continuous, per-source trust scalar from sensor evidence and hardware attestation so that authority is contracted before a corrupted input is acted on, not after an invariant breaks.

TESTED MODELED **IMPLEMENTED**

Name variants: Published article title: Sensor Attestation and Trust Anchoring (Journal of Hardware and Systems Security). Both expansions refer to the same protocol; the catalog uses the program name.

PRIMARY FUNCTION	Continuous sensor-trust scalar in [0,1] from weighted Dempster-Shafer fusion across four diagnostics (internal consistency, cross-sensor agreement, temporal stability, physical plausibility), anchored in an attestation chain design (TPM 2.0 measurements, signatures, monotonic counters).
MISSION AREAS	Cross-domain foundation layer: uncrewed air and ground systems under GNSS denial, counter-UAS track validation, maritime sensor fusion, spacecraft telemetry integrity.
EVIDENCE ANCHORS	Peer-reviewed article, J. Hardware and Systems Security (Springer Nature) 10, 17 (2026), DOI 10.1007/s41635-026-00190-4; Zenodo DOI 10.5281/zenodo.18936251 (specification, reference implementation, reproducibility artifacts); seeded browser simulation.
CURRENT STANDING	Specified and published; protocol state machine model checked as reported in the article; reference implementation and seeded simulation exist; not built on hardware; not independently reproduced.
SELF-ASSESSED TRL	TRL 3 (analytical and simulation proof of concept of the critical function). No formal TRL determination has been performed.
INTELLECTUAL PROPERTY	U.S. provisional application 64/002,453, filed March 11, 2026 (receipt 74808459), reported as filed; unexamined; confers no enforceable rights.

[MISSION PROBLEM]

Autonomy stacks act on sensor reports whose integrity cannot be assumed under jamming, spoofing, calibration drift, or hardware compromise. A control decision that is formally correct on corrupted inputs is still a wrong outcome, and the supervisor typically learns of the corruption only when an invariant is violated. Under GNSS denial or radio-frequency interference the corruption can be gradual and consistent across correlated sources, which defeats simple thresholding and majority voting. The mission consequence is an action executed with full authority on evidence that no longer deserves it.

[OPERATIONAL RELEVANCE]

- Uncrewed air and ground systems in GNSS-denied or contested electromagnetic conditions (decision windows of seconds).
- Counter-UAS track validation where radar, RF identification, and electro-optical reports must be weighed before an engagement recommendation.
- Spacecraft telemetry integrity under light delay, where a supervisor cannot re-check the sensor in real time.
- Resilient PNT: trust in position is treated as evidence with its own validity state rather than as an assumed input.

[CAPABILITY DESCRIPTION]

SATA is the foundation trust layer of the pipeline. For each source it computes a trust scalar in the interval zero to one by combining four diagnostics with weighted Dempster-Shafer belief functions, keeping ignorance (no evidence) distinct from distrust (contradicting evidence). The published protocol binds sensor reports to cryptographically committed attestation records so that a report can be checked for origin and replay, and specifies a formal adversary model under which the signature barrier is analyzed. Every downstream stage consumes the resulting scalar: HMAA converts it to an authority tier and ADARA tests it for manipulation. SATA does not decide what the platform may do; it states how much the evidence deserves to be believed.

[SATA · ARCHITECTURE, ASSURANCE, EVIDENCE]

[TECHNICAL ARCHITECTURE]

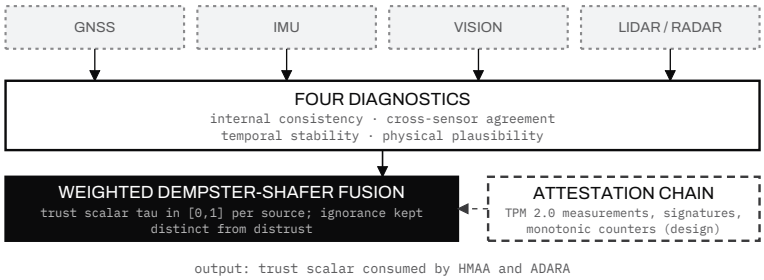

Figure 1. SATA technical architecture. Encoding: solid = implemented; dashed = specified; gray = external; filled = enforcement point.

[HOW IT WORKS]

1. Ingest raw and fused sensor reports at the host stack boundary, together with attestation records where the platform provides them.
2. Run the four diagnostics per source: internal consistency, cross-sensor agreement, temporal stability, physical plausibility.
3. Combine diagnostics with weighted Dempster-Shafer fusion into a trust scalar per source, preserving the ignorance-versus-distrust distinction.
4. Check the attestation chain (measurement, signature, monotonic counter) and reject replayed or unattested reports where the design is instantiated.
5. Emit the trust scalar over the authority interface to HMAA (tier computation) and ADARA (deception prior).
6. Record the scalar, diagnostics, and attestation outcome to the hash-chained ledger for later reconstruction.
7. Continue at the control-loop rate; a falling scalar contracts authority at once, and recovery of the scalar alone never restores authority.

[TECHNICAL DIFFERENTIATION]

Conventional runtime assurance switches to a safe controller when a monitored invariant is about to be violated and assumes the monitor's inputs are valid. Redundancy and voting mask a faulty channel but assume independent failure, which common-mode conditions such as spoofing defeat. SATA computes trust before authority is computed, so degraded inputs lower authority ahead of the violation rather than at the instant it occurs, and treats agreement across heterogeneous sources as evidence about trust rather than as truth. This is a complement to Simplex-style assurance and voting, not a replacement.

[ASSURANCE MECHANISMS: DESIGN PROPERTY VERSUS ESTABLISHED PROPERTY]

Mechanism	Status
Ignorance kept distinct from distrust in fusion	Design property; exercised in simulation
Attestation chain: origin, replay resistance, monotonic counters	Design property (hardware anchoring not built); adversary model stated in the article
Signature-barrier unforgeability bound	Analytical bound reported in the published article under its stated adversary model
Protocol state-machine safety invariants	Reported in the article as model checked at stated bounds (see evidence table); not independently reproduced
Deterministic scalar for identical inputs	Design property; reproducible in the seeded console
Auditability of every trust decision	Implemented in the seeded console (hash-chained record); not implemented on hardware

[TECHNICAL EVIDENCE]

Evidence	Result	Source / artifact
Peer-reviewed publication of the protocol	JHSS 10, 17 (2026); received 4 Jun, accepted 14 Jul, published 21 Jul 2026	DOI 10.1007/s41635-026-00190-4
TLA+ bounded model check of the protocol state machine, as reported in the article	18,892 distinct reachable states (47,616 generated); zero violations of eight safety invariants (as published; not re-executed for this catalog)	Article and Zenodo 10.5281/zenodo.18936251
Monte Carlo conformance and sensitivity campaign, as reported in the article	10,000 runs; bit-exact reproducibility of the packaged artifacts	Article and Zenodo deposit
Seeded browser simulation	Deterministic replay from seed; synthetic data	burakoktenli.com/sata-simulation
Separately archived TLC run (523 distinct states, depth 9)	WITHDRAWN from this catalog: the site-check log records the run as archived (SATA_verification.zip), but the archive could not be produced for this edition; the figure is not carried	Correction log CL-02

[EVIDENCE SNAPSHOT]

EVIDENCE SNAPSHOT	
Implemented	Yes (reference implementation)
Executable artifact	Yes
Formal model	Yes
Simulation evidence	Yes
Hardware validation	No
Field validation	No
Independent validation	No
Government evaluation	No
Operational deployment	No

[LIMITATIONS]

- Hardware anchoring (TPM 2.0 attestation chain) is a design; no platform carrying it has been built, so replay resistance and timing on real hardware are unmeasured.
- Sensor models are parameterized (Gaussian noise, step faults, drift ramps), not physics-based; adaptive adversaries are not modeled.
- The program's formal-coverage record lists no SATA module among the FAITHFUL models; the article-reported check is an as-published result.
- Simulation and formal-model evidence only; all data synthetic; no hardware-in-the-loop test, no field validation, no independent replication, no deployment.
- Not independently reproduced or technically verified by another party; second-environment reruns are confirmation by the same producer.

[TECHNOLOGY READINESS (SELF-ASSESSED; NO FORMAL TRA)]

CURRENT ASSESSED TRL	TRL 3 (self-assessed against the DoD nine-level scale; no formal technology readiness assessment has been performed).
BASIS	The critical function (continuous trust from fused diagnostics with attestation) has an analytical treatment in a peer-reviewed article, a reference implementation, and a seeded simulation that exercises it: analytical and experimental proof of concept. It has not been validated as a component in a laboratory environment on representative hardware, which TRL 4 requires. Disclosure: the published article states TRL 4 on its own reading (simulation-evaluated, with explicit assurance gaps G1 to G4 and hardware not yet measured); this catalog assesses TRL 3 for the reason given, and the difference is recorded rather than reconciled.
EVIDENCE SUPPORTING	concept defined and published; analytical treatment (adversary model, unforgeability bound); software implementation (reference); simulation demonstration
EVIDENCE REQUIRED FOR NEXT TRL	(1) Component validation in a laboratory environment: SATA running on a representative compute node against recorded or injected sensor streams with attestation hardware present. (2) Independent replication of the published model check and Monte Carlo campaign from the packaged deposit. (3) Measured contraction latency and false-restriction rate under representative interference (HIL).

[INTEGRATION PROFILE]

INPUTS	Raw and fused sensor reports; attestation records (TPM measurements, signatures, counters) where available; platform clock.
OUTPUTS	Per-source trust scalar in [0,1]; diagnostic vector; attestation verdict; ledger entry.
INTERFACES	Authority interface exposed by the host stack (message bus or middleware, ROS 2 candidate); TPM 2.0 for attestation (design).
DEPLOYMENT	Software module beside the controller at the sensor-ingest boundary; standalone browser console for evaluation.
DEPENDENCIES	Host sensor reports; optional hardware root of trust; no network required for the console.
SECURITY BOUNDARY	Inside: fusion logic, ledger. Outside: sensors, attestation hardware, controller. A compromised attestation root defeats the anchoring.
DATA REQUIREMENTS	Synthetic today; government-furnished recorded sensor streams (unclassified or CUI as the sponsor determines) would enable Phase 2.

[TRANSITION ROADMAP]

Phase	Activity for this capability	Who can execute it	Exit evidence
1 Technical review	Review of the artifacts listed in the evidence table	Program alone, with a government reviewer	Reviewer findings; correction log entries
2 Government-defined simulation	Sponsor scenarios on the seeded console	Program with sponsor scenarios	Deterministic, replayable results; hash-chained records
3 Degradation and failure injection	Evidence lapse, spoofing, delayed human response	Program with sponsor parameters	False-restriction, miss, and contraction-latency figures
4 Government testbed / SIL	Integration with a representative controller or interface	Requires sponsor interface and testbed	SIL integration report
5 Hardware-in-the-loop	Representative interference; timing on real hardware	Requires sponsor laboratory or rig	Measured latencies; TRL 4 to 5 evidence
6 Relevant-environment demonstration	Independent evaluation and red team	Requires independent evaluator	Independent report
7 Program integration	Only after earlier gates succeed	Requires transition partner	Not claimed to be available

[GOVERNMENT ENGAGEMENT REQUEST (SPECIFIC)]

- Technical peer review of the published protocol against the reviewer's own sensor-integrity threat models.
- Representative recorded sensor datasets with known degradation or spoofing episodes (unclassified is sufficient to start).
- Laboratory access with a ROS 2 platform and a TPM-equipped compute node for Phase 4 to 5 component validation.
- Independent replication of the packaged model-checking and Monte Carlo artifacts by a government laboratory or FFRDC.

[REVIEWER QUESTIONS THIS PROFILE ANSWERS]

What is it: see capability description. Why does DoD care: see operational relevance. Where would it fit: see integration profile. Does it exist: see evidence snapshot. What evidence supports it: see technical evidence. Has anyone independent tested it: no. Has it been tested in the real world: no. What does it rely on: see limitations and security boundary. What can cause it to fail and what does it do then: see assurance mechanisms and the published counterexamples. How mature: see the TRL block. What would the government need to provide: see the request above. Next measurable milestone: first item under evidence required for next TRL.