SIPCORE
Internet-Draft
Intended status: Best Current Practice
Expires: 18 September 2026

Q. Wang
J. Chen
Tsinghua University

17 March 2026

Mitigating SIP Identity Spoofing Caused by Semantic Ambiguities
draft-sip-identity-confusion-bcp-00

Abstract

The Session Initiation Protocol (SIP) carries originator identity through multiple header fields --- From, P-Asserted-Identity, P-Preferred-Identity, Remote-Party-ID, Identity and so on --- whose processing is spread across several specifications. Because these mechanisms leave significant room for implementation choice, SIP servers and user agents frequently disagree on which identity signal to trust and how to parse it. Recent research has demonstrated that these semantic ambiguities enable caller-ID and message-origin spoofing in the majority of tested server-client combinations, even when TLS and authentication are correctly deployed.

This document provides best current practices for mitigating identity spoofing caused by such ambiguities. It defines identity-validation rules at each stage of the SIP signaling path, from the originating server through trust boundaries and transit proxies to the receiving user agent, and includes guidance on deployment, migration, and interoperability with legacy systems.

Status of This Memo

Table of Contents

1. Introduction

The Session Initiation Protocol (SIP), specified in [RFC3261], is the signaling standard for voice, video, and messaging services including VoIP, VoLTE, and RCS. SIP does not directly include a single unified mechanism that controls how originator identity is asserted, validated, transported, and displayed across all entities in the signaling path. This document summarizes common existing guidelines and helps SIP operators and implementers apply coherent identity-handling policies.

A recent systematic study [SIPCONFUSION] evaluated six open-source SIP servers and nine user agents and found that 47 of 54 server-UA combinations were susceptible to ambiguity-based identity spoofing. The same study confirmed real-world impact across commercial VoIP hardware, public SIP services, and carrier-grade RCS messaging platforms. Importantly, these spoofing outcomes persist even when TLS, Digest authentication, and STIR are correctly deployed, because the vulnerabilities lie in how identity information is handled across entities rather than in the absence of authentication primitives. Section 5 catalogues these vulnerability classes and references the corresponding mitigations defined in this document.

2. Scope of the Document

The guidelines defined in this document are intended for SIP deployments that carry originator identity for INVITE and MESSAGE [RFC3428] methods. MESSAGE is explicitly in scope because SIP-based messaging directly drives user trust decisions and has been shown to be vulnerable to identity spoofing [SIPCONFUSION].

This document does not define new SIP headers, does not replace baseline authentication frameworks, and does not mandate any specific national regulatory profile. It does not modify the normative requirements of [RFC3261], [RFC3325], or [RFC8224]; it provides deployment guidance that tightens implementation behavior within the latitude already permitted by those specifications.

This document is complementary to [RFC8862], which binds SIP-layer identity to media-layer keys for media confidentiality. The guarantees of [RFC8862] depend on the identity assertions being correct; this document addresses the identity-handling gaps that can undermine those guarantees (see Section 11 for details).

3. Definitions and Acronyms

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

AoR:
: Address of Record. A SIP URI that points to a domain with a location service that can map the URI to a set of contact URIs where the user might be available.

Authentication point:
: The first trusted SIP server (typically a proxy or registrar) that authenticates the originator of a request and derives a canonical identity. Analogous to the ingress router in BGP that validates route announcements.

Canonical identity:
: The single identity value selected by the authentication point after credential verification and policy checks, bound to the transaction context for downstream use.

Identity-bearing field:

Any SIP header field or URI component used by a receiving entity to infer or display originating identity. Includes: From, P-Asserted-Identity (PAI), P-Preferred-Identity (PPI), Remote-Party-ID (RPID), and the Identity header field.

PAI:
P-Asserted-Identity, defined in [RFC3325].

PPI:
P-Preferred-Identity, defined in [RFC3325].

RPID:
Remote-Party-ID, a non-standardized header for network-asserted identity, originally defined in the expired draft-ietf-sip-privacy-04 and still supported by some implementations.

SBC:
Session Border Controller. A B2BUA or proxy deployed at the edge of an administrative domain to enforce policy at trust boundaries.

Trust Domain:
A set of SIP entities that have been configured to trust one another in accordance with a common specification (Spec(T)), as defined in [RFC3324] and used by [RFC3325].

Trust boundary:
A point in the signaling path where administrative trust changes: between an untrusted endpoint and its domain proxy, between an external peer and a provider edge SBC, or between two administrative domains that have not established mutual trust.

In addition to the list above, the following terms are used with a specific meaning.

Authenticated endpoint:
A UA that has presented valid credentials (e.g., via SIP Digest) to its domain’s authentication point.

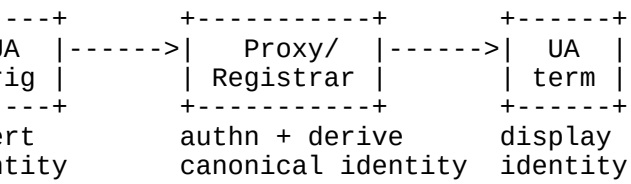
Trusted peer:
An adjacent SIP entity that is a member of the local Trust Domain (explicitly configured with a shared Spec(T)).

Untrusted source:
Any SIP entity that is not a member of the local Trust Domain, including authenticated endpoints (which are trusted to be who they claim but are not trusted to assert other identities), and cross-domain peers without inter-domain authentication.

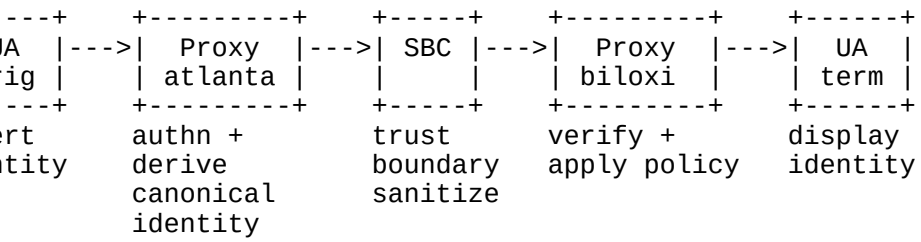
4. SIP Identity Architecture

A SIP request carrying originator identity traverses a chain of entities. The chain varies by deployment scenario. Two common topologies are:

Intra-domain (e.g., alice@example.com calls bob@example.com):



Inter-domain (e.g., alice@atlanta.com calls bob@biloxi.com):



In both topologies, identity-bearing fields enter the chain at the originating UA and are processed at each hop. Each entity has a distinct role:

1. *Assertion*: The originating UA populates From (and optionally PPI) in the outgoing request.

2. *Authentication and derivation*: The first trusted proxy (the authentication point) verifies the originator's credentials and derives a canonical identity. It may insert PAI and/or a STIR Identity header.

3. *Trust-boundary sanitization* (inter-domain only): SBCs and edge proxies at domain boundaries strip or rewrite identity-bearing fields that are not appropriate for the next trust context. In intra-domain scenarios, the originating proxy may serve both the authentication and trust-boundary role.

4. *Inter-domain validation* (inter-domain only): The receiving domain's edge proxy verifies incoming identity assertions (e.g., STIR signature validation) and applies local policy.

5. *Display*: The terminating UA selects which identity-bearing fields to render and presents them to the user.

Spoofing vulnerabilities arise when adjacent entities in this chain disagree on which field carries authoritative identity, how to parse its value, or whether to trust its source. The BCP sections below follow this chain and specify controls at each stage. Intra-domain deployments without SBCs still require the controls in Section 6 (authentication point) and Section 7.1 (inbound sanitization), as these address the risks from authenticated in-domain attackers and parser divergence that exist regardless of whether requests cross domain boundaries.

5. Threats and Vulnerabilities

This section lists known vulnerabilities in SIP identity handling that have been demonstrated in practice. Each description is followed by references to the mitigation sections in this document. The vulnerability classes below are derived from the systematic evaluation in [SIPCONFUSION], which tested six open-source SIP servers and nine user agents and found 47 of 54 server-UA combinations susceptible to at least one class of identity spoofing.

5.1. Identity-Authentication Mismatch

[RFC3261] permits the From header field to differ from authenticated credentials, as this flexibility was designed to support anonymity. However, the same flexibility allows an authenticated user to set the From addr-spec to any in-domain identity that is neither their own nor an anonymous URI. If the server does not enforce alignment between the From value and the authenticated identity, the attacker can impersonate any user in the domain while presenting valid credentials. This was the most broadly successful attack class in [SIPCONFUSION], succeeding against all six tested servers in their default configurations. Notably, some implementations enforce this binding for INVITE but skip it for MESSAGE, allowing message-origin spoofing through the same infrastructure that correctly protects voice calls.

For mitigations, see Section 6.1, Section 6.2, and Section 6.3.

5.2. Parser Differential Exploitation

SIP's text-based syntax allows header fields to be constructed in ways that different parsers interpret differently. When a server's authentication module extracts one identity from a From header but the receiving UA extracts a different identity from the same header, the alignment check at the server passes while the UA displays a spoofed identity. Specific constructs that trigger parser differentials include:

* Ambiguous quoting of display-names containing angle brackets or semicolons.

* Null bytes (0x00) and control characters that cause some parsers to truncate or split the header value.

* Mixed use of display-name quoting styles with embedded URI-like strings.

For mitigations, see Section 7.1.5 and Section 7.1.4.

5.3. Domain Encoding Confusion

Percent-encoding in the host component of SIP URIs is prohibited by Section 19.1.2 of [RFC3261], but some implementations accept and decode it. An attacker can exploit this by encoding parts of a domain name to bypass domain-matching logic while the decoded form matches a legitimate domain at the display layer. Similarly, internationalized domain names can be used to create visual confusion between attacker-controlled and legitimate domains.

For mitigations, see Section 7.1.5.

5.4. Trust Boundary Header Smuggling

P-Asserted-Identity and P-Preferred-Identity are defined for use within a Trust Domain ([RFC3325]). When a server accepts these headers from an untrusted source without stripping them, an external attacker can inject a PAI value claiming any identity, including a local-domain identity. If the receiving UA or a downstream proxy prefers PAI over From for display, the injected identity is presented to the user. This attack is particularly effective against MESSAGE requests, where the identity directly influences user trust decisions for received text messages.

For mitigations, see Section 7.1.3 and Section 7.1.

5.5. Identity Signal Precedence Disagreement

A SIP request may simultaneously carry identity information in From, PAI, RPID, and the STIR Identity header. When the server verifies identity using one field (e.g., From via Digest) but the UA displays a different field (e.g., PAI or RPID), an attacker can spoof the displayed field while satisfying the checked field. This is especially dangerous when non-standardized headers like RPID take precedence at the UA, as RPID is not subject to any authentication or trust-domain controls.

For mitigations, see Section 6.4 and Section 7.1.6.

5.6. Display-Name Spoofing

STIR/PASSporT signs only the addr-spec portion of the From header; the display-name is explicitly excluded (Section 12.6 of [RFC8224]). Because many UAs render the display-name as the primary identity indicator and some render it as the sole indicator, an attacker can set the display-name to a misleading value (e.g., "IT Security <sip:attacker@evil.example>") while the addr-spec passes STIR verification. No currently deployed standard provides integrity protection for display-names.

For mitigations, see Section 6.5 and Section 8.3.

6. Protecting Identity at the Authentication Point

The authentication point is the first line of defense against identity spoofing. It authenticates the originator and establishes the canonical identity that downstream entities rely on.

6.1. Identity-Authentication Binding

Section 8.1.1.3 of [RFC3261] defines the From header field as "the logical identity of the initiator" and does not require it to match the authenticated identity --- the specification provides for anonymity by allowing a non-identifying From value. The specification recommends using "Anonymous" with a meaningless URI (e.g., sip:thisis@anonymous.invalid) when the client’s identity is to remain hidden.

However, SIP does not restrict the From value to either the authenticated identity or a well-formed anonymous identity. This permits an authenticated user to populate From with an arbitrary in-domain identity that the server may accept and forward.

A server at the authentication point MUST enforce the following constraint on requests from authenticated endpoints:

* If the From addr-spec matches the authenticated identity (after URI comparison per Section 19.1.4 of [RFC3261]): accept normally.

* If the From addr-spec is a well-formed anonymous URI (i.e., the host is "anonymous.invalid", or the form conforms to [RFC3323]): accept as a privacy request and apply privacy handling per [RFC3323] and [RFC3325].

* If the From addr-spec matches neither the authenticated identity nor a permitted anonymous form: reject with 403 (Forbidden), unless the From addr-spec is an alias explicitly bound to the authenticated user in the server’s provisioning system (e.g., a role-based address such as helpdesk@example.com assigned to specific users).

The principle is: an authenticated user may speak as themselves, or may speak anonymously, but MUST NOT speak as someone else.

6.2. Canonical Identity Derivation

After verifying the identity-authentication binding, the authentication point MUST derive exactly one canonical identity per request and bind it to the transaction/dialog context.

If PPI is present, the authentication point MAY use it as a hint for selecting among multiple valid identities for the authenticated user (e.g., role aliases). The PPI MUST then be removed before forwarding (Section 6 of [RFC3325]).

Downstream entities MUST use the canonical identity --- not raw header values from the originating UA --- for authorization, routing policy, and display-identity decisions.

6.3. MESSAGE Method Alignment

Implementations MUST apply the same identity-authentication binding and canonical identity derivation to MESSAGE [RFC3428] as to INVITE.

SIP entities MUST NOT skip identity-alignment checks for MESSAGE simply because no media session is established. The [SIPCONFUSION] evaluation found that in some widely deployed stacks, MESSAGE bypasses the alignment checks that are correctly applied to INVITE, creating an inconsistency exploitable for SMS spoofing over RCS and similar messaging paths.

6.4. Identity Signal Precedence

A single SIP request may carry identity information in multiple headers simultaneously. Without a unified precedence rule, different entities may trust different fields, enabling an attacker to spoof the field that the UA displays while authenticating against a different field that the server checks.

Servers that derive a canonical identity MUST apply a documented precedence policy. A RECOMMENDED order is:

1. Cryptographically verified identity --- a valid PASSporT/STIR Identity header whose signature has been verified and whose "orig" claim matches the From addr-spec [RFC8224].

2. Network-asserted identity --- PAI received from an explicitly trusted peer within the same Trust Domain [RFC3325].

3. Locally authenticated identity --- the identity bound to the Digest authentication context.

6.5. Display-Name Scrubbing

PASSporT/STIR signs only the addr-spec of the From header; the display-name is explicitly excluded (Section 12.6 of [RFC8224]). Since many UAs render the display-name as the primary identity indicator, an attacker who controls the display-name can present misleading identity even when the addr-spec is verified.

Trusted servers SHOULD validate the From display-name of authenticated requests against a list of acceptable display-names for the authenticated user. If the display-name does not match policy, the server SHOULD either strip it (forwarding only the addr-spec) or, where the display-name is clearly being used to spoof the identity, reject the request with 403 (Forbidden) as recommended in Section 12.6 of [RFC8224].

When rewriting identity headers for forwarding, servers SHOULD set the display-name to a server-validated value (e.g., the registered display name for the authenticated AoR) or omit it entirely.

7. Protecting Identity at Trust Boundaries

Trust boundaries are the SIP equivalent of BGP peering edges: they are the points where identity assertions from one administrative context enter another. Proper sanitization at these boundaries prevents header smuggling, unauthorized identity assertion, and cross-domain spoofing.

7.1. Inbound Sanitization

Inbound sanitization encompasses both source-based identity filtering (stripping headers by trust relationship) and structural validation (rejecting duplicate headers, malformed syntax, and legacy headers that could undermine canonical identity).

7.1.1. From Authenticated Endpoints

Authenticated endpoints are trusted to be who they claim (via Digest credentials) but are NOT trusted to assert identities other than their own. At the authentication point, the server MUST:

* Remove PPI (it is a hint that has been consumed).

* Remove RPID if present (endpoints are not trusted to insert network-asserted identity).

* Verify From aligns with authenticated identity per Section 6.1.

* Insert PAI reflecting the canonical identity if the request will traverse the Trust Domain.

7.1.2. From Trusted Peers (Intra-Trust-Domain)

Messages from a peer explicitly configured as a Trust Domain member (with a shared Spec(T) per [RFC3325]): the receiving server MAY accept the PAI as authoritative. The server SHOULD still verify consistency between PAI and From, and SHOULD log any discrepancy.

7.1.3. From Untrusted and Cross-Domain Sources

Messages from any source not explicitly in the local Trust Domain --- including cross-domain peers, external SIP trunks, and unknown IP addresses --- MUST be treated as untrusted. The receiving server MUST:

* Remove PAI and RPID: these headers are meaningful only within a Trust Domain (Section 5 of [RFC3325]). A proxy receiving PAI from an untrusted source MUST replace or remove it.

* Remove PPI: this header MUST NOT traverse trust boundaries (Section 6 of [RFC3325]).

* Reject requests where any identity-bearing field claims a local-domain identity unless validated by an inter-domain authentication mechanism (see Section 7.3).

* If forwarding the request, mark it as unverified for all downstream identity-display decisions.

7.1.4. Duplicate Header Filtering

The From header field’s grammar is not a comma-separated list (Section 7.3.1 of [RFC3261]), so multiple From headers are prohibited by the base specification. Implementations MUST enforce this:

* Messages with multiple From headers MUST be rejected as malformed (400).

* Comma-joining MUST NOT be used to attempt recovery of multiple From values.

* Downstream entities MUST NOT apply "first wins" or "last wins" fallback when duplicate singletons are encountered.

This requirement does not restrict fields that legitimately allow multiple values: the Identity header may appear more than once per [RFC8224], and PAI may carry up to two values (one SIP/SIPS URI and one tel URI) per Section 9.1 of [RFC3325].

7.1.5. Parser Hardening

The semantic gap between how a server’s authentication module parses a From header and how a UA renders it is the mechanism that enables parser-differential spoofing. To close this gap, implementations MUST apply strict, deterministic parsing to identity-bearing fields.

Implementations MUST:

* Reject identity-bearing headers containing null bytes (0x00) or other control characters. These are not permitted by [RFC3261] ABNF grammar.

* Reject percent-encoded characters in the host component of SIP URIs, as prohibited by Section 19.1.2 of [RFC3261].

* Reject From headers where the boundary between display-name and addr-spec is ambiguous --- for example, quoted display-names containing unescaped angle brackets or semicolons that could be interpreted as URI delimiters.

* Apply normalization (if any) before identity policy decisions. The same normalized value MUST be used for authentication, canonical identity derivation, and forwarding.

This is an intentional security tightening over permissive parsing. SIP's text-based format makes the robustness principle dangerous for identity handling: accepting and "best-effort" interpreting a malformed identity header is precisely the behavior that creates parser differentials. For identity-bearing fields, the correct posture is: if the input is ambiguous, reject it.

7.1.6. Legacy Header Scrubbing

RPID was never formally standardized by the IETF, yet it remains supported by several widely deployed UAs. Its presence in the signaling path can override canonical identity at UAs that prefer RPID over From or PAI.

RPID and other non-standardized identity-bearing headers SHOULD be disabled by default. If legacy support is required for interoperability with specific peers:

* it MUST be explicitly enabled per-peer;
* it MUST be constrained to requests received from those trusted peers;
* it MUST NOT override canonical identity;
* it MUST be logged for security monitoring.

7.2. Outbound Identity Handling

When forwarding toward an entity outside the Trust Domain, treatment of PAI MUST follow the Privacy header semantics in Section 7 of [RFC3325]:

* If Privacy contains "id": remove PAI before forwarding.
* If Privacy is "none": do NOT remove PAI.
* If no Privacy header is present: behavior is governed by local Spec(T) policy. It is RECOMMENDED that PAI SHOULD NOT be removed unless local privacy policies require it.

7.3. Inter-Domain Identity Validation

When accepting SIP traffic from another domain, operators SHOULD deploy cryptographic identity validation (e.g., PASSporT/STIR) where supported by ecosystem and regulation [RFC8224] [RFC8225] [RFC8226].

If inter-domain identity validation fails, the server MUST apply one of the following actions, chosen according to deployment profile and documented in the local Spec(T):

* Reject: respond with 438 (Invalid Identity Header) when STIR signature verification fails, or 403 (Forbidden) when a cross-domain request claims a local-domain identity without any verifiable proof.

* Downgrade: accept the request but strip all unverified identity-bearing fields and mark the call as "unverified" in all downstream signaling (e.g., via a P-Identity-Status header or equivalent local mechanism).

* Divert: route the request to an operator-defined screening queue or IVR that warns the callee before connecting.

The default action MUST be reject or downgrade; silent acceptance of unverified cross-domain identity claims is NOT RECOMMENDED.

Operators SHOULD monitor cross-domain validation failure rates per peer and generate alerts when anomalous patterns are detected.

Servers that do not support any inter-domain authentication mechanism MUST NOT forward cross-domain requests with local-domain identity claims as though they were locally originated.

7.4. Failure Semantics

When an identity-related check fails, the server MUST respond with a specific SIP response code so that the originating entity can distinguish between failure types. The following mapping defines the expected response for each class of identity failure:

Failure Condition	Response Code	BCP Section
From addr-spec does not match authenticated identity and is not a valid anonymous URI	403 (Forbidden)	Section 6.1
Multiple From headers detected	400 (Bad Request)	Section 7.1.4
Null bytes or control characters in identity field	400 (Bad Request)	Section 7.1.5
Percent-encoded host component in SIP URI	400 (Bad Request)	Section 7.1.5
Ambiguous display-name / addr-spec boundary	400 (Bad Request)	Section 7.1.5
Cross-domain request claiming local identity without inter-domain authentication	403 (Forbidden)	Section 7.1.3
STIR Identity signature verification failure	438 (Invalid Identity Header)	Section 7.3

Table 1: Identity Failure Response Codes

Implementations SHOULD include the Reason header field (e.g., Reason: SIP;cause=403;text="From-auth mismatch") to support automated diagnostics.

8. Protecting Identity at the Receiving User Agent

The receiving UA is the last entity in the identity chain and the point where identity is presented to a human user. The [SIPCONFUSION] evaluation found that none of the twelve tested UAs (nine softphones and three hardware phones) displayed any verification status or trust indicator --- all rendered verified and unverified identities with identical visual weight.

8.1. Identity Selection for Display

A UA MUST apply the same identity-signal precedence as defined in Section 6.4. Deprecated or non-standardized headers (e.g., RPID received from outside a Trust Domain) MUST be ignored for display by default.

8.2. Trust and Provenance Indicators

UAs MUST visually distinguish between at least three identity assurance levels: verified (cryptographic proof via STIR), unverified (no assertion or proof available), and verification failed (STIR signature present but invalid). An unverified identity MUST NOT be rendered with the same visual prominence as a verified identity.

Where the STIR signing domain or PAI domain differs from the From addr-spec domain, UAs SHOULD indicate this provenance difference to the user (e.g., a "via" annotation, analogous to email clients showing "via sendgrid.net" when the DKIM signing domain differs from the From domain). Where the request originates from outside the local administrative domain, UAs SHOULD indicate external origin (e.g., an "[External]" tag).

8.3. Display-Name Handling

Display-names are not integrity-protected by any deployed standard (Section 12.6 of [RFC8224]). UAs MUST NOT render the display-name as the sole identity indicator; the addr-spec MUST always be visible or accessible with minimal interaction. When the display-name resembles a URI or phone number that differs from the actual addr-spec, the UA SHOULD warn the user.

8.4. Duplicate Header Handling at the UA

If a UA receives a request containing multiple From headers (which the server-side controls in Section 7.1.4 should have prevented), the UA MUST NOT silently select one value. The UA MUST either reject the request or display a warning indicating that the identity is malformed.

9. Deployment and Operational Considerations

The controls in this document apply to all SIP deployments, but deployment contexts differ in risk tolerance and legacy constraints. As a general principle:

* Carrier/interconnect environments, where traffic crosses trust boundaries by default, SHOULD apply the strictest controls and treat cross-domain identity assertions as untrusted absent cryptographic verification.

* Enterprise/PBX environments, where most traffic is intra-domain, SHOULD still enforce From-authentication binding and parser hardening but MAY defer cross-domain validation to the SBC or session border element.

* All environments MUST enforce the core safety controls (From-authentication binding, duplicate-From rejection, untrusted-source PAI stripping, and parser hardening) regardless of deployment context.

To support monitoring, auditing, and phased migration, trusted SIP elements SHOULD log the canonical identity result, original and post-sanitization header values, trust-boundary classification of the source, and the reject/allow decision for each identity-bearing request. Operators SHOULD alert on repeated malformed identity attempts from the same source.

9.1. Testing

Every MUST-level requirement in this document SHOULD be covered by at least one negative test case that verifies the expected rejection behavior. Appendix A provides a mapping of normative requirements to minimal test inputs and expected results.

Implementations SHOULD include SIP parser differential test cases inspired by [RFC4475] and the attack patterns in [SIPCONFUSION].

10. Migration and Compatibility

Because strict validation can expose long-standing non-compliant behaviors, operators SHOULD deploy these controls in phases:

1. Monitor-only mode: log all identity anomalies without rejecting traffic. Identify legitimate non-conformant peers.

2. Staged enforcement: reject highest-risk patterns (duplicate From, null-byte injection, cross-domain PAI smuggling) while applying compatibility-mode rewrites (see below) for lower-risk anomalies.

3. Full enforcement: reject all non-conformant identity inputs, with a documented exception process for peers that require additional migration time.

Exception policies MUST be time-bounded, with a maximum duration of 12 months before mandatory re-evaluation.

10.1. Compatibility Mode

To preserve interoperability with legacy peers that generate non-conformant headers, deployments MAY implement a compatibility mode that rewrites the ambiguous origin identity to a well-formed anonymous identity token (sip:anonymous@anonymous.invalid) [RFC3323] before forwarding.

If compatibility mode is enabled, the server MUST:

* Mark the request as unverified for all downstream display and policy logic.

* Strip or overwrite all conflicting identity-bearing fields.

* Prevent privileged treatment based on the rewritten identity.

* Emit auditable logs including the original header values and the reason for rewrite.

11. Security Considerations

This document is entirely about improving security posture for SIP identity handling. The primary security objective is preventing identity spoofing at each stage of the signaling path described in Section 4.

The identity-authentication binding (Section 6.1) enforces that the only legitimate alternative to presenting one’s authenticated identity is presenting an anonymous identity. This aligns with the original design intent of Section 8.1.1.3 of [RFC3261] while closing the loophole that allows impersonation disguised as a "different logical identity."

Applying this BCP reduces impersonation opportunities for both authenticated in-domain and unauthenticated cross-domain attackers. Residual risk remains for environments that retain permissive parsing, legacy identity features, or undefined trust boundaries. Implementers and operators should review guidance from [RFC3552] when performing threat analysis.

The media confidentiality framework in [RFC8862] relies on STIR to provide integrity protection for SDP key fingerprints. If identity-bearing fields are spoofed before STIR signing occurs (Section 5.1), or if trust-boundary sanitization fails to prevent header smuggling

(Section 5.4), an attacker may cause STIR to sign a manipulated identity, undermining the end-to-end assurance that [RFC8862] provides. Conversely, even with correct identity handling per this BCP, media confidentiality requires the additional mechanisms specified in [RFC8862]. The two documents together address identity integrity (this document) and identity-to-media binding ([RFC8862]).

12. Privacy Considerations

The identity-authentication binding in Section 6.1 explicitly preserves the anonymity mechanism of [RFC3261] and [RFC3323]. Users who wish to remain anonymous may use anonymous URI forms. The restriction is only on impersonating another specific identity.

Identity sanitization and logging can process personally identifiable information. Operators SHOULD minimize retained data, apply role-based access controls, and follow jurisdictional privacy requirements.

13. IANA Considerations

This document has no IANA actions.

14. References

14.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <https://www.rfc-editor.org/rfc/rfc2119>.

[RFC3261] Rosenberg, J., Schulzrinne, H., Camarillo, G., Johnston, A., Peterson, J., Sparks, R., Handley, M., and E. Schooler, "SIP: Session Initiation Protocol", RFC 3261, DOI 10.17487/RFC3261, June 2002, <https://www.rfc-editor.org/rfc/rfc3261>.

[RFC3323] Peterson, J., "A Privacy Mechanism for the Session Initiation Protocol (SIP)", RFC 3323, DOI 10.17487/RFC3323, November 2002, <https://www.rfc-editor.org/rfc/rfc3323>.

[RFC3324] Watson, M., "Short Term Requirements for Network Asserted Identity", RFC 3324, DOI 10.17487/RFC3324, November 2002, <https://www.rfc-editor.org/rfc/rfc3324>.

[RFC3325] Jennings, C., Peterson, J., and M. Watson, "Private Extensions to the Session Initiation Protocol (SIP) for Asserted Identity within Trusted Networks", RFC 3325, DOI 10.17487/RFC3325, November 2002, <https://www.rfc-editor.org/rfc/rfc3325>.

[RFC3428] Campbell, B., Ed., Rosenberg, J., Schulzrinne, H., Huitema, C., and D. Gurle, "Session Initiation Protocol (SIP) Extension for Instant Messaging", RFC 3428, DOI 10.17487/RFC3428, December 2002, <https://www.rfc-editor.org/rfc/rfc3428>.

[RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174, May 2017, <https://www.rfc-editor.org/rfc/rfc8174>.

14.2. Informative References

[RFC3552] Rescorla, E. and B. Korver, "Guidelines for Writing RFC Text on Security Considerations", BCP 72, RFC 3552, DOI 10.17487/RFC3552, July 2003, <https://www.rfc-editor.org/rfc/rfc3552>.

[RFC4475] Sparks, R., Ed., Hawrylyshen, A., Johnston, A., Rosenberg, J., and H. Schulzrinne, "Session Initiation Protocol (SIP) Torture Test Messages", RFC 4475, DOI 10.17487/RFC4475, May 2006, <https://www.rfc-editor.org/rfc/rfc4475>.

[RFC8224] Peterson, J., Jennings, C., Rescorla, E., and C. Wendt, "Authenticated Identity Management in the Session Initiation Protocol (SIP)", RFC 8224, DOI 10.17487/RFC8224, February 2018, <https://www.rfc-editor.org/rfc/rfc8224>.

[RFC8225] Wendt, C. and J. Peterson, "PASSporT: Personal Assertion Token", RFC 8225, DOI 10.17487/RFC8225, February 2018, <https://www.rfc-editor.org/rfc/rfc8225>.

[RFC8226] Peterson, J. and S. Turner, "Secure Telephone Identity Credentials: Certificates", RFC 8226, DOI 10.17487/RFC8226, February 2018, <https://www.rfc-editor.org/rfc/rfc8226>.

[RFC8862] Peterson, J., Barnes, R., and R. Housley, "Best Practices for Securing RTP Media Signaled with SIP", BCP 228, RFC 8862, DOI 10.17487/RFC8862, January 2021, <https://www.rfc-editor.org/rfc/rfc8862>.

[SIPCONFUSION]
Wang, Q., Chen, J., Yang, J., Zhang, J., Yang, Y., and H. Duan, "SIPConfusion: Exploiting SIP Semantic Ambiguities for Caller ID and SMS Spoofing", NDSS Symposium 2026, February 2026, <https://www.ndss-symposium.org/ndss-paper/sipconfusion-exploiting-sip-semantic-ambiguities-for-caller-id-and-sms-spoofing/>.

Appendix A. Requirement-to-Test Matrix

This appendix maps each MUST-level normative requirement to a minimal negative test input, the expected SIP response code, and the expected downstream header state after correct processing. Each row constitutes a minimum-viable test case.

A.1. Server-Side Authentication and Sanitization Tests

Req ID	Requirement	BCP Section
R-AUTH-1	From must match auth or be anonymous	Section 6.1
R-AUTH-2	Anonymous URI accepted as privacy	Section 6.1
R-CANON-1	PPI removed after consumption	Section 6.2
R-MSG-1	MESSAGE gets same checks as INVITE	Section 6.3
R-STRIP-1	Strip PAI from untrusted source	Section 7.1.3
R-STRIP-2	Strip PPI from untrusted source	Section 7.1.3
R-STRIP-3	Reject cross-domain local-domain claim without STIR	Section 7.1.3

Table 2: Authentication and Sanitization Tests

Req ID	Test Input	Expected Result
R-AUTH-1	From: bob, Digest user=alice	403; not forwarded
R-AUTH-2	From: anonymous, Digest user=alice	2xx; PAI=alice, From=anon
R-CANON-1	PPI present from auth'd endpoint	2xx; PPI absent downstream
R-MSG-1	MESSAGE From: bob, Digest user=alice	403; not forwarded
R-STRIP-1	PAI from external IP	2xx; PAI absent or replaced
R-STRIP-2	PPI from external IP	2xx; PPI absent downstream
R-STRIP-3	From: local-domain from evil.example	403; not forwarded

Table 3: Authentication and Sanitization Test Inputs and Results

A.2. Filtering and Parser Tests

Req ID	Requirement	BCP Section
R-DUP-1	Reject multiple From headers	Section 7.1.4
R-PARSE-1	Reject null bytes in identity fields	Section 7.1.5
R-PARSE-2	Reject %-encoded host in SIP URI	Section 7.1.5
R-PARSE-3	Reject ambiguous display-name/addr-spec	Section 7.1.5
R-XDOM-1	STIR signature failure triggers 438	Section 7.3
R-XDOM-2	Cross-domain without STIR must not forward as local	Section 7.3

Table 4: Filtering and Parser Tests

Req ID	Test Input	Expected Result
R-DUP-1	Two From: headers in request	400; not forwarded
R-PARSE-1	From with null byte (0x00)	400; not forwarded
R-PARSE-2	From host %-encoded	400; not forwarded
R-PARSE-3	Ambiguous delimiters in From	400; not forwarded
R-XDOM-1	Invalid PASSporT signature	438; not forwarded
R-XDOM-2	External INVITE, no Identity/PAI	2xx or 403; mark "unverified"

Table 5: Filtering and Parser Test Inputs and Results

A.3. UA Display Tests

Req ID	Requirement	BCP Section
R-UA-1	UA must distinguish unverified identity	Section 8.2
R-UA-2	UA must show "via" for differing signing domain	Section 8.2
R-UA-3	UA must not render display-name as sole indicator	Section 8.3
R-UA-4	UA must reject or warn on duplicate From	Section 8.4

Table 6: UA Display Tests

Req ID	Test Input	Expected UA Behavior
R-UA-1	INVITE with From only, no PAI/Identity	"Unverified" label
R-UA-2	From: alice@a.example, Identity signed by b.net	"alice@a.example via b.net"
R-UA-3	Display-name looks like a different address	addr-spec visible; warning
R-UA-4	Request with two From: headers	Reject or "malformed" warning

Table 7: UA Display Test Inputs and Results

Implementations SHOULD include all rows above in automated regression suites. Additional test vectors based on [RFC4475] and [SIPCONFUSION] attack patterns are RECOMMENDED.

Appendix B. Example Attack Flows

B.1. In-Domain Impersonation (Authentication Point Failure)

An authenticated user (alice) sends:

```
INVITE sip:victim@example.com SIP/2.0
From: "admin" <sip:admin@example.com>;tag=a1b2c3
Proxy-Authorization: Digest username="alice", ...
```

The server authenticates alice but does not check From alignment. The request is forwarded with the spoofed From.

Mitigation: Section 6.1 requires rejection (From is neither alice's identity nor anonymous).

B.2. Parser-Differential Bypass (Filtering Failure)

```
From: "\"<sip:admin@example.com>;"<sip:alice@example.com>
```

The server's authentication module extracts alice@example.com (correct); the UA extracts admin@example.com (incorrect). The alignment check passes at the server, but the UA displays the wrong identity.

Mitigation: Section 7.1.5 requires rejection of ambiguous delimiter constructs.

B.3. Cross-Domain PAI Smuggling (Trust Boundary Failure)

```
MESSAGE sip:victim@example.com SIP/2.0
From: <sip:someone@evil.example>;tag=d4e5f6
P-Asserted-Identity: <sip:admin@example.com>
```

The server forwards without stripping PAI from the untrusted source. The UA prefers PAI and displays admin@example.com.

Mitigation: Section 7.1.3 requires stripping PAI from untrusted sources.

Acknowledgements

The editors thank the SIP security research and operations communities for demonstrating practical identity-confusion attack scenarios and sharing deployment lessons that informed this BCP.

Authors' Addresses

Qi Wang
Tsinghua University
China
Email: qi-wang23@mails.tsinghua.edu.cn

Jianjun Chen
Tsinghua University
China
Email: jianjun@tsinghua.edu.cn