**ATIS-0x0000x**

ATIS Standard on

**Signature-based Handling of Asserted Information Using Tokens**

**Alliance for Telecommunications Industry Solutions**

Approved Month DD, YYYY

**Abstract**

Signature-based Handling of Asserted information using Tokens (SHAKEN) is an industry framework for managing the deployment of Secure Telephone Identity (STI) technologies with the purpose of providing end-to-end cryptographic authentication and verification of the telephone identity and other information in an IP-based service provider voice network. This specification defines the framework for telephone service providers to create signatures in SIP and defines the key Network-to-Network Interface (NNI) requirements, Network Elements, the X.509 certificate framework to validate the initiator of the signature, and the various classes of signers and how the verification of a signature can be used towards the mitigation and identification of illegitimate use of national telecommunications infrastructure and protecting its users.

**Foreword**

The Alliance for Telecommunications Industry Solutions (ATIS) serves the public through improved understanding between carriers, customers, and manufacturers. The [**COMMITTEE NAME**] Committee [**INSERT MISSION**]. [**INSERT SCOPE**].

The mandatory requirements are designated by the word *shall* and recommendations by the word *should*. Where both a mandatory requirement and a recommendation are specified for the same criterion, the recommendation represents a goal currently identifiable as having distinct compatibility or performance advantages. The word *may* denotes an optional capability that could augment the standard. The standard is fully functional without the incorporation of this optional capability.

Suggestions for improvement of this document are welcome. They should be sent to the Alliance for Telecommunications Industry Solutions, [**COMMITTEE NAME**], 1200 G Street NW, Suite 500, Washington, DC 20005.

At the time of consensus on this document, [**COMMITTEE NAME**], which was responsible for its development, had the following leadership:

[**LEADERSHIP LIST**]

The **[SUBCOMMITTEE NAME]** Subcommittee was responsible for the development of this document.

**Revision History**

| **Date** | **Version** | **Description** | **Author** |
| --- | --- | --- | --- |
| March 24, 2016 | 0.1 | Initial Draft | Chris Wendt |
| August 25, 2016 | 0.2 | Baseline Draft | Chris Wendt |

**Table of Contents**

[INSERT]

**Table of Figures**

[INSERT]

**Table of Tables**

[INSERT]

# Scope & Purpose

## Scope

This document is intended to provide telephone services providers with a framework and guidance on how to utilize Secure Telephone Identity (STI) technologies toward the validation of legitimate calls and the mitigation of illegitimate spoofing of telephone identities on the VoIP Telephone Network. The focus of this document is the network signaling.

## Purpose

Using the protocols defined in draft-ietf-stir-rfc4474bis and draft-ietf-stir-passport, this document will define the signature-based handling of asserted information using tokens (SHAKEN) framework. This framework is targeted at telephone service providers delivering telephone calls over VoIP, addressing the implementation and usage of the IETF STIR WG protocols and the architecture and use of STI-related certificates on VoIP networks. It also discusses the general architecture of service provider authentication and verification services and identifies NNI and peering impacts and dependencies. Finally, it provides high level guidance on the use of positive or negative verification of the signature to mitigate illegitimate telephone identity in general.

Illegitimate Caller ID spoofing is a growing concern for North American telephone service providers and their customers. There are many Caller ID spoofing mechanisms, and illegitimate spoofing can evolve to evade mitigation techniques. Service provider solutions must therefore be flexible to respond to evolving threats in much the same way as cyber security solutions. In addition, the integration of new technologies into established VoIP networks imposes many interoperability and interworking challenges. As a result, this document is a baseline document on the implementation of the protocol related requirements for STI. The objective is to provide a baseline that can evolve over time, incorporating more comprehensive functionality and a broader scope in a backward compatible and forward looking manner.

# Normative References

The following standards contain provisions which, through reference in this text, constitute provisions of this Standard. At the time of publication, the editions indicated were valid. All standards are subject to revision, and parties to agreements based on this Standard are encouraged to investigate the possibility of applying the most recent editions of the standards indicated below.

draft-ietf-stir-passport

draft-ietf-stir-rfc4474bis

draft-ietf-stir-certificates

IETF RFC 3325 - Private Extensions to SIP for Asserted Identity within Trusted Networks

# Definitions, Acronyms, & Abbreviations

For a list of common communications terms and definitions, please visit the *ATIS Telecom Glossary*, which is located at < <http://www.atis.org/glossary> >.

## Definitions

**Caller ID**: the originating or calling party’s telephone number used to identify the caller carried either in the P-Asserted ID or From header.

## Acronyms & Abbreviations

|  |  |
| --- | --- |
| ATIS | Alliance for Telecommunications Industry Solutions |
| NNI | Network-to-Network Interface |
| PSTN | Public Switched Telephone Network |
| STI | Secure Telephone Identity |
| VoIP | Voice over Internet Protocol |

# Overview

This document presents the SHAKEN framework. SHAKEN is defined as a framework that utilizes protocols defined in the IETF STIR working group (WG) that work together in an end-to-end architecture for the authentication and assertion of a telephone identity by an originating service provider and the validation of the telephone identity by a terminating service provider.

Today, assertion of telephone identity in VoIP networks between peering service providers, particularly in a 3GPP IMS environment, typically uses the P-Asserted-ID as defined in RFC3325 as a network self-asserted identity. This usage assumes an inherent trust model between peering providers. However, in many telephone calling scenarios where there are many indirect call path relationships between the originating and terminating providers, these trust relationships are often simply not verifiable and do not allow for identification of the true origination of the call. Currently, the P-Asserted-ID header field can be populated by an enterprise PBX and passed on without validation by the service provider. Secure Telephone Identity (STI) as defined in the STIR WG and the usage of cryptographic digital signatures to validate the originator of a signed identity can provide a verifiable mechanism to identify the authorized originator of a call into the telephone network with non-repudiation and assignment of an attestation indicator and a unique originating identifier depending on how and where the call is originated in the VoIP network. This attestation and identifier represent the originating signer’s ability to vouch for the accuracy of the source of origin of the call. For example, if the service provider has an authenticated direct relationship with the origination of the call this attestation is categorized differently than calls that are originated from different networks or gateways that the service provider may have received from an unauthenticated network or that are unsigned. Verification of signatures will use these attestations as information to provide trace back mechanisms as well as information to feed into any call spam identification techniques the service provider has enabled on behalf of their customer.

## STIR Overview

The documents draft-ietf-stir-rfc4474bis and draft-ietf-stir-passport define a set of protocol level tools that can be used in SIP for applying digital signatures to the Caller ID or telephone number of the calling party.

### PASSporT Token

The document draft-ietf-stir-passport defines a token based signature that combines the use of JSON Web Tokens, JSON Web Signatures, and X.509 certificate key pairs, or PKI, to create a trusted signature. The authorized owner of the certificate used to generate the signature can be validated and traced back to the known trust anchor who signed the certificate. The PASSporT token includes a number of claims the signer of the token is asserting. The associated public certificate is used to verify the digital signature and the claims included in the PASSporT token. The public certificate is also used to validate the entity that signed the token through a SPID identifier, as defined in draft-ietf-stir-certificates. The validated claims, and the validated identity of the entity signing the claims, can both be used to determine the level of trust in the originating entity and their asserted calling party information. Call blocking applications or other mitigation techniques could use the information over time to determine “reputation” of the entity signing the token, which could provide further input to determine the level of trust for the calling party information. Note that PASSporT tokens and signatures themselves are agnostic to network signaling protocols but are used in draft-ietf-stir-rfc4474bis to define specific SIP usage described in next section.

### RFC4474bis

The document draft-ietf-stir-rfc4474bis defines a SIP based framework for an authentication service and verification service for using the PASSporT signature in a SIP INVITE. It defines a new Identity header field that delivers the PASSporT signature and other associated parameters. The authentication service adds the Identity header field and signature to the SIP INVITE generated by the originating provider. The INVITE is delivered to the destination provider which uses the verification service to verify the signature using the identity in the P-Asserted-ID header field or From header field.

## SHAKEN Architecture

There are a number of architectural components required for an end-to-end framework for STI.

The figure below shows the SHAKEN reference architecture. This is a logical view of the architecture and doesn’t mandate any particular deployment and/or implementation. For reference, this architecture is specifically based on the 3GPP IMS architecture with an IMS application server, and is only provided as an example to set the context for the functionality described this document. The diagram shows the two IMS instances that comprise the IMS half-call model; an originating IMS network hosted by Service Provider A, and a terminating IMS network hosted by Service Provider B.

 

Figure : SHAKEN reference architecture

This SHAKEN reference architecture includes the following elements:

* SIP UA - SIP User Agent that is authenticated by the service provider network is considered secure and the calling party identity is “known” since it is under direct management by the telephone service provider. It initiates the SIP INVITE as the calling party.
* IMS/CSCF - This component represents the SIP registrar and routing function. It also has a SIP application server interface.
* IBCF/TrGW - This function is at the edge of the service provider network and represents the NNI or peering interconnection point between telephone service providers. It is the ingress and egress point for SIP calls between providers.
* Authentication Service (STI-AS) - The SIP application server that performs the function of the authentication service defined in 4474bis. It SHOULD either itself be highly secured and contain the Secure Private Key Store (SKS) or have an authenticated, TLS encrypted interface to the SKS which stores the secret private key(s) used to create the PASSporT signature.
* Verification Service (STI-VS) - The SIP application server that performs the function of the verification service defined in 4474bis. It has an HTTPS interface to the TN Certificate Repository, that is referenced in the Identity header field to retrieve the provider public key certificate.
* Call Validation Treatment (CVT) - This is a logical function that could be an application server function or a third party application for applying anti-spoofing mitigation techniques once the signature is positively or negatively verified and then provides a response to signal the display response for the end user.
* SKS – Secure Key Store is a logical highly secure element to store private key(s) for the authentication service (STI-AS) to access.
* Certificate Provisioning Service – A logical service used to provision certificate(s) used for STI.
* TN Certificate Repository (TN-CR): This represents the publically accessible store for public key certificates. This should be an HTTPS web service that can be validated back to the owner of the public key certificate.

The focus of this document is on the STI-AS and STI-VS functionality and the relevant SIP signaling and interfaces. Detailed functionality for the Certification Provisioning Service, the TN-CR, the SKS and the CVT will be provided in separate document(s).

## SHAKEN call flow

 

Figure : SHAKEN reference call flow

1. The originating SIP UA, which first REGISTERs and is authenticated to the CSCF, creates a SIP INVITE with a telephone number identity.
2. The CSCF of the originating provider adds a P-Asserted-Identity header field asserting the Caller ID of the originating SIP UA. The CSCF then initiates an originating trigger to the STI-AS for the INVITE.

Note: The originating triggers should be configured such that the STI-AS is invoked after any AS that updates To, From, or Date header fields.

1. The STI-AS in the originating SP (i.e., Service Provider A) securely requests its private key from the SKS.
2. The SKS provides the private key in the response, and the STI-AS signs the INVITE and adds an Identity header field per RFC 4474bis using the Caller ID in the P-Asserted-Identity header field.
3. The STI-AS passes the INVITE back to the SP A CSCF.
4. The originating CSCF, through standard resolution, routes the call to the egress IBCF.
5. The INVITE is routed over the NNI through the standard inter-domain routing configuration.
6. The terminating SP (Service Provider B) ingress IBCF receives the INVITE over the NNI.
7. The terminating CSCF initiates a terminating trigger to the STI-VS for the INVITE.
8. The terminating SP STI-VS uses the “info” parameter information in the Identity header field per RFC 4474bis to determine the TN-CR URI and makes an HTTPS request to the TN-CR.
9. The STI-VS validates the certificate, which can include these steps: check the validity dates, check the certificate’s signature, check chain of trust, and check certificate validity via CRLs and/or OCSP. It then extracts the public key. It constructs the RFC 4474bis format and uses the public key to verify the signature in the Identity header field, which validates the Caller ID used when signing the INVITE on the originating service provider STI-AS.
10. The CVT is an optional function that can be invoked to perform call spam analytics or other mitigation techniques and return a response related to what should be displayed to the user for a legitimate or illegitimate call. The CVT may be integrated in the service provider network or outside the service provider network by a third party.
11. Depending on the result of the STI validation, the STI-VS determines that the call is to be completed with any appropriate indicator (that may be defined outside of this document) and the INVITE is passed back to the terminating CSCF which continues to set up the call to the terminating SIP UA. Note: Error cases where verification fails are discussed in Section 6.
12. The terminating SIP UA receives the INVITE and normal SIP processing of the call continues, returning “200 OK”, or optionally setting up media end-to-end.

 Note: This document assumes that the verification occurs prior to terminating call features or B2BUAs that modify the To or Date header fields.

# STI SIP Procedures

Both draft-ietf-stir-4474bis and draft-ietf-stir-passport define a base set of procedures for how STI fits into the SIP call flow. 4474bis defines an authentication service, corresponding to STI-AS in the SHAKEN reference architecture, as well as a verification service or STI-VS. This section will detail the procedures required for the STI-AS to create the identity header required.

## PASSporT Token Overview

STI as defined in draft-ietf-stir-passport specifies the process of the PASSporT token.

PASSporT tokens have the following form:

* A protected header with the value BASE64URL(UTF(JWS Protected Header))
* A payload with the value BASE64URL(JWS Payload)
* A signature with the value BASE64URL(JWS Signature)

An example of each is as follows:

*Protected Header*

{

 "typ":"passport",

 "alg":"ES256",

 "x5u":"https://cert.example.org/passport.crt"

}

*Payload*

{

"iat":"1443208345",

 "orig":{“tn”:"12155551212"},

 "dest":{“tn”:"12155551213"}

}

draft-ietf-stir-passport has specific examples of a Passport token.

## 4474bis Authentication procedures

### PASSporT and identity header construction

For the SHAKEN framework, standard PASSporT base claims should be used as defined in both PASSporT and RFC4474bis documents.

The ‘orig’ claim and ‘dest’ claim MUST be of type ‘tn’.

The ‘orig’ claim ‘tn’ value should be derived using the following rules:

* The P-Asserted-ID header field MUST be used as the telephone identity, if present, otherwise the From header field MUST be used.
* If there is more than one P-Asserted-ID, Authentication service MUST have logic to choose the most appropriate based on service provider policy.

In the case where no called or calling telephone number is available (e.g., the To header field does not contain a telephone number), the authorization service is terminated without signing the Caller ID.

### PASSporT extension ‘shaken’

The base passport set of claims cover the assertion of the telephone number along with date and destination telephone numbers to avoid replay attacks using valid Identity header fields. This section will detail a specific extension to the PASSporT to cover the following requirements of SHAKEN.

1. The ability to provide an attestation indicator for the context of how the call was originated.
2. The ability to provide a unique identifier that can service as an opaque indication of where in the service provider network the call was originated.
	1. This identifier MUST be globally unique and consistent so it can be used in analytics for tracking the reputation of a particular originating service.
	2. This identifier MUST be globally unique and consistent so it can be used for any traceback efforts if a particular originator is a consistent or pervasive “bad actor”.

Draft-wendt-stir-shaken-passport-extension defines a PASSporT extension IANA registered names for “shaken” extension and the claims that represent both an attestation indicator (attest) and Origination Identifier (origID). The format of these claims are defined in the following sections, but the PASSporT token would have the form:

*Protected Header*

{

 "alg":"ES256",

 "typ":"passport",

 “ppt”:”shaken”,

"x5u":"https://cert.example.org/passport.crt"

}

*Payload*

{

 “attest”:”A”

 "dest":{“uri”:"sip:alice@example.com"}

"iat":"1443208345",

 "orig":{“tn”:"12155551212"},

 “origid”:”123e4567-e89b-12d3-a456-426655440000”

}

### Attestation Indicator

This indication allows for both identifying the service provider that is vouching for the call as well as a clear indication of what information the service provider is attesting to.

In the SHAKEN framework we will use the for attestation in the following scenarios:

A.  **Full Attestation:** The signing provider:

* is responsible for the origination of the call onto the IP based service provider voice network
* has a direct authenticated relationship with the customer and can identify the customer
* has established a verified association with the telephone number used for the call.

Note: The signing provider is asserting that their customer can “legitimately” insert the number that appears as the calling party (i.e., the Caller ID). The legitimacy of the telephone number(s) the originator of the call can use is subject to signer specific policy, but could use mechanisms such as the following:

* The number was assigned to this customer by the signing service provider.
* This number is one of a range of numbers assigned to an enterprise or wholesale customer.
* The signing service provider has ascertained that the customer is authorized to use a number (e.g. by business agreement or evidence the customer has access to use the number). This includes numbers assigned by another service provider.
* The number is not permanently assigned to an individual customer but the signing provider can track the use of the number by a customer for certain calls or during a certain timeframe.

Note: ultimately it is up to service provider policy to decide what constitutes “legitimate right to assert a telephone number” but the service provider’s reputation may be directly dependent on how rigorous they have been.

B. **Partial Attestation:** The signing provider:

* is responsible for the origination of the call onto its IP based voice network
* has a direct authenticated relationship with the customer and can identify the customer
* has NOT established a verified association with the telephone number being used for the call

Note: When partial attestation is used, each customer will have a unique origination identifier created and managed by the service provider, but the intention is that it will not be possible to reverse engineer the identity of the customer purely from the identifier or signature. As described in section 5.2.4, the unique origination identifier provides a consistent identifier to allow data analytics to establish a reputation profile and assess the reliability of information asserted by the customer assigned this unique identifier. The identifier also provides a reliable mechanism to determine the customer for forensic analysis or legal action where appropriate.

C. **Gateway Attestation:** The signing provider:

* is the entry point of the call onto its IP based voice network
* has no relationship with the initiator of the call (e.g., international gateways).

Note: The token will provide a unique originating identifier of the node in the “origID” claim. (The signer is not asserting anything other than “this is the point where the call entered my network”.)

For the PASSporT extension claim, the “attest” key value pair MUST be set to “A”, “B”, or “C” corresponding to the appropriate attestation listed above.

### Origination Identifier (origID)

In addition to attestation, the unique origination identifier (origID) is defined as part of SHAKEN. This unique origination ID should be a globally unique string corresponding to a UUID (RFC4122) that is set as the serial number attribute in the subject name.

The purpose of the unique origination identifier is to assign an opaque identifier corresponding to the service provider initiated calls themselves, customers, classes of devices, or other groupings that a service provider should use for determining things like reputation or trace back identification of customers or gateways.

For Full Attestation, in general, a single identifier will be used as part of the certificate representing direct service provider initiated calls on its IP based voice network. A service provider though may choose to have a pool of identifiers to differentiate regions or classes of customers for example. Best practices will likely develop as the traceback and illegitimate call identification practices evolve.

For Partial Attestation, a single identifier per customer is required in order to differentiate calls both for traceback and for reputation segmentation so one customers reputation doesn’t affect other customers or the service provider’s call reputation. A service provider may choose to be more granular (e.g., per node per customer) depending on size and classes of services that the service provider offers as well.

For Gateway Attestation, best practices will dictate that the origID should be sufficiently granular to identify the originating node or trunk to allow for trace back identification and reputation scoring.

## 4474bis Verification procedures

Draft-ietf-stir-rfc4474bis defines the procedures for verification services including the methods used to verify the signature contained in the Identity header field.

### PASSporT and identity header verification

The certificate referenced in the info parameter of the Identity header field MUST be validated by performing the following:

* check the validity dates
* check the certificate’s signature
* check chain of trust
* check certificate validity via CRLs and/or OCSP

The PASSporT token provided in the Identity header field of the INVITE MUST validate the presence of all of the baseline claims as well as SHAKEN extension claims. It MUST also follow RFC4474bis defined verification procedures to check the corresponding Date, Originating Identity and Destination Identity.

The ‘orig’ claim and ‘dest’ claim MUST be of type ‘tn’.

The ‘orig’ claim ‘tn’ value validation MUST be performed as follows:

* The P-Asserted-ID header field MUST be checked as the telephone identity to be validated if present, otherwise the From header field MUST be checked.
* If there is more than one P-Asserted-ID, verification MUST check each P-Asserted-ID value until it finds one that is valid.
* If there are one or more P-Asserted-ID header fields present, but none are valid, then the Caller ID verification process is deemed to have failed (i.e., don’t check the From in this case).

### Verification Error conditions

If the authentication service functions correctly, and the certificate is valid and available to the verification service, the SIP message can be delivered successfully. However, if these conditions are not satisfied, errors can be generated as defined in rfc4474bis. This section identifies important error conditions and specifies procedurally what should happen, if they occur. Error handling procedures should consider how best to always deliver the call per current regulatory requirements, while providing diagnostic information back to the signer.

There are five main procedural errors defined in rfc4474bis that can identify issues with allowing the validation of the Identity header field to occur. They are:

**403** - ‘Stale Date’ – sent when the verification service receives a request with a Date header field value that is older than the local policy for freshness permits.

**428** – ‘Use Identity Header’ is not recommended for SHAKEN until a point where all calls on the VoIP network are mandated either by local or global policy to be signed.

**436** – ‘Bad-Identity-Info’ – the URI in the info parameter cannot be dereferenced (i.e., the request times out or receives a 4xx or 5xx error)

**437** – ‘Unsupported credential’ – this error occurs when a credential is supplied by the info parameter but the verifier doesn’t support it or it doesn’t contain the proper certificate chain in order to trust the credentials.

**438** – ‘Invalid Identity Header’ – this occurs if the signature verification fails

If any of these error conditions occur, a reason code MUST be used in either a 18x provisional response or any final response that indicates one of the five above scenarios back to the authentication service of any error conditions.

Example:

Reason: SIP ;cause=436 ;text=“Bad Identity Info”

In addition, if any of the base claims or shaken extension claims are missing from the PASSporT token claims when the full form of PASSporT is received, the verification service MUST treat this as a 438 ‘Invalid Identity Header’ error and proceed as defined above.

### Use of the compact form of PASSporT

For initial SHAKEN deployment, the compact form of PASSporT MUST NOT be used to avoid any potential SIP network element interaction with headers, in particular the Date header field, which could lead to large numbers of 438, Invalid Identity Header errors.

## SIP Identity Header Example for SHAKEN

Draft-ietf-stir-rfc4474bis defines the Identity header field for SIP. It uses the PASSporT token as a basis for creation of the Identity header field for INVITE, MESSAGE, and NOTIFY SIP messages.

The procedure is detailed in rfc4474bis, but an example of an INVITE with an Identity header field is as follows:

INVITE sip:+12155551213@tel.example1.net SIP/2.0
Via: SIP/2.0/UDP 10.36.78.177:60012;branch=z9hG4bK-524287-1---77ba17085d60f141;rport
Max-Forwards: 69
Contact: <sip:+12155551212@69.241.19.12:50207;rinstance=9da3088f36cc528e>
To: <sip:+12155551213@tel.example1.net>
From: "Alice"<sip:+12155551212@tel.example2.net>;tag=614bdb40
Call-ID: 79048YzkxNDA5NTI1MzA0OWFjOTFkMmFlODhiNTI2OWQ1ZTI

P-Asserted-Identity: "Alice"<sip:+12155551212@tel.example2.net>,<tel:+12155551212>
CSeq: 2 INVITE
Allow: SUBSCRIBE, NOTIFY, INVITE, ACK, CANCEL, BYE, REFER, INFO, MESSAGE, OPTIONS
Content-Type: application/sdp
Date: Tue, 16 Aug 2016 19:23:38 GMT
Identity: eyJ0eXAiOiJwYXNzcG9ydCIsImFsZyI6IkVTMjU2IiwieDV1IjoiaHR0cHM6Ly9jZXJ0LWF1dGgucG9jLnN5cy5jb21jYXN0Lm5ldC9leGFtcGxlLmNlcnQifQ.eyJkZXN0Ijp7InVyaSI6WyJzaXA6MTAwMEBzaXB0ZXN0LmNvbWNhc3QubmV0Il19LCJpYXQiOiIxNDcxMzc1NDE4Iiwib3JpZyI6eyJ1cmkiOiJzaXA6NTcxMjIyMzMzM0BzaXB0ZXN0LmNvbWNhc3QubmV0In19.rq3pjT1hoRwakEGjHCnWSwUnshd0-zJ6F1VOgFWSjHBr8Qjpjlk-cpFYpFYsojNCpTzO3QfPOlckGaS6hEck7w;info=<https://cert.example2.net/example.crt>;alg=ES256
Content-Length: 153
v=0
o=- 13103070023943130 1 IN IP4 10.36.78.177
c=IN IP4 10.36.78.177
t=0 0
m=audio 54242 RTP/AVP 0
a=sendrecv