ETSI TS 103 816-4 V1.1.1 (2021-07)

Publicly Available Specification (PAS);
CYBER;
Connecting Products based on MIKEY-SAKKE;
Part 4: Group Voice Communication

CAUTION

The present document has been submitted to ETSI as a PAS produced by Secure Chorus and approved by the ETSI Technical Committee Cyber Security (CYBER).

Reference
DTS/CYBER-0065-4

Keywords
cyber security, mobile, PAS

ETSI

650 Route des Lucioles
F-06921 Sophia Antipolis Cedex - FRANCE

Tel.: +33 4 92 94 42 00 Fax: +33 4 93 65 47 16

Siret N° 348 623 562 00017 - APE 7112B
Association à but non lucratif enregistrée à la
Sous-Préfecture de Grasse (06) N° w061004871

Contents

Intellectual Property Rights

Essential patents

IPRs essential or potentially essential to normative deliverables may have been declared to ETSI. The declarations pertaining to these essential IPRs, if any, are publicly available for **ETSI members and non-members**, and can be found in ETSI SR 000 314: *"Intellectual Property Rights (IPRs); Essential, or potentially Essential, IPRs notified to ETSI in respect of ETSI standards"*, which is available from the ETSI Secretariat. Latest updates are available on the ETSI Web server (https://ipr.etsi.org/).

Pursuant to the ETSI Directives including the ETSI IPR Policy, no investigation regarding the essentiality of IPRs, including IPR searches, has been carried out by ETSI. No guarantee can be given as to the existence of other IPRs not referenced in ETSI SR 000 314 (or the updates on the ETSI Web server) which are, or may be, or may become, essential to the present document.

Trademarks

The present document may include trademarks and/or tradenames which are asserted and/or registered by their owners. ETSI claims no ownership of these except for any which are indicated as being the property of ETSI, and conveys no right to use or reproduce any trademark and/or tradename. Mention of those trademarks in the present document does not constitute an endorsement by ETSI of products, services or organizations associated with those trademarks.

DECT™, **PLUGTESTS™**, **UMTS™** and the ETSI logo are trademarks of ETSI registered for the benefit of its Members. **3GPP™** and **LTE™** are trademarks of ETSI registered for the benefit of its Members and of the 3GPP Organizational Partners. **oneM2M™** logo is a trademark of ETSI registered for the benefit of its Members and of the oneM2M Partners. **GSM**® and the GSM logo are trademarks registered and owned by the GSM Association.

Foreword

This Technical Specification (TS) has been produced by ETSI Technical Committee Cyber Security (CYBER).

The present document is part 4 of a multi-part deliverable covering Connecting Products based on MIKEY-SAKKE, as identified below:

Part 1: "KMS Certificate Definition";

Part 2: "One-to-One Voice Communication";

Part 3: "One-to-One Messaging";

Part 4: "Group Voice Communication";

Part 5: "Discovery".

Modal verbs terminology

In the present document "**shall**", "**shall not**", "**should**", "**should not**", "**may**", "**need not**", "**will**", "**will not**", "**can**" and "**cannot**" are to be interpreted as described in clause 3.2 of the ETSI Drafting Rules (Verbal forms for the expression of provisions).

"**must**" and "**must not**" are **NOT** allowed in ETSI deliverables except when used in direct citation.

1 Scope

The present document is intended to specify the group conferencing interface used for encrypted voice communications. It is intended for use in connecting products based on Multimedia Internet Keying Sakai-Kasahara Key Encryption (MIKEY-SAKKE) domains and to validate products.

2 References

2.1 Normative references

References are either specific (identified by date of publication and/or edition number or version number) or non-specific. For specific references, only the cited version applies. For non-specific references, the latest version of the referenced document (including any amendments) applies.

Referenced documents which are not found to be publicly available in the expected location might be found at https://docbox.etsi.org/Reference.

NOTE: While any hyperlinks included in this clause were valid at the time of publication, ETSI cannot guarantee their long term validity.

The following referenced documents are necessary for the application of the present document.

[1] IETF RFC 3261 (June 2002): "SIP: Session Initiation Protocol". J. Rosenberg et al.

NOTE: Available at rfc3261 (ietf.org).

[2] IETF RFC 3264 (June 2002): "An Offer/Answer Model with the Session Description Protocol (SDP)". J. Rosenberg and H. Schulzrinne.

NOTE: Available at rfc3264 (ietf.org).

[3] IETF RFC 3389 (September 2002):"Real-time Transport Protocol (RTP) Payload for Comfort Noise (CN)". R. Zopf.

NOTE: Available at rfc3389 (ietf.org).

[4] IETF RFC 3550 (July 2003): "RTP: A Transport Protocol for Real-Time Applications". H. Schulzrinne et al.

NOTE: Available at rfc3550 (ietf.org).

[5] IETF RFC 3711 (March 2004): "The Secure Real-time Transport Protocol (SRTP)". M. Baugher et al.

NOTE: Available at rfc3711 (ietf.org).

[6] IETF RFC 3830 (August 2004): "MIKEY: Multimedia Internet KEYing". J. Arkko et al.

NOTE: Available at rfc3830 (ietf.org).

[7] IETF RFC 3966 (December 2004): "The tel URI for Telephone Numbers". H. Schulzrinne.

NOTE: Available at rfc3966 (ietf.org).

[8] IETF RFC 4566 (July 2006): "SDP: Session Description Protocol". M. Handley et al.

NOTE: Available at rfc4566 (ietf.org).

[9] IETF RFC 5727 (March 2010): "Change Process for the Session Initiation Protocol (SIP) and the Real-time Applications and Infrastructure Area", J. Peterson et al.

NOTE: Available at rfc5727 (ietf.org).

[10] IETF RFC 6507 (February 2012): "Elliptic Curve-Based Certificateless Signatures for Identity-Based Encryption (ECCSI)". M. Groves.

NOTE: Available at rfc6507 (ietf.org).

[11] IETF RFC 6508 (February 2012): "Sakai-Kasahara Key Encryption (SAKKE)". M. Groves.

NOTE: Available at rfc6508 (ietf.org).

[12] IETF RFC 6509 (February 2012): "MIKEY-SAKKE: Sakai-Kasahara Key Encryption in Multimedia Internet KEYing (MIKEY)". M. Groves.

NOTE: Available at rfc6509 (ietf.org).

[13] IETF RFC 6665 (July 2012): "Session Initiation Protocol (SIP)-Specific Event Notification". A.B. Roach.

NOTE: Available at rfc6665 (ietf.org).

[14] IETF RFC 6716 (September 2012): "Definition of the Opus Audio Codec". J.M. Valin et al.

NOTE: Available at rfc6716 (ietf.org).

[15] IETF RFC 7714 (December 2015): "AES-GCM Authenticated Encryption in the Secure Real-time Transport Protocol (SRTP)". D.McGrew and K. Igoe.

NOTE: Available at rfc7714 (ietf.org).

[16] IETF RFC 8126 (June 2017): "Guidelines for Writing an IANA Considerations Section in RFCs". M. Cotton et al.

NOTE: Available at rfc8126 (ietf.org).

[17] IANA Registry Multimedia Internet KEYing (MIKEY) Payload Name Spaces.

NOTE: Available at Multimedia Internet KEYing (MIKEY) Payload Name Spaces (iana.org).

[18] IANA Registry Session Initiation Protocol (SIP) Parameters.

NOTE: Available at Session Initiation Protocol (SIP) Parameters (iana.org).

[19] IANA Registry tel URI Parameters.

NOTE: Available at tel URI Parameters (iana.org).

[20] ETSI TS 103 816-1: "Publicly Available Specification (PAS); CYBER; Part 1: KMS Certificate Definition".

[21] ETSI TS 103 816-2: "Publicly Available Specification (PAS); CYBER; Part 2: One-to-One Voice Communication".

[22] ETSI TS 133 179 (V13.9.0) (July 2019): "LTE; Security of Mission Critical Push To Talk (MCPTT) over LTE (3GPP TS 33.179 version 13.9.0 Release 13)".

[23] ETSI TS 133 180 (V15.5.0) (July 2019): "LTE; Security of the mission critical service (3GPP TS 33.180 version 15.5.0 Release 15)".

2.2 Informative references

References are either specific (identified by date of publication and/or edition number or version number) or non-specific. For specific references, only the cited version applies. For non-specific references, the latest version of the referenced document (including any amendments) applies.

NOTE: While any hyperlinks included in this clause were valid at the time of publication, ETSI cannot guarantee their long term validity.

The following referenced documents are not necessary for the application of the present document but they assist the user with regard to a particular subject area.

Not applicable.

3 Definition of terms, symbols and abbreviations

3.1 Terms

Void.

3.2 Symbols

Void.

3.3 Abbreviations

For the purposes of the present document, the following abbreviations apply:

#CS	Number of Crypto Sessions
3GPP	3rd Generation Partnership Project
5G	5th Generation
AEAD	Authenticated Encryption with Associated Data
AES	Advanced Encryption Standard
AES-GCM	AES using GCM
CS	Crypto Sessions
DoS	Denial of Service
ECCSI	Elliptic Curve-based Certificateless Signatures for IBE
ETSI	European Telecommunications Standards Institute
GCM	Galois/Counter Mode
HDR	Common Header payload
IANA	Internet Assigned Numbers Authority
IBE	Identity Based Encryption
IDR	Identity with Role indicator
IETF	Internet Engineering Task Force
IV	Initialisation vector
KMS	Key Management Service
MIKEY	Multimedia Internet KEYing
PIN	Personal Identification Number
PRF	Pseudo-Random Function
PTT	Push To Talk
RFC	Request For Comments
RTP	Real-time Transport Protocol
SAKKE	SAkai-Kasahara Key Encryption
SDP	Session Description Protocol
SID	Silence Insertion Descriptor
SIP	Session Initiation Protocol
SRTP	Secure Real-time Transport Protocol
SSV	Shared Secret Value
tel URI	URI encoding of a telephone number
UAC	User Agent Client
URI	Uniform Resource Identifier

4 Overview

4.1 Overview of Group Connections

The present document pre-supposes adoption of a Vendor Product which is based on MIKEY-SAKKE [12] and makes use of ETSI TS 103 816-2 [21].

Group voice communications based on MIKEY-SAKKE can be established using a group-keyed approach as described in ETSI TS 133 180 [23] However, that standard is for cases such as Push-To-Talk (PTT) in which only a single user is communicating at any time and traffic can be carried encrypted end-to-end (one end to multiple ends).

The present document is intended for group communications such as conference calls, in which multiple users are connected and may be speaking simultaneously. If packets containing voice communications were sent from all user devices in the group to all other user devices in the group, then the bandwidth and processing requirements at each user device would increase linearly with the number of group members, which would not be a scalable solution.

An approach that would lead to a bandwidth scalable solution would be for all users to communicate only with a server and to terminate all encryption at the server, combine the audio at the server, and re-encrypt the combined audio to each user. Note that to avoid echo (a sound or sounds caused by the reflection of sound waves back to the listener), the combined signal sent to a user would exclude its own signal, so each user would receive a different audio signal, and each would be encrypted specifically for that endpoint. This means that although this approach is bandwidth scalable, the processing load on the server would still increase linearly with the number of group members.

However, the most significant limitation on this approach is that the server would be trusted with the unencrypted audio signal. This can be acceptable in some circumstances, but it is not the security model assumed in the present document, and therefore such a server model is not considered here. Instead only the user devices are permitted to have access to the unencrypted voice signals.

The approach defined in this version of the present document only includes user devices. Each group will be managed by a group leader, which is one of those user devices. Some limitations that apply when using only user devices are described in clause A.1.

An evolutionary path to a hybrid system in which some of the group functionality is delegated to a server that does not have access to the unencrypted voice signals is described in clause A.2. This might reduce some of the hardware limitations, the server being assumed to be a more capable device, and also a better- connected device. However, it would introduce some additional scalability issues.

In the approach defined in the present document, signalling between clients, as for one-to-one communication, will use the Session Initiation Protocol (SIP) IETF RFC 3261 [1]. The encrypted communication between user devices that this enables will use the Secure Real-time Transport Protocol (SRTP) IETF RFC 3711 [5], also as for one-to-one communication. SIP describes such a user device, an endpoint of the SIP signalling and of the SRTP encryption, as a User Agent Client (UAC). ETSI TS 103 816-2 [21] and the present document use the term client; the term user will refer to the human user and the term user device will refer to the physical device implementing the client. The standardization issues in using SIP are described in clause A.4.

The present document defines the signalling between clients that this solution shall use; details of implementation in user devices, and the group communication interface that they offer to the user, are outside the scope of the present document, other than the need for the signalling to provide the clients with any required information.

A group can be established as a collection of one-to-one communication links set up as described in ETSI TS 103 816-2 [21]; this approach is suggested in IETF RFC 6509 [12]. In its most basic form, a single client, referred to in the present document as the group leader, establishes all of those one-to-one communication links. For example, Figure 1 shows the connectivity of such a group containing four users, "A" through "D", with "A" acting as the group leader. "A" sets up one- to-one connections with the other clients, "B" through "D" using ETSI TS 103 816-2 [21]. This uses a sequence of MIKEY-SAKKE I_MESSAGEs from "A" to the other clients. Viewed as a conference call initiated by "A", all connections are established by "dialling out".

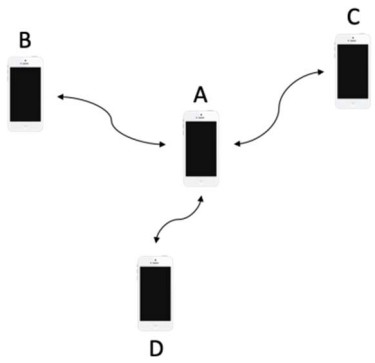

Figure 1: Example single client-initiated group connections

MIKEY-SAKKE [12] describes how such individual one-to-one communications can be linked by using a common Shared Secret Value (SSV) in the MIKEY-SAKKE I_MESSAGEs, and this use of a common SSV is adopted in the present document. Note that this common SSV does not produce common encryption in the group, as each session key used derives from both the SSV and the RAND field in the I_MESSAGE and the latter shall contain a different random value from all other I_MESSAGEs. This is because using a common RAND value would lead to also using a common initialisation vector (IV) in the encryption used by SRTP, something that is specifically prohibited - see the Security Considerations clause of IETF RFC 3711 [5].

The use of a common SSV - the motivation for which is not described in MIKEY-SAKKE [12] - does not help enable group setup, but enables the group member verification that is described below.

With such connectivity, the group leader shall perform the audio signal combining. IETF RFC 6716 [14], which defines the codec used by ETSI TS 103 816-2 [21], refers (in its clause 3.2.1) to *"repacketization by … conference bridges"* so it is expected that this can be performed with acceptable quality.

The privileged position of the group leader in the group allows the user of that device to act as a group host and, if implemented on that device, to use features such as muting other users, for example in Figure 1 sending only client "C" audio signal to client "D". However, the implementation and use of such features is outside the scope of the present document.

The group leader is essential not just to set up the group, but throughout the group's existence. If the group leader becomes disconnected (e.g. due to signal loss on what is expected to be a mobile device) then the group will fail. The evolutionary path suggested in clause A.2 considers a possible mitigation of this possibility.

MIKEY-SAKKE [12] also notes that a group can be established where making an invitation to join the group, and acting as a point of audio combination, is not limited to a single client. Clients that are already part of the group connection can invite in additional clients. An example of how this might be done in principle is shown in Figure 2, in which the group leader "A" invites clients "B", "C" and "D" to join the call, then client "C" invites client "E" to join the call and client "D" invites clients "F" and "G" to join the call.

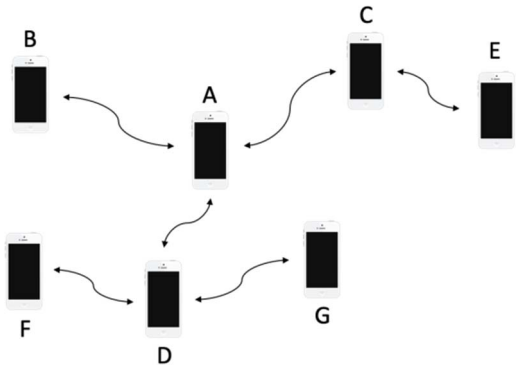

Figure 2: Example multiple client-initiated group connections

However, using this approach introduces additional problems into the group, also described in clause A.1, and is also not consistent with the evolutionary path described in clause A.2. This approach, with what are, in effect, multiple group leaders, is therefore not included in the present document.

The approach of using multiple one-to-one connections is sufficient to create group communications that are secure against eavesdropping. It does not however provide all of the features that are usually required for secure group communications. In particular it is usually required that all clients, and hence their users, know who all the other group participants are, and in this approach only the group leader even knows that it is in a group call.

At a minimum, all clients shall be informed that they are in a group call. This is done by the definition and distribution of a group identity. In order to ensure that this identity is authenticated it needs to be included in the MIKEY-SAKKE I_MESSAGE. However, it cannot replace either of the existing sender or receiver identities in this message, as they are still needed. This identity thus needs to be a new identity both in its role as a group identity and in its format. A new format is needed because the identity format used in ETSI TS 103 816-2 [21] is a tel Uniform Resource Identifier (URI) [7] that includes only a phone number (and a month in which the identifier's key material is valid). However, the group's identity will not be a simple phone number, especially when allowance is made for future extensibility - see clause A.2 and consistency with usual conference call protocols. See clause A.5.

Distribution of a group identity only allows a client to know that it is in a group, which is not sufficient. Usual group communications requirements include that all participants in a group call know who the other participants are; this is the case for ETSI TS 133 180 [23], for example. For a secure auditable system such as Vendor Products, it is expected to be required to provide some form of verification of this group member information. However, it should be noted that this cannot be a complete record of who had access to the call material - a user can always allow an unknown user to join a call, if just by repeating the audio signal.

This distribution of group membership information shall have the following steps:

1) The SIP INVITE message sent by the group leader to indicate that this is a group call in its identity information.

2) Clients that are members of the group will periodically provide verifiable information to the group leader confirming its group membership in a defined format. This formatted information is described in the present document as a *tag*.

3) The received tag is sent by the group leader to all members if the group. The group leader will also send its own tag in the same manner.

The format of the tag is defined in clause A.6.

The information, including the tag, to be distributed by steps 2 and 3 is carried using SIP messages. This means that the following desirable properties are ensured:

- The information is inherently tied to the SIP sessions created by the SIP INVITE messages to each invited client.

- These messages are available to the same auditing processes as the SIP INVITE message, in particular being routed via the SIP server(s) in use.

The SIP messages to be used are SIP SUBSCRIBE and NOTIFY messages, using the SIP SUBSCRIBE/NOTIFY event notification protocol defined in IETF RFC 6665 [13]. The SIP NOTIFY messages are the tag carrying messages; in order to use the standard SIP SUBSCRIBE/NOTIFY protocol, clients shall subscribe to the event so as to receive these SIP NOTIFY messages; this shall be done by using a SIP SUBSCRIBE message. These are sent:

- by each client other than the group leader to the group leader;

- by the group leader to each other client.

Using the usual IP SUBSCRIBE/NOTIFY protocol, the SIP NOTIFY messages sent by a client are periodic, and thus a message interval shall be established. A short message interval - between one and ten seconds - should be used, because:

- It is believed that the bandwidth overhead is acceptable at that rate.

- This interval allows a stateless implementation at the group leader, other than the need to record all clients from which a SIP SUBSCRIBE message has been received, which should be all clients with which it is one-to-one communication with and will include no other clients. On receiving a SIP NOTIFY message the group leader will forward it to all clients that have subscribed to that event, including itself, and then discard it. The delay before a client, and its user, is able to know what other clients are members of the group that it has just joined is no more than the message interval.

- This interval might permit a client to avoid the need to send silent SRTP packets in order that the group leader considers that the client is still active. Note that the usual need to send "comfort noise" when silent described in IETF RFC 3389 [3] is not desirable in a group call, except by the group leader if it identifies that all clients, other than the one being sent combined voice to, are silent.
- This interval represents the granularity of group membership information available to an auditing process, see clause A.3 and the interval suggested above is believed to be as fine a granularity as is likely to be useful.
- This interval makes the need for a SIP NOTIFY message to report a group member leaving unnecessary, this will be identified by the cessation of SIP NOTIFY messages. If a group member temporarily leaves the group - e.g. through movement into a signal shadow - and then returns, then it will naturally be reported as again being a group member.

For details of the contents of these messages see clause A.7.

5 Criterion

5.1 Initialization

All the initialization requirements that shall be used for one-to-one communications as defined in ETSI TS 103 816-2 [21] shall also be used for the present document. This includes provisioning of all clients by a suitable Key Management Service (KMS) , and that clients are registered with a SIP server, both being configured as described in ETSI TS 103 816-2 [21].

A user device that is capable as acting as a group leader shall have additional capability both to set up the multiple one-to-one connections and to act appropriately during the voice session.

5.2 Profiles

Use of SIP, Session Description Protocol (SDP) and SRTP shall follow the profiles described in ETSI TS 103 816-2 [21]. Use of SIP compared to that specification shall be extended by the use in the present document of SIP SUBSCRIBE and NOTIFY messages defined in IETF RFC 6665 [13].

The SIP INVITE message described in ETSI TS 103 816-2 [21] can be authenticated because it carries a signed MIKEY-SAKKE I_MESSAGE. However, direct authentication of this and other SIP messages is not described in ETSI TS 103 816-2 [21]. SIP offers a number of alternative approaches to authentication described in IETF RFC 3261 [1] and updates, SIP messages - including the added SIP SUBSCRIBE and NOTIFY messages - should be authenticated as would be implemented using ETSI TS 103 816-2 [21], but no approach is mandated by the present document. Note that IETF RFC 6665 [13] indicates that failure to authenticate SIP SUBSCRIBE/NOTIFY messages provides a Denial of Service (DoS) opportunity to an attacker.

5.3 Initiating a Group Call

A client that initiates a group call, the group leader, shall do so as a sequence of separate one-to-one connections as described in ETSI TS 103 816-2 [21], modified as described below. Each connection proceeds as shown in Figure 3 of the present document (which is the same as Figure 2 of ETSI TS 103 816-2 [21]).

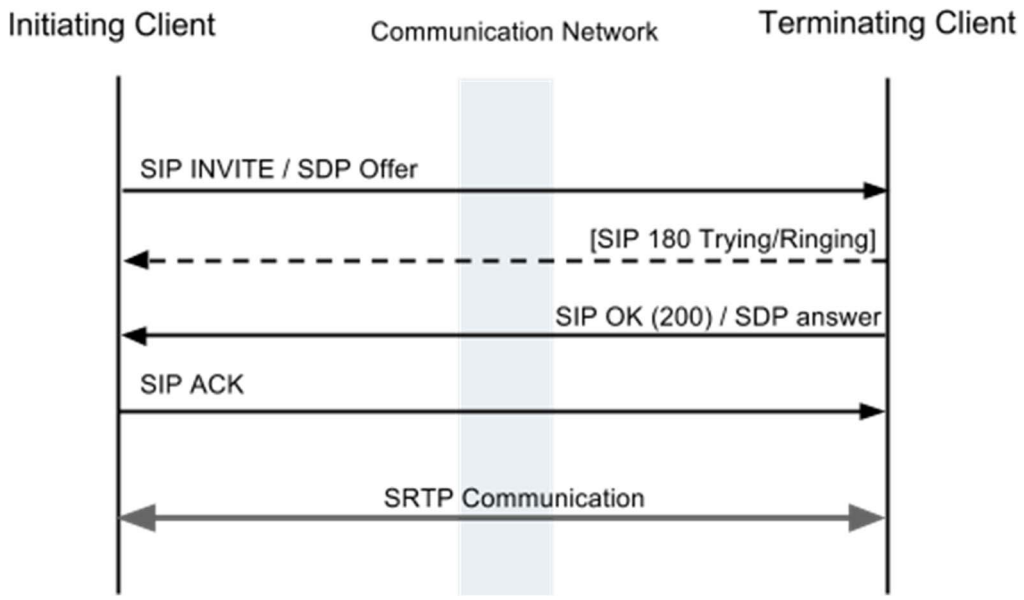

Figure 3: Setup procedure for client-to-client communications

The separate one-to-one connections may be set up sequentially or in parallel, and may be delayed, for example when a user joins a group call late. To set up the connections in parallel, SIP INVITE messages (containing MIKEY-SAKKE I_MESSAGEs) may be sent to multiple other clients before a SIP OK is received. To set up the connections sequentially the group leader may wait until after the SIP ACK is sent for the previously set up one-to-one connection. When delayed, for whatever reason, a further SIP INVITE may be sent during the period shown as SRTP communication in Figure 3.

The group leader shall also act differently to when it is included in a single one-to-one connection during the period of SRTP communication shown in Figure 3 both when receiving and when sending SRTP packets.

During that SRTP communication period the group leader sends and receives SRTP packets to and from multiple other clients. Each received packet shall first be separately decrypted. As such audio signals can be received from more than one other client, but a single audio signal is sent to each other client, the group leader device shall decode the received audio signals using the codec in use ETSI TS 103 816-2 [21]. It then shall combine these linearly encoded audio signals, including its own and received signals, for sending to each other client, encoding the combined audio signal using the codec in use ETSI TS 103 816-2 [21], and then encrypting using the SRTP session key for the client being sent to. Each other client shall be sent a different audio signal, because a client shall not have its own signal returned to it. Some clients may not send an audio signal, because they are silent, and the combined audio signal may thus include a variable number of received signals for each send packet.

5.4 Modification of One-to-One Messages

Reporting that a required one-to-one connection is part of a group call shall be in the SIP INVITE message that initiates the connection, and this is done in the MIKEY-SAKKE I_MESSAGE that is included. That message shall include an additional IDentity with Role Indicator (IDR) payload that reports the group identity as described in clause A.5. The tel URI included in this group identity is that of the group leader. This is an extension of the requirements for what to include in the session description for one-to-one communication described in ETSI TS 103 816-2 [21].

5.5 Group Membership Notification

The SIP SUBSCRIBE and NOTIFY messages required by the present document are for the SIP event type defined in clause A.4. The message formats shall be as defined in clause A.7.

All group members, other than the group leader, shall send a SIP SUBSCRIBE message to the group leader after responding, with a SIP OK message, to a SIP INVITE message from the group leader.

After the group leader has sent a SIP INVITE message to and received a SIP OK message in response from a group member, then after sending a SIP ACK message, the group leader shall send a SIP SUBSCRIBE message to that group member.

All clients, including the group leader, shall send regular SIP NOTIFY messages to all other clients that have subscribed to this information. For the group leader this should be to all other clients in the group and no others, for all other clients this should be only to the group leader.

When the group leader receives a SIP NOTIFY message from another client it shall forward a copy of that message to all other clients that have subscribed to that information - which should be all other clients in the group other than the client from which that SIP NOTIFY message was received and no others.

All SIP SUBSCRIBE and NOTIFY messages shall be acknowledged as usual with a SIP OK message and handled as usual if that SIP OK message is not received. These SIP OK messages do not require a SIP ACK message.

In this manner, all group members that have properly subscribed to this information will receive a SIP NOTIFY message, and hence the tag that it contains, for each other member of the group. The group leader may - and auditing rules might REQUIRE that - terminate the group membership of a client that does not provide timely SIP NOTIFY messages, as such a client is attempting to be a member of the group without leaving an audit trail of its agreed membership.

The sending of SIP NOTIFY messages is periodic, with a message interval defined in the SIP SUBSCRIBE message. The group leader shall specify a suitable period for this interval, a figure between one and ten seconds should be used, these bounds might be tightened further in a later version of the present document. All clients shall send periodic SIP NOTIFY messages respecting the required information validity time. They should not send such messages significantly faster than the indicated rate, but should allow for any possible problems that might delay such messages.

The exchange of these SIP messages (INVITE, SUBSCRIBE, NOTIFY, OK and ACK) in a group containing a group leader "A" and two other clients, "B" and "C", is shown in Figure 4. The message ordering shown is not the only one possible, in particular this figure shows that although "B" and "C" are invited by "A" to join the group at the same time, "C"'s response is delayed until after the initial dialogue with "B" is complete. The messages shown by dashed lines are produced on a periodic schedule, not triggered by message reception. It is assumed that "B" is the first client to join the group and thus "A"'s first SIP NOTIFY message shown is its first in the group and establishes that schedule. "C" will also establish a schedule, but that is not shown in Figure 4 as its first SIP NOTIFY message reaches all other group members. The SIP SUBSCRIBE messages are shown for each other client as from "A" before to "A", but this could be in the opposite order.

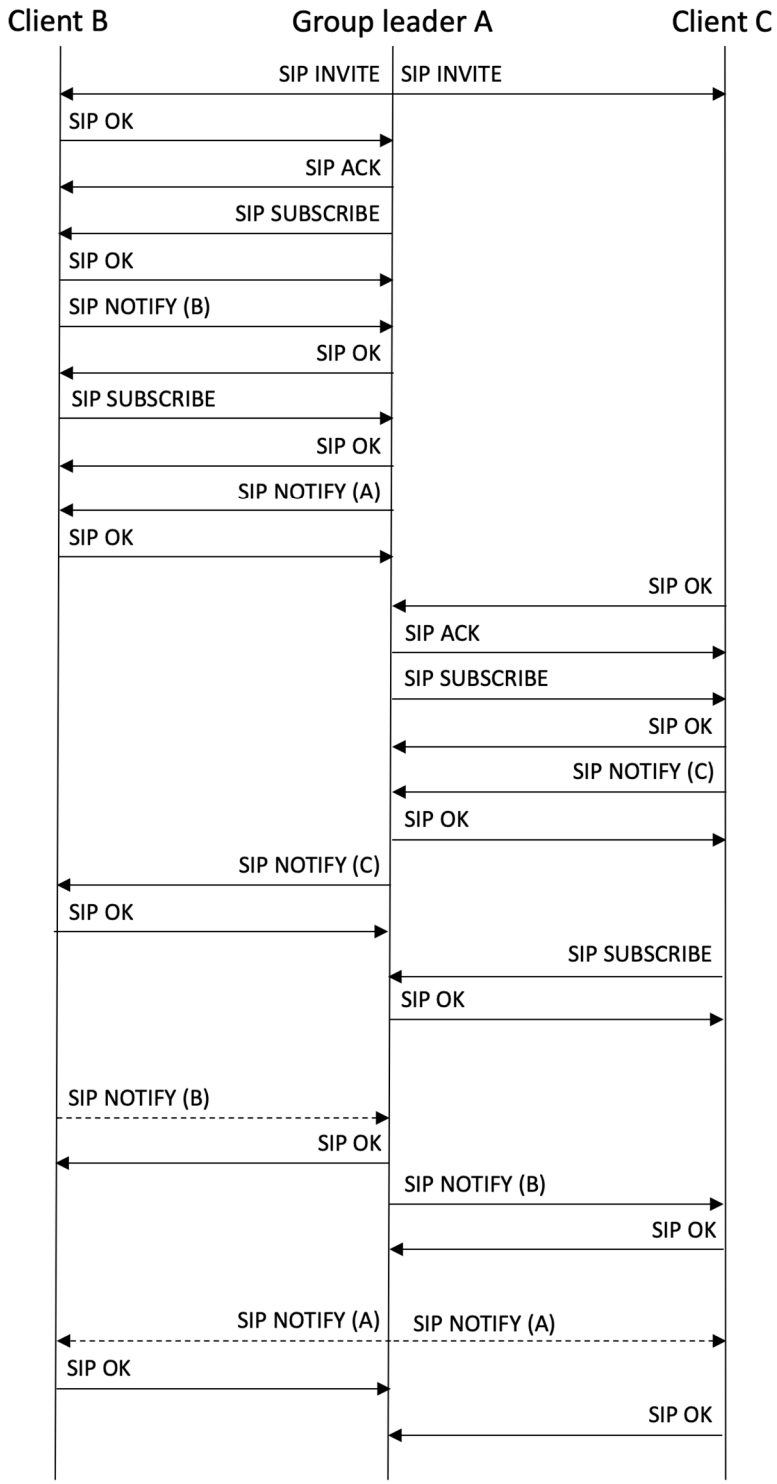

Figure 4: Setup procedure for group with member information exchange

Each received tag shall be authenticated as per clause A.6. If that authentication fails, then the message containing the tag shall be discarded and no action taken due to its receipt. Other information in the tag should be validated as far as possible - for tags received by and from the group leader this extends to most of the information in the tag, which shall match the corresponding SIP INVITE message that the group leader sent. The group leader may take action to exclude any group member on the basis of invalid information in an authenticated tag.

How to use that group membership information, in particular how to make that information visible to the user of the client device, is outside the scope of the present document. See also clause A.3 for how this information might be of value to any auditing process.

Annex A (normative): Implementation considerations and formats

A.1 Hardware Issues

Some limitations of a client-based approach to a group voice call arise from the hardware on which the client is implemented - here considered to include built in and unchangeable software such as the operating system. This hardware is expected to be a smart phone or similar device such as a tablet.

These limitations - some of which would also apply to a server-based approach, as indicated - include:

- Processing power - a modern smartphone is a very powerful device, but what is being requested of it is significant, and can only occupy some of the smartphone processing power while other applications and other software will also have to run. Not all smartphones that can be used will be cutting edge devices.
- Bandwidth - each voice call requires the bandwidth of a one-to-one connection. This is not great compared to e.g. a video call, but phones are not always used where available bandwidth is at its best.
- Delay - the delays of the simplest group call is that of two one-to-one connections, plus any additional central processing. What can be the most important factor in delay, the voice codec frame duration, is required twice even in the best circumstances; it can be worse if voice frames are not synchronized, and there is no reason to assume that they would be.
- Availability - mobile phone signals drop-out, especially in less well provisioned areas or in situations such as on trains. This is also the case for one-to-one calls, and thus there is no additional problem in this regard, except drop-outs by the device implementing the group leader. This is a single point of failure for the group as a whole, and thus for multiple users, who also may not immediately know that this group failure has occurred, although they should do so shortly due to the loss of all SIP NOTIFY messages. Re-establishing the group can require starting group establishment again - no alternative mechanism for group recovery is defined in the present document.

Adoption of the topology shown in Figure 2 - which is not implemented in the present document - would introduce additional problems. It would increase the delay of the voice signals. It would add additional voice codec decoding and encoding, which would have an impact on voice quality. It would introduce a problem of knowing how to balance different voice signals received.

A.2 Evolutionary Path

This clause describes a possible evolutionary path for the present document. There is no commitment that the present document will be developed in this manner, and not all details of this evolutionary path are fully developed in this version of the present document.

Relying on the continued availability of the group leader, which is expected to be a mobile client, is a weakness in the approach defined in the present document. A more reliable approach would be to require only that the group leader is present at the creation of the group or when another client joins the group, and not to require the group leader to be present at all times. Instead, a server would be present at all times, and all clients other than the group leader would communicate only with this server. The server would be assumed to be more reliable due to being fixed and mains powered or to be using some other reliable means of implementation.

The server could not however perform all the functions of the group leader, as that would require it to be fully trusted with the decrypted voice signals, which is contrary to the security model assumed for the present document, see clause A.3. Voice combination would therefore have to be on client devices. However, a simple implementation of such an approach would not be scalable as the bandwidth requirement from the server to each client would increase linearly with the number of clients in the group, and the work required by each client, not just the group leader, would similarly increase.

In order to make this solution scale better than that, note that in a usual conference call not all users speak at once - if they do so then the conference call becomes unmanageable. Instead a small number of users - typically no more than two or three - should be active at any time. If unnecessary voice data could be suppressed and not sent to all clients, then the group call should be reasonably scalable.

Rules for SRTP in such cases are those for Real-time Transport Protocol (RTP), IETF RFC 3550 [4], of which SRTP is defined as a profile IETF RFC 3711 [5]. IETF RFC 3389 [3] states that "RTP allows discontinuous transmission (silence suppression) on any audio payload format. The receiver can detect silence suppression on the first packet received after the silence by observing that the RTP timestamp is not contiguous with the end of the interval covered by the previous packet even though the RTP sequence number has incremented only by one."

It is however possible that it may be preferred to continue to send SRTP packets during a period of silence - which is likely be longer for an individual client than is usual in a one-to-one voice call - in order to indicate that the client is still present on the call. Alternatively, this function is provided by the SIP NOTIFY messages, provided these are sufficiently frequent. If packets that represent silence are sent, then for scalability the server would have to recognize those packets and to not forward them to clients. A means that would allow the server to identify these packets has not yet been identified, and thus consideration would have to be given to disallowing the sending of such silent packets.

Note that the use of "comfort noise" providing a Silence Insertion Descriptor (SID) defined by IETF RFC 3389 [3] as an option to recognize silence is not applicable as the SID appears in the encrypted packet payload, not in the header as would allow simple identification; also note that the use of a SID as comfort noise would prevent the scalability recovery that is the purpose of this approach, and is not needed with multiple speakers.

Dropping packets and, more seriously, forwarding packets without verification in each case would open up the server to a potential DoS attack. It would be desirable to verify SRTP packets at the server to limit the effectiveness of this attack. The original design of SRTP separated the encryption key from the authentication key. Even though both keys would be derived from the same MIKEY-SAKKE SSV, the cryptographically strong hash-based Pseudo-Random Function (PRF) used would make either unpredictable from the other. This would allow the server to authenticate but not decrypt.

However, the one-to-one standard ETSI TS 103 816-2 [21] adopted a later approach to SRTP security defined in IETF RFC 7714 [15], using a combined authentication and encryption process, the Galois Counter Mode (GCM) of AES (Advanced Encryption Standard). This mode cannot be authenticated separately from being decrypted, and allowing the latter would break the security model defined in clause A.3.

Thus, currently, no mitigation of this DoS attack is apparent. It can be noted that the effect of a single spoofed SRTP packet sent to the server is to require all clients to decrypt and fail to verify that packet. This is a per constant linear multiplication of effort by the system compared to the attacker, but this effort is distributed across that many clients. This attack also uses client bandwidth, which is expected to be more valuable than server bandwidth.

In order to maintain the common SSV used in the group, required for verifiable group member information, the SIP INVITE message carrying the MIKEY-SAKKE I_MESSAGE that includes the encrypted SSV needs to be sent from the group leader to the server and then forwarded to the other clients. The server is used in order that all communication to and from clients is from and to the server.

However, this is not sufficient. Each client is now receiving the SRTP stream from all other clients and needs to be able to decrypt that stream. The client knows the SSV that has been used to establish the SRTP encryption (and authentication) key, but that key is also dependent on the RAND field in the SIP INVITE message used to create the connection. Each client therefore needs to know the RAND field used by each other client.

One possible way in which this could be distributed is in the SIP NOTIFY messages sent by each client, including by the group leader. This is possible because the RAND field is carried in a MIKEY payload, and thus could be added to the tag, see clause A.6 - "Tag Format". However, that tag already includes a RAND payload for its own purposes. The issues involved in carrying two RAND payloads and distinguishing between them have not been considered in this version of the present document.

Inclusion of that RAND field in the tag, however done, would mean that the SIP NOTIFY message interval is now a delay before which a client can effectively join a group. The RAND field is thus transmitted more than once, but it is only used once, so does not break the relevant security requirement. Note that a client is now sending SRTP packets to the server, encrypted with its own key, and receiving SRTP packets forwarded from the server from other clients, using those other clients' keys. The original source of the SRTP packets received from the server therefore needs to be known.

In order that the group leader can send that SIP INVITE message it needs to know who the other group participants are. This could be by prearrangement, when the group is defined, typically by the user whose client is the group leader. Note that the equivalent of that arrangement is used for a one-to-one call. Because of the inclusion of the group identity described in clause A.5 in the MIKEY-SAKKE I_MESSAGE, the other client would know the server's tel URI, and hence its phone number.

Alternatively, and fitting a more common use of a conference call, the invited clients could send another, as yet unidentified, suitable SIP message to the server to "dial-in" to the call. The server would forward that message to the group leader; the server would acknowledge that SIP message as usual.

In this case, the server should verify that dial-in message; possibly using a simple whitelist of expected callers plus a Personal Identification Number (PIN)based mechanism, as for conventional conference calls, or something more secure may be needed, depending on the identified threat model. (Note that even if a weak PIN-based approach is used, an attacker in possession of the PIN could not enter the call and receive voice information, it could only attempt to disrupt the call.)

A.3 Security Model

The present document assumes that all clients are configured with a certificate for their security domain and private key material, both provided by their KMS. They may also receive certificates for other security domains. The means by which those certificates and key material are received is outside the scope of the present document, which assumes only that clients have that information and can trust it. Clients are trusted not to divulge their private key material to any other parties. A KMS is trusted not to divulge either its master secret or client private information - except possibly for auditing purposes that are beyond the scope of the present document. The strength of the cryptographic processes used (including SAKKE, ECCSI and AES) is assumed to be sufficient to thwart all direct attacks when not in possession of the corresponding key material. These assumptions are all the same as those for one-to-one communications [21].

To enable the formation of a group call controlled by a group leader, the group leader shall be trusted by all clients in the call to perform the required functions of group initiation and voice data mixing and distribution. The group leader is trusted with all decrypted voice data but in this regard does not have a privileged position among clients because all clients in the group are so trusted.

All clients are capable of distributing that voice data further than is intended, i.e. to other parties than group members. This cannot be fully prevented; for example, simply putting a phone into "speaker" mode - and possibly then relaying via a further unconnected phone - is a form of unintended distribution. Clients have to be trusted not to take such actions - although it is possible that design of client software implementing the present document may constrain some actions by users. This also is the same as for one-to-one communications.

As well as constructing a group call using multiple one-to-one connections, the present document includes messages for the purpose of distributing group member information. Using a common SSV for all one-to-one connections enables group member information that is signed - in an unforgeable manner given the assumptions in the first paragraph of the present clause - by the group member. Use of periodic SIP NOTIFY messages - which are each timestamped - sent by each client provides evidence that the client continued to participate in the call for as long as those messages were being sent. These messages cannot however prevent the group leader from taking actions such as blocking any distribution of material from any group member, the group leader shall be, as noted above, trusted to act properly in that regard. Any client that becomes unsatisfied with the behaviour of the group leader or other clients should leave the group, which will be apparent to any auditing process.

Given that reliance on the group leader, why not trust the group leader to sign the group member messages? This might be acceptable to many group members, although some might prefer the added confidence of endpoint signatures. But in addition, such signalling is also intended for the benefit of auditing processes. The adopted approach including signed periodic SIP NOTIFY messages means that a group member cannot repudiate that it was a member of the group at any times that it sent those messages. A group leader may take action to ensure that a client that is not sending such messages is excluded from the group in order that this audit trail is always present. To verify what voice material is transferred would require auditing the SRTP streams.

The use of a common shared secret, originally suggested in IETF RFC 6509 [12], does not compromise the security of the group. While the present document does not support it, it is possible that some private information could be transferred using the established SRTP keys between two clients (most conveniently if one is the group leader). Even with a common SSV, no third client can decrypt that information - or any packets not intended for it in normal use.

This is because the SRTP keys (or a single key when using AES-GCM as defined in ETSI TS 103 816-2 [21]) are created using a PRF whose inputs include not only the SSV but also a nonce (number used once) carried in the RAND payload of the initiating MIKEY-SAKKE message, and all RAND payloads for the present document (and other standards) shall be unique, not repeated to other clients.

The use of periodic SIP NOTIFY messages that contain group members' identities and RAND values does not compromise the security of the group in an absolute sense because all of that information is already sent unencrypted in the SIP INVITE message that created the group. It does make an eavesdropper's job easier because that information continues to be sent throughout the duration of a call, not just once. The eavesdropper is assumed to be collecting group membership information in order to associate users, it is not able - under the assumptions of the security of SAKKE and AES and proper key handling - to decrypt any information.

The issues that would be involved in encrypting the tag carried in the SIP NOTIFY message to limit this eavesdropping advantage have not been fully considered, and such encryption is not defined in the present document. Note that fully encrypting the tag would prevent the approach for distributing RAND data in the tag suggested in clause A.2, as without that RAND data there is no established symmetric key to decrypt the tag. That approach would thus require that the RAND data not be encrypted - but it would still need to be authenticated. It is possible that use of a suitable AEAD (authenticated encryption with associated data) algorithm could be used, the RAND information being the associated data, which could be used (together with the SSV, without which this approach is obviously flawed) to establish the symmetric key to decrypt and authenticate the tag data and authenticate the RAND information. Such an approach would need a careful review of its security properties.

If the present document is developed further as described in clause A.2 to include a server, then the intent is that the server is not fully trusted. It would be trusted to perform its defined functions of distributing SIP messages and SRTP packets, but not with the unencrypted contents of voice data. A badly-behaved server could thus compromise the availability, but not the confidentiality, of the group call. In this way the server could possibly be a third-party service. It could be trusted with authentication keys, but as clause A.2 notes, use of AES-GCM for SRTP packets means that these are inseparable from decryption keys, which it is not trusted with. A mechanism to prevent a denial of service attack on such a server has not been identified.

A.4 SIP Standardisation Issues

The SIP SUBSCRIBE/NOTIFY process described in clause A.7 shall use a SIP Event Package, and no existing SIP Event Package is suitable. The present document defines a new SIP Event Package. Note that there is no private space in this registry that can be used; instead the new SIP Event Package has been defined outside the Internet Engineering Task Force (IETF) and IANA Internet Assigned Number Authority (IANA approval process described in IETF RFC 5727 [9].

Those approval processes are at the most difficult end of the range of requirements that IANA registrations require, including but not limited to requiring an RFC (not just an Internet Draft or an external standard). As of this date, fifteen such events have been defined in five RFCs - see the "Header Field Parameters and Parameter Values" sub-registry in the IANA registry "Session Initiation Protocol (SIP) Parameters" [18]. It is possible that the use of SDP in the Event Package could be a complicating factor were IETF/IANA approval to be requested.

The SIP Event Package that is defined by the present document is given the name **MIKEY-group-tag**. The behaviour of this event is that in response to a SIP SUBSCRIBE message for this event, a client shall send a SIP NOTIFY message for each client whose identity it is responsible for reporting. For a group leader this is all members of the group, for any other client that is just itself.

The manner in which that information is included in the SIP NOTIFY message is defined in clause A.7.

A.5 Group Identity

User identities for Vendor Products for use in one-to-one communication are defined in ETSI TS 103 816-1 [20]. Two options are supported, one as defined for MIKEY-SAKKE [12] using a tel URI [7], and one defined in an earlier version of ETSI TS 133 179 [22]. The present document only considers an evolution of the former option.

The identity defined in IETF RFC 6509 [12] is based on the tel URI defined in IETF RFC 3966 [7], preceded by a monthly timestamp. This form of identity is retained for users and KMSs, but a new form of identity shall be used for the group. Tel URIs have the option of including parameters. These are however explicitly disallowed by IETF RFC 6509 [12] with the comment: *These constraints on format are necessary so that all parties can unambiguously form the "tel" URI*.

For a group identity this constraint is relaxed, to allow specific parameters to be included in the group identity, in an unambiguous manner. Note that this group identity is not a MIKEY-SAKKE identity, there is no cryptographic information associated with it, and it is not known to any KMS, those being factors in adoption the strict rules in IETF RFC 6509 [12] that do not apply to this use in the present document. The monthly timestamp is thus unnecessary, and probably of an unsuitable granularity, and is not included.

The usual identification of a conference call is a combination of a conference call number and a PIN, and the group identity will include the former, but not the latter since that would defeat the point of the PIN as known only to participants. The group identity also needs to include a phone number at which the group can be found. At this point this phone number is that of the group leader but could in the future become that of a server, see clause A.2. With a phone number included, the natural encoding is as a tel URI for that number, with the group call identity completed using an added tel URI parameter IETF RFC 3966 [7] for the conference call number.

There is no suitable tel URI parameter defined in [19], thus the present document proposes a new tel URI parameter **group-identity** for the conference call number or other identifying data. The IANA registration procedure for a new tel URI parameter is Specification Required, a procedure that, as defined in IETF RFC 8126 [16], requires an expert review but can be specified in a non-IETF document and has already included a 3GPP document. Such a specification could possibly include a future version of the present document.

The group identity used by the present document shall include the group-identity tel URI parameter. The value of this parameter shall satisfy tel URI parameter rules IETF RFC 3966 [7] but is otherwise unconstrained. The value of this parameter is selected by, or externally provided to, the group leader. As included in SIP messages, the group identity may include other tel URI parameters, but these other parameters shall be ignored in matching the received identity with that of the group call being formed.

A.6 TAG Format

This annex defines the format of a tag that can be created by a first client that, when distributed to a second client can be used by that second client to verify that the first client, and no other party could have reported that event (other than as leaked by the first client or authorized by its KMS, both of which are assumed not to do that).

The first client is expected to be reporting an event specific to itself, that it was invited into the group and accepted that invitation. However, it is possible that a future use of this tag could allow verification by a different party, in particular by the group leader, and thus that option is supported by this definition.

The two things that make the provision of verifiable event information possible are:

- the common SSV shared with all clients in the group. Without this an attacker could indicate that it is a member of a group into which it has not been invited.
- the ability of any client with known identity, to create an Elliptic Curve-Based Certificateless Signature for IBE - Identity-Based Encryption (ECCSI) IETF RFC 6507 [10] signature that only it could create. Note that it is assumed that a certificate that includes the required public information for that client is available.

The tag shall not include the SSV, which is never sent unencrypted in any communication. The tag is formatted as a sequence of MIKEY payloads, as defined in IETF RFC 3830 [6] and other RFCs that extend it. Payloads such as T (timestamp) and RAND shall be constrained as specified in ETSI TS 103 816-2 [21]. The format allows the reported group member and the signer to be different, however these shall be the same for all messages defined in the present document.

The following payloads shall be used; other MIKEY payloads may be included:

- Common Header payload. A new data type is defined for this use, see clause A.8. This payload should avoid including unnecessary crypto session information; this can be done by including a #CS (number of Crypto Sessions) field containing zero.

- An IDR payload reporting the identity of the group.
- An IDR payload reporting the identity of the group member.
- An IDR payload reporting the identity of the signer.
- An IDR payload reporting the identity of the signer's KMS.
- A timestamp (T) payload.
- A RAND payload. This shall be unique to this message, not reused.
- A signature (SIGN) payload. This shall be last. It is the signature of data that consists of the other payloads, concatenated with the unencrypted SSV, formatted as defined in IETF RFC 6508 [11]. The signature method shall be ECCSI. This assumes that the nature and format of the SSV is known, even if the use of SAKKE is later replaced in an evolution of ETSI TS 103 816-2 [21], because that format is not recorded in the tag. That information is available to all clients from the earlier SIP INVITE message, which it shall match. Note that the requirement to use side channel information (the SSV) that is not in the tag to sign the tag is inherent in the concept and usefulness of this tag.

A.7 SIP SUBSCRIBE/NOTIFY Messages

The SIP SUBSCRIBE/NOTIFY model described in IETF RFC 6665 [13] provides a message flow pattern that is appropriate to the transfer of group membership information.

Using this model, the first client uses a SIP SUBSCRIBE message to inform the second client which information it requires, and the second client sends a SIP NOTIFY message when an event occurs that requires notification. Each is acknowledged with a SIP OK message. The expectation of this model, which is followed here, is that information in a SIP NOTIFY message has a validity time, before the expiry of which a new SIP NOTIFY message is sent. Use of the model by the present document includes the group leader as both first and second client.

Each SIP SUBSCRIBE message will result in multiple SIP NOTIFY messages, both because it requests notification from each other group member and because the SIP NOTIFY messages are sent periodically.

SIP Event types are described in, currently, five RFCs, see [18]. None are applicable to this requirement, and a new SIP Event type is defined, see clause A.4.

The SIP SUBSCRIBE message identifies the information that it requires through this event type and a Request URI. The Request URI is defined as the group identity, which is a URI, see clause A.5.

The format of the NOTIFY message body is either specified by an Accept header in the SUBSCRIBE message or is a default for the event type; the latter is used in the present document. The message body carries a tag as defined in clause A.6. The tag is a modified version of a MIKEY I_MESSAGE and is included in the SIP NOTIFY message in the same way that a MIKEY I_MESSAGE is included in a SIP INVITE message, by using SDP (Session Description Protocol) IETF RFC 4566 [8]. The details of this inclusion shall follow the rules defined in ETSI TS 103 816-2 [21] for the use of SDP.

Note that this, like the use of SDP to carry a MIKEY-SAKKE I_MESSAGE in a SIP INVITE message, is using SDP as a format by which information can be carried in a SIP message, not as part of an SDP session as described in IETF RFC 3264 [2]. No existing SIP NOTIFY message uses SDP (see the RFCs referenced for event types at [18]) but this should not be a technical obstacle to this use; however, see also clause A.4.

A.8 MIKEY Standardisation Issues

A group identity is included in a MIKEY I_MESSAGE in order that the identity is verified and cannot be modified. This has required three new things to be defined for use in a MIKEY IDR payload:

- A group identity role number. Existing roles are standardized in the ID Role registry at [17].
- An identity scheme number for the group identity. Existing schemes are standardized in the ID Scheme registry at [17].

- The group identity format. See clause A.5.

There is a private use number range in each of the first two cases that can be used temporarily, and in the longer term those registries have a registration process of "Specification Required", which as defined in IETF RFC 8126 [16] requires an expert review, but this can be specified using a non-IETF document and those registries already include allocations that were made in a version of ETSI TS 133 180 [23].

In addition, a new MIKEY common Header payload (HDR) data type shall be used, to report the reason for this message: that a group member is present. This data type also has a private number range that can be used temporarily.

The proposed numbers to be used from private spaces are ID Role: **254**, ID Scheme: **254**, Common Header Payload Field Names: **255**.

NOTE: The first two of these number ranges reserve (rather than making private) the value 255, which therefore cannot be used.

History

Document history		
V1.1.1	July 2021	Publication