3GPP TR 22.873 V18.0.0 (2021-06)

Technical Report

3rd Generation Partnership Project;

Technical Specification Group Services and System Aspects;

Study on evolution of IMS multimedia telephony service

(Release 18)

The present document has been developed within the 3rd Generation Partnership Project (3GPP TM) and may be further elaborated for the purposes of 3GPP.
The present document has not been subject to any approval process by the 3GPPOrganizational Partners and shall not be implemented.
This Report is provided for future development work within 3GPPonly. The Organizational Partners accept no liability for any use of this Specification.
Specifications and Reports for implementation of the 3GPP TM system should be obtained via the 3GPP Organizational Partners' Publications Offices.

Keywords

<keyword[, keyword]>

***3GPP***

Postal address

3GPP support office address

650 Route des Lucioles - Sophia Antipolis

Valbonne - FRANCE

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

Internet

http://www.3gpp.org

***Copyright Notification***

No part may be reproduced except as authorized by written permission.
The copyright and the foregoing restriction extend to reproduction in all media.

© 2021, 3GPP Organizational Partners (ARIB, ATIS, CCSA, ETSI, TSDSI, TTA, TTC).

All rights reserved.

UMTS™ is a Trade Mark of ETSI registered for the benefit of its members

3GPP™ is a Trade Mark of ETSI registered for the benefit of its Members and of the 3GPP Organizational Partners
LTE™ is a Trade Mark of ETSI registered for the benefit of its Members and of the 3GPP Organizational Partners

GSM® and the GSM logo are registered and owned by the GSM Association

Contents

Foreword 5

1 Scope 6

2 References 6

3 Definitions, symbols and abbreviations 6

3.1 Definitions 6

3.2 Symbols 6

3.3 Abbreviations 7

4 Overview 7

5 Use cases 7

5.1 Conference call with AR holography 7

5.1.1 Description 7

5.1.2 Pre-conditions 7

5.1.3 Service Flows 7

5.1.4 Post-conditions 8

5.1.5 Existing features partly or fully covering the use case functionality 8

5.1.6 Potential New Requirements needed to support the use case 8

5.2 Real-time speech translation 8

5.2.1 Description 8

5.2.2 Pre-conditions 8

5.2.3 Service Flows 8

5.2.4 Post-conditions 8

5.2.5 Existing features partly or fully covering the use case functionality 8

5.2.6 Potential New Requirements needed to support the use case 9

5.3 AR call 9

5.3.1 Description 9

5.3.2 Pre-conditions 9

5.3.3 Service Flows 9

5.3.4 Post-conditions 9

5.3.5 Existing features partly or fully covering the use case functionality 10

5.3.6 Potential New Requirements needed to support the use case 10

5.4 Real-time screen sharing 10

5.4.1 Description 10

5.4.2 Pre-conditions 10

5.4.3 Service Flows 10

5.4.4 Post-conditions 10

5.4.5 Existing features partly or fully covering the use case functionality 10

5.4.6 Potential New Requirements needed to support the use case 11

5.5 Multimedia CLIP (Calling Line Identification Presentation) and COLP (Connected Line Identification Presentation) 11

5.5.2 Pre-conditions 12

5.5.3 Service Flows 12

5.5.4 Post-conditions 12

5.5.5 Existing features partly or fully covering the use case functionality 12

5.5.6 Potential New Requirements needed to support the use case 13

5.6 3rd party specific user identities 14

5.6.1 Description 14

5.6.2 Pre-conditions 14

5.6.3 Service Flows 14

5.6.4 Post-conditions 15

5.6.5 Existing features partly or fully covering the use case functionality 15

5.6.6 Potential New Requirements needed to support the use case 15

5.7 Real-time visual interactive menu 15

5.7.1 Description 15

5.7.2 Pre-conditions 16

5.7.3 Service Flows 16

5.7.4 Post-conditions 16

5.7.5 Existing features partly or fully covering the use case functionality 16

5.7.6 Potential New Requirements needed to support the use case 17

6 Consolidated potential requirements 17

6.1 Media handling capability 17

6.2 Authorization and authentication 17

Annex A: Change history 18

# Foreword

This Technical Report has been produced by the 3rd Generation Partnership Project (3GPP).

The contents of the present document are subject to continuing work within the TSG and may change following formal TSG approval. Should the TSG modify the contents of the present document, it will be re-released by the TSG with an identifying change of release date and an increase in version number as follows:

Version x.y.z

where:

x the first digit:

1 presented to TSG for information;

2 presented to TSG for approval;

3 or greater indicates TSG approved document under change control.

y the second digit is incremented for all changes of substance, i.e. technical enhancements, corrections, updates, etc.

z the third digit is incremented when editorial only changes have been incorporated in the document.

# 1 Scope

The present document studies new scenarios (e.g. AR/VR communication) for IMS Multimedia Telephony service, identifies potential service requirements, and identifies additional potential requirements on 5G system.

This document also does a gap analysis to existing requirements in 3GPP to identify potential requirements for introduction into 3GPP.

NOTE: Use cases in this document have references to those already specified in 3GPP TR26.918 [2] and TR26.928[3]. This allows an easier breakdown of these use cases to existing features and potential new requirements as well as relevant performance figures. New use cases in this TR are highlighted if they are not covered in 3GPP SA4 work.

# 2 References

The following documents contain provisions which, through reference in this text, constitute provisions of the present document.

- References are either specific (identified by date of publication, edition number, version number, etc.) or non‑specific.

- For a specific reference, subsequent revisions do not apply.

- For a non-specific reference, the latest version applies. In the case of a reference to a 3GPP document (including a GSM document), a non-specific reference implicitly refers to the latest version of that document *in the same Release as the present document*.

[1] 3GPP TR 21.905: "Vocabulary for 3GPP Specifications".

[2] 3GPP TR 26.918: "Virtual Reality (VR) media services over 3GPP ".

[3] 3GPP TR 26.928: "Extended Reality (XR) in 5G ".

[4] 3GPP TR 22.823: " Study on Enhancements to IMS for new real time communication Services ".

[5] 3GPP TS 22.173: " IP Multimedia Core Network Subsystem (IMS) Multimedia Telephony Service and supplementary services ".

# 3 Definitions, symbols and abbreviations

## 3.1 Definitions

For the purposes of the present document, the terms and definitions given in 3GPP TR 21.905 [1] and the following apply. A term defined in the present document takes precedence over the definition of the same term, if any, in 3GPP TR 21.905 [1].

**example:** text used to clarify abstract rules by applying them literally.

## 3.2 Symbols

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

<symbol> <Explanation>

## 3.3 Abbreviations

For the purposes of the present document, the abbreviations given in 3GPP TR 21.905 [1] and the following apply. An abbreviation defined in the present document takes precedence over the definition of the same abbreviation, if any, in 3GPP TR 21.905 [1].

AR Augmented Reality

AI Artificial Intelligent

VR Virtual Reality

XR Extended reality

# 4 Overview

With the evolution and capability enhancement of telecommunication system, telephony service has evolved from CS voice call to IMS multimedia telephony service. Nowadays 3GPP has introduced new network capabilities and new types of devices (e.g. AR/VR devices, robot, etc.), which can bring promising improvements to IMS multimedia telephony service.

XR (an umbrella term for different types of realities, such as AR/VR) technologies have proven themselves to be valuable additions to the fields of digital entertainment, information and workspace.

While more and more individual consumers enjoy MMTEL services across the globe, MMTEL services become popular also among business customers. There are several primary business functions that organizations use MMTEL services for, including internal communication, talking with prospects (sales call), contacting current customers and clients, customer support, and contact centre (or call centre) activities. While business customers consider the MMTEL services offer attractive features to their business, they also experience some practical issues that may benefit from support from the 5G system.

# 5 Use cases

## 5.1 Conference call with AR holography

### 5.1.1 Description

International companies have offices all over the world. The staffs of the company may have conference calls very often. If AR holography can be applied in the conference call, and AR holography data can be transmitted in the conference call session, all the staffs can feel the expression and action of the other parties, or even communicate with each other through body gestures. It just feels like having face-to-face meeting at the same location. It can really improve the immersive experience of the conference.

### 5.1.2 Pre-conditions

User A, user B and user C are the staffs of the same company but in different location. They always have to discuss their product schemes through conference call. Each of their own office has AR holography camera and glasses.

### 5.1.3 Service Flows

1. User A, user B and user C plan to have a conference call to discuss the latest product scheme.
2. Each of them goes to their own office in which there is an AR holography camera.
3. When meeting begins, they all initiate the AR holography camera. The AR cameras scan their whole bodies, then have the data processed in a dedicated device to build up the models of themselves.
4. All the AR holography data and audio/video will be uploaded to the IMS Application Server. The Application Server makes media processing, including synchronizing AR holography data and audio/video, AR rending, etc.
5. After media processing, the IMS Application Server will generate new media streams (i.e. video, audio) and transmit them to each end user’s AR glasses in the conference call.
6. When AR glasses receive the generated media streams, they decode the media streams and present the product. Users can see the AR holography images of the other parties and hear their sound from the smartphone. The images can be updated in real time when conference call is processing.

### 5.1.4 Post-conditions

People in the IMS conference call can see the action and expression of the other parties. They can present the product with their body gestures. They can communicate and interact with each other (e.g., shaking hands, etc.). It just feels like having face-to-face meeting at that moment.

### 5.1.5 Existing features partly or fully covering the use case functionality

Part of requirements for AR is covered in TR22.823 [4] already.

### 5.1.6 Potential New Requirements needed to support the use case

The MMTEL service shall support synchronization between media/data streams belonging to the same MMTEL session.

## 5.2 Real-time speech translation

### 5.2.1 Description

People may need to communicate with others speaking different languages. With AI technology(e.g. Skype translator), it is possible to provide instantaneous voice translation for the calling parties in a call and further perform speech-to- text or speech-to-speech for the translated voice.

### 5.2.2 Pre-conditions

User A and user B speak different languages.

### 5.2.3 Service Flows

1. User A makes an audio or video call to user B.
2. User B chooses the translation service and text presentation. User A’s voice is recognized and translated to user B’s chosen language. Then the translation will be presented as texts or subtitles on user A’s smartphone screen during the call.
3. User A chooses the translation service and audio presentation. User B’s voice is recognized and translated to user A’s chosen language. Then the translation will be presented as audio replacing user B’s original audio during the call.
4. User A and user B are able to communicate in their native languages.

### 5.2.4 Post-conditions

User A can hear the translated audio and user B can see the translated texts on his screens synchronized with the voice and video.

### 5.2.5 Existing features partly or fully covering the use case functionality

Currently in order to support real time voice translation in a MMTEL call, the MMTEL session shall end when the user trigger the translation service and another MMTEL session to the user shall be set up for transmitting translated audio.

The MMTEL service can’t incorporate the translation service within the end-to-end call, just like transcoding.

### 5.2.6 Potential New Requirements needed to support the use case

Editor’s Note: text to be provided.

## 5.3 AR call

### 5.3.1 Description

This use case focuses on the multimedia call, equipped with AR to enrich the user experience and to be more helpful for users. AR call can be used in many areas like: person-to-person communication, emergency call, remote cooperation, and consumer-to-business call.

Take remote cooperation supporting AR as an example, when local technicians have some difficulty in repairing consumers’ cars, they can cooperate with the remote engineer of technical support department via multimedia telephony communication supporting AR. When making the AR call, UEs produce some data needed to perform rendering computing, all or part of which is done in the cloud/edge severs in this use case.

The car technician makes a video call with the remote engineer to get more support. The technician can capture the car parts as video contents and mark possible points of failure in-call, in order to enable remote engineer to understand the problem. During the discussion, both parties can modify the video contents about the car in-call via their UEs, such as marking graphics and overlaying AR models. At the same time, both parties can receive the updated video contents in real time with the continuous modification. It looks like that the remote engineer is beside the technician, discusses and solves the problems together. AR call provides a solution which is less time-consuming and less costly.

### 5.3.2 Pre-conditions

User A is a local technician in a car repair shop and user B is a remote engineer arranged by the technical support department.

### 5.3.3 Service Flows

1. User A has some difficulty in repairing consumer’s cars and hopes to get help from the technical support department by phone.
2. User A makes a video call to user B, then he captures the car being repaired using his phone camera during the call.
3. User A further turns on the AR call function and marks possible points of failure to the captured video contents.
4. The AR data are produced within user A’s UE and these data are sent to the AR cloud/edge server. The cloud server performs media processing (e.g. AR rendering, synchronizing multiple media streams) and produce multimedia. The multimedia is sent to user B’s UE.
5. User B sees the captured video contents related to the car being repaired with user A’s marks via user B’s call screen.
6. In order to demonstrate the repairing procedures, user B deletes user A’s marks and overlays AR models to the video contents captured by user A.
7. User A can see the virtual demonstration, and follows the instructions step by step.
8. User A and user B are sure that the problem is solved, and then hang up the video call.

### 5.3.4 Post-conditions

Both User A and User B can see and modify the marks and AR models on their respectively in-call screen. User A can cooperate with the remote engineer to repair the car well.

### 5.3.5 Existing features partly or fully covering the use case functionality

Traditional video call is fully covered by existing IMS functionalities.

Part of requirements for AR is covered in TR22.823 already.

### 5.3.6 Potential New Requirements needed to support the use case

The IMS multimedia telephony service shall support AR media processing (e.g. MRFC and MRFP).

## 5.4 Real-time screen sharing

### 5.4.1 Description

This use case is about the real-time screen sharing experience in a video call or a video conference. During a video call or a video conference, you can not only view real-time information on each other's screens but also directly perform “hands-on” operations on the shared screen. You can explain to other participants on the shared screen with doodles in real-time as if you were having a face-to-face white-board discussion. For example, you can remotely attend a customer meeting and present the solution to the customers via the shared mobile screen. In addition, you can also mark the key points on the slides displayed on the shared screen.

With the support of IMS network that has been enhanced for the mobile screen sharing, the UE captures the screenshots and sends the data to the IMS network. The IMS network transfers the data to all other parties. After receiving the data, the receiving UEs can adapt the screen resolution and display the data clearly.

### 5.4.2 Pre-conditions

Mr. Anderson is having a multi-party video call with his son Mike and daughter Joan, asking for instructions on how to purchase tickets on the airline website.

All UEs (smart phones, tablets) used in this call are able to capture screenshots, which may have subscriptions with different MNOs.

### 5.4.3 Service Flows

1. Mr. Anderson initiates a video call, inviting his son Mike and his daughter Joan to the call, and asking them for help to purchase tickets online.
2. Mike asks if he could share his mobile screen to show his old father step by step. Mr. Anderson then “requests” Mike to share his mobile screen.
3. Mike accepts the request and starts mobile screen sharing.
4. Mike browses the airline website, selects an itinerary, and purchases a ticket. His smart phone, meanwhile, collects the screenshots and sends the data to the IMS network.
5. The smart phones of Mr. Anderson and Joan receive the screen sharing data and display the data locally.

### 5.4.4 Post-conditions

Both Mr. Anderson and Joan can view the online ticket purchase process on Mike’s screen by using the screen-sharing function.

### 5.4.5 Existing features partly or fully covering the use case functionality

The related existing requirements can be found in TS 22.173[5] “Multimedia Telephony Service and supplementary services”:

*IMS Multimedia Telephony service includes the following standardized media capabilities:*

*- Full duplex speech;*

*- Real time video (simplex, full duplex), synchronized with speech if present;*

*- Real-Time Text communication;*

*- File transfer;*

*- Video clip sharing, picture sharing, audio clip sharing. Transferred files may be displayed/replayed on receiving terminal for specified file formats*

*- Fax;*

*- Data (CS).*

*The support of each of these media capabilities is optional for a UE.*

Real-time screen-sharing is similar to capture a sequence of screenshots and transfer them in real-time and on the receiving party display the screenshots in the right order in real-time.

The MMTEL service supports a user participating in a MMTEL session to deliver to all other participants simultaneously a sequence of files in real-time.

In addition, the features of “flexible data channel handling” cover the required functionality – “Data channels do not require use of any codec but allows for real-time interaction in parallel to the conversational media”. The details of data channel can be found in TS 26.114 “IP Multimedia Subsystem (IMS); Multimedia Telephony; Media handling and interaction”. Alternatively this functionality can also be supported in IMS with MSRP (Message Session Relay Protocol) using RFC 5547 (“A SDP Offer/Answer Mechanism to Enable File Transfer”).

### 5.4.6 Potential New Requirements needed to support the use case

Editor’s Note: text to be provided.

## 5.5 Multimedia CLIP (Calling Line Identification Presentation) and COLP (Connected Line Identification Presentation)

5.5.1 Description

The traditional calling card function provides the calling party identification information to the called party, and it also provides the calling party with the possibility to receive the line identity of the connected party. Such information can be enriched to include a subject, location, picture, and even video clips. There have been on-going efforts across the globe in enhancing the protection against unwanted robocalls, among which STIR/SHAKEN are most well-known solutions to verify the caller, minimizing unwanted robocalls and protecting consumers from fraudulent intentions from bad actors. IETF-defined STIR (Secure Telephone Identity Revisited) is considered one of the most promising technologies to combat unwanted robocalls, while SHAKEN (Signature-based Handling of Asserted information using tokens) framework provides a mechanism to implement STIR to authenticate calls. The STIR/SHAKEN works for authenticating and thus proving that a call has not been spoofed, but it does not determine caller intent. This means that bad actors may continue making unwanted calls by registering telephone numbers, which, as long as registered, are authentically theirs.

Meanwhile in vertical applications (such as hospitality industry and tourism industry), multimedia Calling Line Identification Presentations (CLIP) and multimedia Connected Line Identification Presentation (COLP) can help companies to promote and advertise their business with promotional videos / pictures, office location information, etc. While MMTEL and the supplementary services offers attractive features, the biggest practical issue for lots of businesses is that calls sometimes are rejected as fraudulent robocalls. It would be beneficial to enable business customers to be involved in controlling/managing the business specific information of their users (e.g. the identity of its own employees), which will help to confirm the authenticity of the business specific identity information and to provide additional support in preventing fraudulent robocalls.

### 5.5.2 Pre-conditions

Hotel California has a service contract with an operator iCA, including IMS subscription for the MMTEL services and the supplementary services such as CLIP and COLP.

### 5.5.3 Service Flows

Example 1:

1. Eagles called the hotel reservation hot line to book rooms for the band to stay in Baja California.

2. When connected, they could see the promotional video of the hotel, as well as the identity information to confirm the authenticity of the called number.

Example 2:

1. As part of pre-check-in service, the hotel customer service called Eagles who would check in the next week to learn about the room requirements.

2. When receiving the call, the guests could see the video introduction of the hotel with the available facilities and the scheduled entertainment events, as well as the identity information to confirm the authenticity of the calling number.

### 5.5.4 Post-conditions

When calling the hotel reservation hot line, customers can see the promotional video of the hotel.

When receiving a call from the hotel customer service centre, guests can see the video introduction of the hotel.

### 5.5.5 Existing features partly or fully covering the use case functionality

Note that the general requirements of Calling Line Identification Presentation (CLIP) and (Connected Line Identification Presentation) COLP are defined in TS 22.081 “Line identification Supplementary Services”, and for MMTEL in TS 22.173, clause 8.2.1 “Originating Identification Presentation (OIP)” and clause 8.2.3 “Terminating Identification Presentation (TIP)”.

The existing service requirements relate to the use case functionality can be found in TS 22.183 “Customized Ringing Signal (CRS) Requirements” clause 4.1 “Basic functionality of CRS service”:

*The service shall support the CRS content types of audio, video, image, graphic, text, location, and electronic business card.*

*CRS can contain one or several content types, and at most one element of each type.*

*The called party should be able to simultaneously experience multiple CRS contents, e.g. audio and video.*

Rich Communication Services (RCS) defined in GSMA partially cover the required use case functionality, where the call composer functionality enables caller to enrich session establishment by sending e.g. reason, priority, location and other data to called party.

In addition, the features of “flexible data channel handling” partially cover the required use case functionality. The details of data channel can be found in TS 26.114 “IP Multimedia Subsystem (IMS); Multimedia Telephony; Media handling and interaction”:

*Data channels do not require use of any codec but allows for real-time interaction in parallel to the conversational media (see clause 6.2.10)...*

*The data channel application is created prior to the DCMTSI call where it is intended to be used, by means left out of scope for this specification. The data channel application workflow is depicted by Figure 6.2.10.1-1 below.*



*Figure 6.2.10.1-1 Data Channel Workflow*

Figure 5.5.5-1: copy of Figure 6.2.10.1-1 of TS 26.114 on Data Channel Workflow

The existing features related Caller ID authentication are STIR/SHAKEN Standards. STIR (Secure Telephone Identity Revisited) provides the ability within SIP to authenticate Caller ID, and SHAKEN (Signature-based Handling of Asserted information using tokens) defines the end-to-end architecture to implement Caller ID authentication using STIR in the telephone network. Note that the related work for STIR/SHAKEN support in 3GPP systems is being referred to as caller ID attestation in the downstream WGs (i.e., CT work item on enhancements to call spoofing functionality eSPECTRE).

### 5.5.6 Potential New Requirements needed to support the use case

The following requirements apply to the originating side.

[5.5.6-1] Subject to regulatory requirements and operator policy, the 3GPP network shall provide a means for third parties (e.g. enterprises) to be authorized to attest to the use of calling identity information by its authorized representatives.

[5.5.6-2] Subject to regulatory requirements and operator policy, the 3GPP network shall provide a means for authorized third parties to attest that an authenticated representative is authorized to include or reference the pre-established calling identity information included in the call setup or retrieved by the called party.

[5.5.6-3] Subject to regulatory requirements and operator policy, the 3GPP network shall provide a means to demonstrate the authenticity of the pre-established stored identity content that is referenced by the call setup and retrieved by the called party.

The following requirements apply to the terminating side.

[5.5.6-4] Subject to regulatory requirements and operator policy, the 3GPP network shall provide a means for third parties (e.g. enterprises) to be able to securely verify the caller’s authorization to use the caller identity information either in addition to any verification by the terminating PLMN or instead of verification performed by the terminating PLMN (PLMN skips the verification).

[5.5.6-5] Subject to regulatory requirements and operator policy, the 3GPP network shall provide a means to securely verify the authenticity of any stored identity content referenced by the call setup to be presented to the called party.

## 5.6 3rd party specific user identities

### 5.6.1 Description

The majority of businesses use MMTEL services not just for making voice calls, but also for other services, such as online meetings, AR/VR calls. There are several primary business functions that organizations use MMTEL services for, including internal communication, talking with prospects (sales call), contacting current customers and clients, customer support, and contact centre (or call centre) activities. While business customers consider the MMTEL services offer attractive features to their business, they also point out some practical issues:

#1 internal communication – MMTEL for internal communication can include voice calling, hosting online meetings, messaging co-workers, and team collaboration features like screen sharing. It could be challenging to manage the individual subscriptions for employees, especially considering employees may leave or join the companies or (for international companies) relocation to different countries.

#2 communication with current or potential customers – while MMTEL offers attractive pre-call, in-call and post-call features, the biggest issue for lots of businesses is that calls sometimes are rejected as fraudulent robocalls.

To target these pain points, this use case illustrates how businesses can be involved in the management and authentication of the user identities for its own employees when accessing MMTEL services.

### 5.6.2 Pre-conditions

Swift Delivery, a parcel delivery company, has a service contract with an operator iCA, including IMS subscription for the MMTEL services.

### 5.6.3 Service Flows

1. Rob Smith joined Swift Delivery recently, the company enabled him to use MMTEL services for business by updating the IMS business subscription with iCA with a user identity created for him

2. When Rob calls the customer, the customer can easily identify that it is a legitimate call from the Swift Delivery and not a fraudulent robocall.

The figure below illustrates an example where the operator authenticated caller identity can be displayed to the called party.

Figure 5.6.3-1 an example of CLIP for authenticated business callers

### 5.6.4 Post-conditions

With the company specific user identity, Rob can use the MMTEL services for internal communication with co-workers and external communications with customers.

### 5.6.5 Existing features partly or fully covering the use case functionality

The related existing features partially covering the required use case functionality can be found in TS 22.173 “IP Multimedia Core Network Subsystem (IMS) Multimedia Telephony Service and supplementary services” clause 4.6 “Multi-device and Multi-Identity”:

*The owner (e.g. individual subscriber, operator/company) of the added identity needs to give its permission for others to use the identity.*

*The owner (e.g. individual subscriber, operator/company) of the identity shall be able to revoke the permission for others to use the identity.*

With regards to Caller ID authentication, there have been on-going efforts across the globe in enhancing the protection against unwanted robocalls, among which STIR/SHAKEN are most well-known solutions to verify the caller, minimizing unwanted robocalls and protecting consumers from fraudulent intentions from bad actors. IETF-defined STIR (Secure Telephone Identity Revisited) is considered one of the most promising technologies to combat unwanted robocalls, while SHAKEN (Signature-based Handling of Asserted information using tokens) framework provides a mechanism to implement STIR to authenticate calls. The STIR/SHAKEN works for authenticating and thus proving that a call has not been spoofed, but it does not determine caller intent. This means that bad actors may continue making unwanted calls by registering telephone numbers, which, as long as registered, are authentically theirs.

### 5.6.6 Potential New Requirements needed to support the use case

[5.6.6-1] Subject to regulatory requirements and operator policy, the 3GPP network shall provide suitable and secure means to allow an authorized and trusted third-party (e.g. business customers) to update the MMTEL subscription and create a specific identity to an authorized user.

NOTE: The third party is authorized to change user identities for those subscriptions authorized by the operator.

[5.6.6-2] Subject to regulatory requirements and operator policy, the 3GPP network shall provide suitable and secure means to allow an authorized and trusted third-party (e.g. business customers) to authorize the use of identity information the third party is authorized to use by the third-party's users when that user is originating MMTEL services.

[5.6.6-3] Subject to regulatory requirements and operator policy, the 3GPP network shall provide suitable and secure means to allow an authorized and trusted third-party (e.g. business customers) to attest that the third party's user is authorized to use the identity information when that user is originating MMTEL services.

[5.6.6-4] Subject to regulatory requirements and operator policy, the 3GPP network shall provide suitable and secure means to allow an authorized and trusted third-party (e.g. business customers) to authenticate the identity information of its users at the terminating network.

## 5.7 Real-time visual interactive menu

### 5.7.1 Description

This use case is about the call centre that provides customers with self-service information, such as hotels, banks, and restaurants. Such information is provided via the voice platform with the voice navigation prompts to guide a customer. The customer can then select the desired services according to the voice prompts. The traditional method of displaying multiple options in sequence on the voice interface may appear confusing or even misleading. Customers often need to listen to the voice navigation very carefully in order to follow the instructions correctly and obtain the desired services. Unclear voice navigation and/or misunderstanding of voice navigation are among the main reasons for customer service processing delay and suboptimal customer relations.

To overcome this issue, the interactive menu can be made available to the customers corresponding to the voice navigation prompts. The customer can then interact with the call centre through the menu selection in a faster and error-free manner, improving the customer service satisfaction.

### 5.7.2 Pre-conditions

Isabelle is a customer of Universal Bank. Universal Bank provides online self-service via its customer service hot lines. Universal Bank has the subscription with the visual interactive menu services.

### 5.7.3 Service Flows

1. Isabelle calls the bank’s customer service hot line.
2. After the call is connected, the customer service centre answers with voice navigation prompt via the voice platform. At the same time, the customer service also sends the visual interactive menu of the service options via the data channel added via APIs provided by MMTEL service.
3. Isabelle listens to the announcement about the available banking services while viewing the service options on her smart phone.
4. She then selects the "debit card service" option on her phone following the instruction. The voice platform receives her selection and plays the next-level service options available for the debit card service. At the same time, Isabelle can also see the next-level service options on her phone.
5. She then selects the "service bundle query" option on the phone. The related information on the subscribed service bundle is then displayed on her phone.

### 5.7.4 Post-conditions

During a MMTEL call with the bank’s customer service hot line (out of 3GPP system), Isabelle carries out the self-service banking activities following the instruction provided by both the voice navigation prompts and the corresponding visual menu displayed on her smart phone.

### 5.7.5 Existing features partly or fully covering the use case functionality

The features of “flexible data channel handling” partially cover the required use case functionality. The details of data channel can be found in TS 26.114 “IP Multimedia Subsystem (IMS); Multimedia Telephony; Media handling and interaction”:

*Data channels do not require use of any codec but allows for real-time interaction in parallel to the conversational media (see clause 6.2.10). The User interface in Figure 4.2 interacts with a web page and a related script received through a downlink data channel to handle the data channel I/O and data formatting. All conversational media components are transported over RTP with each respective payload format mapped onto the RTP (RFC 3550 [9]) streams. The data channels are using SCTP (RFC 4960 [173]) over DTLS (IETF RFC 8261 [174]), used as specified for WebRTC data channels ([175]).*

**

*Figure 4.3: User plane protocol stack for a basic MTSI client*

Figure 5.7.5-1: copy of Figure 4.3 of TS 26.114 on User plane protocol stack for a basic MTSI client

In addition, the related existing service requirements about service exposure with 3rd party service providers can be found in TS 22.101 “Service principles” clause 29.2 “Exposed Services and capabilities”:

*The 3GPP Core Network shall be able provide a standardized interface to enable exposure of the following services and capabilities to 3rd party service providers:*

*Support of 3rd party requested session QoS and priority*

*- The 3GPP Core Network shall enable a 3rd party service provider to request setting up data sessions with specified QoS (e.g. low latency or jitter) and priority handling to a UE that is served by the 3rd party service provider.*

The IMS Multimedia Telephony service supports providing suitable APIs for a trusted third-party (e.g. enterprise call center) to control media components (e.g. establish, modify or terminate voice, video ,data channels ) during an ongoing MMTEL session.

Note: third-party application may be out of scope.

### 5.7.6 Potential New Requirements needed to support the use case

Void.

# 6 Consolidated potential requirements

## 6.1 Media handling capability

The IMS multimedia telephony service shall support AR media processing.

## 6.2 Authorization and authentication

Requirements in this section are subject to regulatory requirements and operator policy.

[6.2-1] The 5G system shall provide means to allow a trusted third-party to update the MMTEL subscription and allocate a third-party specific identity to an authorized user.

NOTE 1: The third party is authorized to change user identities for those subscriptions authorized by the operator.

The following requirements apply to the originating side.

[6.2-2] The 5G network shall provide a means for third parties (e.g. enterprises) to be authorized to verify the use of calling identity information by its authorized users.

[6.2-3] The 5G network shall provide a means for authorized third parties to verify that an authenticated user is authorized to include or reference the pre-established calling identity information included in the call setup or retrieved by the called party.

[6.2-4] The 5G network shall provide a means to verify the authenticity of the pre-established stored identity information that is referenced by the call setup and retrieved by the called party.

The following requirements apply to the terminating side.

[6.2-5] The 5G network shall provide a means for third parties (e.g. enterprises) to be able to verify the caller’s authorization to use the identity information either in addition to or instead of verification performed by the terminating PLMN.

[6.2-6] The 5G network shall provide a means to verify the authenticity of any stored identity information referenced by the call setup to be presented to the called party.

Annex A:
Change history

|  |
| --- |
| **Change history** |
| **Date** | **Meeting** | **TDoc** | **CR** | **Rev** | **Cat** | **Subject/Comment** | **New version** |
| 2019-11 | SA1#88 | S1-193221 |  |  |  | First revision | 0.2.0 |
| 2020-09 | SA1#90e | S1-203380 | - | - | - | Including agreed docs: [S1-203277](file:///D%3A%5C3GPP_SA1%5CTSGS1_91e_ElectronicMeeting%5CAgendas%5Cdocs%5CS1-203277.zip), [S1-203308](file:///D%3A%5C3GPP_SA1%5CTSGS1_91e_ElectronicMeeting%5CAgendas%5Cdocs%5CS1-203308.zip) | 0.3.0 |
| 2021-01 | SA1#93e | S1-210011 | - | - | - | Identical to v.0.4.0 which should have been provided in S1-204350 but problem with upload. | 0.4.1 |
| 2021-03 | SA1#93e | [S1-210300](file:///C%3A%5CUsers%5Calmodovarchicojl%5CDesktop%5CTSGS1_93e_Electronic_Meeting%5Cdocs%5CS1-210300.zip) |  |  |  | Including agreed docs: [S1-210106](file:///C%3A%5CUsers%5Calmodovarchicojl%5CDesktop%5CTSGS1_93e_Electronic_Meeting%5CDocs%5CS1-210106.zip), [S1-210407](file:///C%3A%5CUsers%5Csultan%5CAppData%5CLocal%5CMicrosoft%5CWindows%5CINetCache%5CContent.Outlook%5C4XT0V618%5Cdocs%5CS1-210407.zip), [S1-210408](file:///C%3A%5CUsers%5Csultan%5CAppData%5CLocal%5CMicrosoft%5CWindows%5CINetCache%5CContent.Outlook%5C4XT0V618%5Cdocs%5CS1-210408.zip), [S1-210370](file:///C%3A%5CUsers%5Csultan%5CAppData%5CLocal%5CMicrosoft%5CWindows%5CINetCache%5CContent.Outlook%5C4XT0V618%5Cdocs%5CS1-210370.zip), [S1-210371](file:///C%3A%5CUsers%5Csultan%5CAppData%5CLocal%5CMicrosoft%5CWindows%5CINetCache%5CContent.Outlook%5C4XT0V618%5Cdocs%5CS1-210371.zip), [S1-210372](file:///C%3A%5CUsers%5Csultan%5CAppData%5CLocal%5CMicrosoft%5CWindows%5CINetCache%5CContent.Outlook%5C4XT0V618%5Cdocs%5CS1-210372.zip), [S1-210373](file:///C%3A%5CUsers%5Csultan%5CAppData%5CLocal%5CMicrosoft%5CWindows%5CINetCache%5CContent.Outlook%5C4XT0V618%5Cdocs%5CS1-210373.zip) | 0.5.0 |
| 2021-05 | SA1#94e | S1-211300 | - | - | - | Output of SA1#94e | 0.6.0 |
| 2021-06 | SA#92e | SP-210505 | - | - | - | Raised by MCC to v.1.0.0 for one-step approval | 1.0.0 |
| 2021-06 | SA#92e | SP-210505 | - | - | - | Raised by MCC to v.18.0.0 following one-step approval | 18.0.0 |