



- (51) **International Patent Classification:**
H04M 3/56 (2006.01) *H04M 3/38* (2006.01)
- (21) **International Application Number:**
PCT/GB2012/051433
- (22) **International Filing Date:**
21 June 2012 (21.06.2012)
- (25) **Filing Language:** English
- (26) **Publication Language:** English
- (30) **Priority Data:**
1110494.0 21 June 2011 (21.06.2011) GB
- (71) **Applicant (for all designated States except US):**
METASWITCH NETWORKS LTD; 100 Church Street,
Enfield, Middlesex EN2 6BQ (GB).
- (72) **Inventor; and**
- (75) **Inventor/Applicant (for US only):** **COLLINS, Neil James**
[GB/GB]; c/o Metaswitch Networks Limited, 100 Church
Street, Enfield, Middlesex EN2 6BQ (GB).
- (74) **Agent:** **FORSYTH, Neil**; EIP, Fairfax House, 15 Fulwood
Place, London, Greater London WC1V 6HU (GB).

(81) **Designated States** (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) **Designated States** (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— without international search report and to be republished upon receipt of that report (Rule 48.2(g))

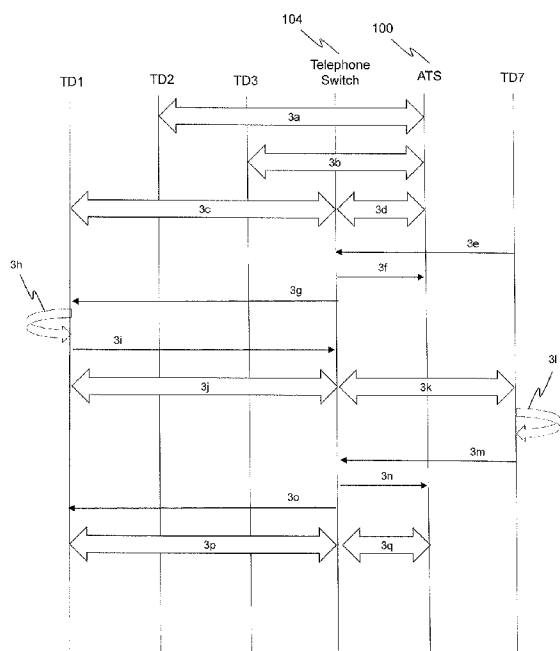
(54) **Title:** MULTI-PARTY TELECONFERENCE METHODS AND SYSTEMS

FIG. 3

(57) **Abstract:** A first call leg for transmittal of audio data associated with a multi-party teleconference call is established between a first telephony device and a multi-party teleconference service. The multi-party teleconference service is capable of connecting the first telephony device with at least two other telephony devices during the multi-party teleconference call. State information indicating that the multi-party teleconference call is of a second call type is stored. During the multi-party teleconference call, an incoming telephone call setup request for the first telephony device is received which is associated with an incoming telephone call of a first call type from a further telephony device which is not a participant in the multi-party teleconference call. The incoming telephone call setup request and the multi-party teleconference call are handled in accordance with the stored state information indicating that the multi-party teleconference call is of the second call type.

Multi-party teleconference methods and systems

Technical Field

5 The present disclosure relates to multi-party teleconference methods and systems.

Background

10 Multi-party audio teleconferences today are a vital business tool for allowing people to “meet” without being in the same location. A multi-party audio teleconference typically comprises a number (greater than two) of endpoint telephony devices, either wire line or wireless, and an audio teleconference bridge. The endpoint telephony devices dial into the multi-party audio teleconference bridge, enter an access code and can then talk to other people in the multi-party audio teleconference call. The way this tends to work
15 is that the multi-party audio teleconference bridge detects several of the loudest (or “active”) speakers, mixes the audio from these speakers and sends out the mixed audio to all participants but subtracts the audio that each individual is sending to the audio teleconference bridge to avoid echo or feedback.

20 With the wide deployment of fast, reliable Internet services, many people are now choosing to work at home. It makes sense for them in terms of work/life balance and can also boost productivity. One major downside of working at home, however, is that people can become quite detached from the social aspects of working in an office. In an office, if someone passes another person’s desk and stops for a chat it is quite natural, but calling that person on
25 the telephone for the same type of chat is not. Likewise, in an office, it is normal to overhear the conversations of others and join in if desired, but such an opportunity does not arise for people working from home. It is in these kinds of environments, and others, in which improved multi-party teleconference methods could be employed to great effect.

30 It would therefore be desirable to provide improved multi-party teleconference methods and systems.

Summary

In accordance with first embodiments, there is a method of handling an incoming telephone call of a first call type during a multi-party teleconference call which is of a second call type, different from the first call type, the method comprising:

5 establishing a first call leg for transmittal of audio data associated with a multi-party teleconference call between a first telephony device and a multi-party teleconference service, the multi-party teleconference service being
10 capable of connecting the first telephony device with at least two other telephony devices during the multi-party teleconference call;

storing state information indicating that the multi-party teleconference call is of the second call type;

during the multi-party teleconference call, receiving an incoming
15 telephone call setup request for the first telephony device, the incoming telephone call setup request being associated with an incoming telephone call from a further telephony device which is not a participant in the multi-party teleconference call; and

handling the incoming telephone call setup request and the multi-party
20 teleconference call in accordance with the stored state information indicating that the multi-party teleconference call is of the second call type.

Hence, embodiments provide for handling a call, such as a particular, new type of multi-party teleconference call, which is interruptible, differently to a standard call type, for example. The multi-party teleconference call is
25 interruptible such that when a given participant in the multi-party teleconference call receives an incoming telephone call originating from outside of the multi-party teleconference call, the participant is removed from the multi-party teleconference call and the incoming telephone call is directed to the participant's telephony device. Embodiments therefore avoid the telephony
30 devices of multi-party teleconference call participants from being engaged and hence unavailable for incoming calls.

In some embodiments, the handling comprises disabling the first call leg between the first telephony device and the multi-party teleconference service. Hence, when a multi-party teleconference call is interrupted for a given user, audio data from the user's telephony device is not transmitted into the multi-party teleconference call and no audio data from the multi-party teleconference is transmitted to the user's telephony device.

In some embodiments, the disabling comprises tearing down the first call leg between the first telephony device and the multi-party teleconference service. Hence, network resources in and out of multi-party teleconference service can be released and used for other purposes.

In other embodiments, the disabling comprises maintaining the network resources reserved for establishing the first call leg, but not transmitting any audio data from the first telephony device to the multi-party teleconference service. Hence, the time taken in reserving network resources when re-entering the multi-party teleconference call is reduced.

In some embodiments, the handling comprises directing the incoming telephone call setup request to the first telephony device, whereby the first telephony device enters a ringing state in relation to the incoming telephone call. Hence, when a multi-party teleconference call is interrupted in favour of an incoming call for the user, the user is notified of the incoming call by a ringing tone emitting from their telephony device. The user need take no action to exit the call multi-party teleconference call, i.e. no user input is required for the incoming call to be directed to their telephony device.

Embodiments comprise, in response to a user of the first telephony device answering the incoming call, establishing a second call leg for transmittal of audio data associated with the answered telephone call between the first telephony device and the further telephony device. Hence, the user can choose to take the incoming call as normal.

Other embodiments comprise, in response to termination of the answered telephone call, tearing down the second call leg between the first telephony device and the further telephony device; and establishing a third call leg for

transmittal of audio data associated with the multi-party teleconference call between the first telephony device and the multi-party teleconference service. Hence, once the incoming call has finished, the user can be reconnected back into the multi-party teleconference call.

5 In some embodiments, the establishing of the third call leg comprises applying modified signalling to the third call leg such that the third call leg will be automatically answered by the first telephony device. Hence, no user input, other than hanging up the incoming call is required from the user in order to re-enter them into the multi-party teleconference call.

10 In embodiments, the establishing of the third call leg comprises instructing the first telephony device to operate in speakerphone mode. Hence, when a user enters the multi-party teleconference call, they will automatically hear any audio produced by all other multi-party teleconference call participants without having to pick up their telephony device handset. Also, any audio
15 produced by a user will be transmitted to all other multi-party teleconference call participants without the user having to speak directly into the handset of their telephony device.

 In some embodiments, the first telephony device is capable of operating either in handset mode or speakerphone mode, during the multi-party
20 teleconference call, the first telephony device operates in speakerphone mode, and the handling comprises instructing the first telephony device to operate in handset mode. Hence, when a multi-party teleconference call is interrupted for a user, they will automatically hear their telephony device ringing without any user input being required on the part of the user.

25 Embodiments comprise receiving an incoming multi-party teleconference call setup request from the first telephony device, the incoming multi-party teleconference call setup request comprising an identifier associated with the multi-party teleconference call, and the multi-party teleconference call in respect of which the first call leg is established is identified on the basis of the
30 received identifier. Hence, the user will be entered into the correct multi-party

teleconference call, for example by recognition of a telephone dialling number associated with the first telephony device.

In some embodiments, establishing the third call leg comprises utilising the maintained network resources. Hence, re-entry back into the multi-party
5 teleconference call can be expedited.

Some embodiments comprise, in response to a user of the first telephony device not answering the first telephony device in the ringing state for a predetermined time period, aborting the handling and re-enabling the first call leg between the first telephony device and the multi-party teleconference
10 service. Hence, if a user chooses not to or is unable to take the incoming call, they will re-enter the multi-party teleconference call.

In accordance with second embodiments, there is a method of controlling multi-party teleconference calls in a telecommunications network, said method comprising:

15 establishing a multi-party teleconference call between at least three telephony devices in said network;

receiving an audio stream from each of said at least three telephony devices in said network;

processing said received audio data streams such that the received audio
20 data streams are combined into mixed audio data streams, the received audio data streams having an incoming volume and one or more mixing volume adjustment level settings which identify a volume adjustment level of the received audio data stream, relative to one or more other received audio data streams, in said mixed audio data streams; and

25 outputting said mixed audio data streams,
wherein said method comprises:

during at least a part of said multi-party teleconference call, transmitting to a first telephony device, of said at least three telephony devices, a first mixed audio data stream generated using a first audio mixing mode, in which said first
30 mixed audio data stream is generated from a plurality of received audio data streams, including a first received audio data stream and a second received audio

data stream each having, in said first audio mixing mode, a respective volume adjustment level setting which results in a non-zero contribution to the first mixed audio data stream, the volume adjustment level settings of said first and second received audio data streams having first respective levels in said first
5 audio mixing mode; and

in response to detecting a trigger during said multi-party teleconference call, transmitting to said first telephony device a second mixed audio data stream generated using a second audio mixing mode, in which said second mixed audio data stream is generated from said plurality of received audio data streams,
10 including said first received audio data stream and said second received audio data stream, said first and second received audio data streams each having, in said second audio mixing mode, a respective volume adjustment level setting which results in a non-zero contribution to the second mixed audio data stream, the volume adjustment level settings of said first and second received audio data
15 streams having second respective levels in said second audio mixing mode, at least one of which is different to one of said first respective levels.

Hence, embodiments provide for volume adjustment of the received data streams in order to provide improved functionality upon the occurrence of certain trigger events.

20 In various embodiments, the first respective volume adjustment levels comprise substantially equal volume adjustment levels, whereby each audio data stream in the plurality has a substantially equal volume adjustment level in the first mixed audio data stream. Hence, audio data due to each multi-party teleconference call participant can be heard substantially equally by all other
25 multi-party teleconference call participants. A given call participant, however, typically does not hear their own speech via the multi-party teleconference, to avoid echo effects.

In various embodiments, the second respective levels comprise a relatively high volume adjustment level for the first received audio data stream
30 and a relatively low volume adjustment level for the second received audio data stream, in the second audio mixing mode. The second respective levels may

comprise a plurality of substantially equal volume adjustment levels, including the adjustment level for the second received audio data stream.

This allows one, or more, participants to have their contribution highlighted amongst the other contributions. For example, if one multi-party
5 teleconference call participant raises their voice, then this may be reflected in their voice being not only relatively louder in the mixed audio data stream transmitted to other participants, but also receiving a relatively high volume adjustment level. Such emphasis of the contribution from one participant simulates, and emphasizes, the situation where one person in a room raises their
10 voice, for example to make an important announcement, in a multi-party teleconference environment.

In some embodiments, the second respective levels comprise a plurality of substantially equal volume adjustment levels, including the adjustment level for said first received audio data stream. Hence, when two or more participants
15 are identified as conversing with each other rather than to other participants in general, their relative contribution to the transmitted mixed audio data streams can be boosted to emphasize that conversation.

In embodiments, substantially equal volume adjustment levels of the first respective volume adjustment levels may receive a higher volume adjustment
20 level than de-emphasized, substantially equal volume adjustment levels, of the second respective volume adjustment levels. Hence, the volume adjustment level contribution to the mixed audio data streams due to for example background chatter can be lowered when a user raises their voice or two users are conducting a direct conversation via the multi-party teleconference.

Embodiments comprise transmitting the second mixed audio data stream
25 generated using the second audio mixing mode to the first telephony device for a predetermined time period after detection of the trigger. Hence, a user's voice may be boosted, for example, over the duration of an announcement.

In some embodiments, the trigger comprises a mean input volume of the
30 first received audio data stream rising above a predetermined mean input volume.

Embodiments involve storing data representative of one or more predetermined words, and monitoring the received audio data streams for the utterance of any of the one or more predetermined words represented in the stored data, and the trigger comprises the monitoring detecting utterance of a given one or more predetermined words represented in the stored data. Hence, 5 embodiments can dynamically adjust the processing of audio data in the multi-party teleconference call to implement predetermined functionality. The predetermined functionality can be triggered by one of the participants uttering one or more specific key words or phrases.

10 In some embodiments, the given one or more predetermined words are uttered by a first user associated with a second telephony device of the at least three telephony devices, the given one or more predetermined words comprise an identifier for a second user associated with a third telephony device of the at least three telephony devices.

15 In some embodiments, the given one or more predetermined words are uttered by a first user associated with a second telephony device of the at least three telephony devices, the given one or more words comprise an identifier for a second user associated with a third telephony device of the at least three telephony devices, the given one or more words further comprise an indication 20 that the first user wishes to conduct a telephone call with the second user separate to the multi-party teleconference call, and the second mixed audio data stream is generated in the second, different respective volume adjustment levels from audio data streams in the plurality apart from audio data stream received from the second and third telephony devices. Hence, two participants can have 25 an exclusive conversation with each other without the conversation being heard by other participants of the multi-party teleconference. These, other participants however, receive different respective volume adjustment levels when the two private call participants leave the teleconference, even if temporarily.

Embodiments comprise, in response to detecting the trigger during the 30 multi-party teleconference call, transmitting the audio data stream received from the second telephony device to the third telephony device, transmitting the audio

data stream received from the third telephony device to the second telephony device, and not transmitting the audio data streams received from the second and third telephony devices to the other participants.

Hence, two participants may have an exclusive conversation with each other, whilst the multi-party teleconference call may continue for the remaining other multi-party teleconference call participants.

In accordance with third embodiments, there is a method of controlling multi-party teleconference calls in a telecommunications network, said method comprising:

10 storing data representative of one or more predetermined words or phrases;

establishing a multi-party teleconference call between at least three telephony devices in said network;

15 receiving an audio stream from each of said at least three telephony devices in said network;

processing said received audio data streams such that the received audio data streams are combined into mixed audio data streams using one of a plurality of available mixing modes;

20 monitoring said received audio data streams for the utterance of at least one of said one or more predetermined words or phrases represented in said stored data; and

altering a mixing mode applied during the processing of said received audio data streams in response to said monitoring detecting the utterance of said one or more predetermined words represented in said stored data.

25 Hence, a speech recognition function can be used to allow a participant to alter the mixing mode applied, for example by the utterance of a word or phrase.

In accordance with fourth embodiments, there is apparatus adapted to perform the method of the first, second or third embodiments.

30 In accordance with fifth embodiments, there is computer software adapted to perform the method of the first, second or third embodiments.

In accordance with sixth embodiments, there is provided a computer program product comprising a non-transitory computer-readable storage medium having computer readable instructions stored thereon, the computer readable instructions being executable by a computerized device to cause the computerized device to perform the method of the first, second or third
5 embodiments.

Further features and advantages will become apparent from the following description of preferred embodiments, given by way of example only, which is made with reference to the accompanying drawings.

10

Brief Description of the Drawings

Figure 1 shows a system diagram according to embodiments.

Figure 2 shows a block diagram according to embodiments.

Figure 3 shows a flow diagram according to embodiments.

15 Figure 4 shows a graph according to embodiments.

Figure 5 shows a graph according to embodiments.

Figure 6 shows a graph according to embodiments.

Detailed Description

20 Figure 1 shows a system diagram according to some embodiments. Figure 1 depicts a telecommunications network 110 in which audio teleconferencing server (ATS) 100 is responsible for providing multi-party audio teleconference services to telephony devices TD1, TD2, TD3, TD4, TD5 and TD6. ATS 100 may also be responsible for providing audio teleconference
25 services to other telephony devices (not shown).

Each of TD1, TD2, TD3, TD4, TD5 and TD6 is associated with a user who works for or is otherwise associated with a given group or organisation. Some of telephony devices TD1, TD2, TD3, TD4, TD5 and TD6 may be located in one office of the organisation, some may be located in one or more other
30 offices of the organisation and some may be located at home locations where the respective users reside. A user may be associated with more than one telephony

device, for example a user may have one telephony device located on their desk at work and another telephony device located at their home.

Telephony device TD1 is provided with telephony services by telephone switch 104 which is connected to ATS 100 via network 102. Network 102 may
5 comprise one or more Public Switch Telephone Networks (PSTNs) and/or the Internet.

A user of telephony device TD1 also has access to a computing device 114 with display and user input capabilities, for example a personal computer, laptop, or personal digital assistant, etc. which is located locally to telephony
10 device TD1.

Telephone switch 104 is responsible for providing switching of telephone calls for a number of telephony devices such as telephony device TD1 including provision of dial tone, ringing tone, etc. Telephone switch 104 may also include the ability to select processes that can be applied to such calls,
15 routeing for such calls based on signalling and subscriber database information, the ability to transfer control of calls to other network elements and management functions such as provisioning, fault detection and billing. Telephone switch 104 may also be referred to as a local telephone exchange, central office, class 5 switch or softswitch. Telephone switch 104 includes a database 108 for storing
20 call related data, including call state information, for example in relation to telephone calls incoming to or outgoing from telephony device TD1.

Telephony devices TD2, TD3, TD4, TD5 and TD6 may similarly be provided with telephony services by one or more other telephone switches (not shown). In this example, telephony device TD7 is provided with telephony
25 services by the same telephone switch 104 as telephony device TD1, but such services could be provided by a different telephone switch.

Figure 2 shows a block diagram according to embodiments. Figure 2 shows some components of ATS 100 of Figure 1. ATS 100 comprises a processor 200 (or processors) for carrying out data processing functionality.
30 ATS 100 includes a data store 206 for storing data representative of predetermined words or phrases, and a data store 208 for storing user

configuration data. ATS 100 has an audio monitoring module 204 with speech recognition capabilities for monitoring audio data streams received from telephony devices TD1 to TD6. ATS 100 has a web interface 210 for providing user access to various multi-party audio teleconference functionalities such as display and user configuration. ATS 100 has a mixing module 202 for processing audio data streams received from telephony devices TD1 to TD6 including combining the received audio data streams into one or more mixed audio data streams according to different audio mixing modes. Mixing module 202 includes codec functionality for decoding audio data streams received from telephony devices in different data formats and recoding mixed audio data streams for transmittal out to the telephony devices in the appropriate data formats.

Some embodiments relate to handling an incoming telephone call of a first call type during a multi-party audio teleconference call which is of a second call type, different from the first call type. The second call type may for example be a call that can be interrupted. The first call type may for example be a call that can not be interrupted.

ATS 100 provides a multi-party audio teleconference service and a multi-party audio teleconference call is currently being conducted between TD2, TD3, TD4, TD5 and TD6.

The user of TD1 wishes to join the multi-party audio teleconference so dials an appropriate number associated with ATS 100 which results in an incoming multi-party audio teleconference call setup request being received at telephone switch 104. Telephone switch 104 establishes a first call leg for transmittal of audio data associated with the multi-party audio teleconference call between TD1 and ATS 100. ATS 100 connects TD1 to the audio teleconference call being conducted between TD2 to TD6. Telephone switch 104 stores, in state information store 108, state information indicating that the call between TD1 and TD2 to TD6 is of the second call type. In this embodiment, the second call type is an interruptible multi-party audio teleconference call type. This indicates that it can be interrupted by an incoming

call to TD1. Alternatively, it may be indicative of a more general call type, for example an interruptible call type.

During the multi-party audio teleconference call between TD1 and TD2 to TD6, TD7 dials a telephone dialling number associated with TD1 which results in an incoming call setup request from TD7 being received at telephone switch 104. TD7 is not a participant in the multi-party audio teleconference call currently being conducted between TD1 and TD2 to TD6. Telephone switch 104 inspects the stored state information for the multi-party audio teleconference call and realises that the multi-party audio teleconference is of the second call type so can be interrupted with respect to TD1 by the incoming call from TD7. Telephone switch 104 therefore proceeds to handle the incoming telephone call setup request and the multi-party audio teleconference call in accordance with the stored state information indicating that the multi-party audio teleconference call is of the second call type.

Telephone switch 104 directs the incoming telephone call setup request to TD1 which results in TD1 ringing, i.e. TD1 enters a ringing state in relation to the incoming telephone call from TD7.

Telephone switch 104 disables the first call leg between TD1 and ATS 100. The audio teleconference call between TD2 to TD6 carries on as normal, but without TD1 participating.

If the user of TD1 chooses to answer the incoming call, telephone switch 104 establishes a second call leg for transmittal of audio data associated with the answered telephone call between TD1 and TD7. The users of TD1 and TD7 are thus able to conduct a telephone call to each other. During this telephone call, ATS 100 continues to provide a multi-party audio teleconference service to TD2 to TD 6 and the audio teleconference call between TD2 to TD6 continues as normal.

When the user of TD1 or the user of TD7 terminates the telephone call between them, for example by hanging up their respective telephony device, telephone switch 104 receives signalling information indicating such and tears down the second call leg between TD1 and TD7. Telephone switch 104

establishes a third call leg for transmittal of audio data associated with the multi-party audio teleconference call between TD1 and ATS 100. TD1 thus re-enters the multi-party audio teleconference call being conducted between TD2 to TD6, i.e. the audio teleconference call then has all of TD1 to TD7 participating again.

5 In some embodiments, TD1 is capable of operating either in a handset mode or a speakerphone mode. Speakerphone mode is a mode of operation where a loudspeaker on TD1 outputs audio from TD1 without the user having to pick up the handset and put it to their ear/mouth. Similarly, in speakerphone mode, an external (i.e. external to the handset) microphone on TD1 picks up
10 audio generated by the user (possibly also audio generated by others in close proximity and other background noise), without the user having to pick up the handset and put it to their ear/mouth.

 When the user of TD1 initially joins the multi-party audio teleconference, the user dials an appropriate number associated with ATS 100
15 without picking up the handset of TD1. TD1 thus enters the multi-party audio teleconference in speakerphone mode, i.e. during the multi-party audio teleconference call, TD1 operates in speakerphone mode. The appropriate number associated with ATS 100 may for example be stored in a speed-dial function on TD1 such that the user of TD1 just needs to press a single button to
20 join the multi-party audio teleconference initially.

 When telephone switch 104 handles the incoming call setup request received from TD7, telephone switch 104 instructs TD1 to operate in handset mode. This means that TD1 will cease to pick up audio from its external microphone. Further, no audio data from the disabled first call leg with ATS
25 100 will be output by the loudspeaker of TD1. However, when telephone switch 104 directs the incoming telephone call setup request to TD1 and it enters a ringing state, TD1 is in handset mode and will emit a ringing sound from its loudspeaker.

 When the user of TD1 picks up the handset of TD1 to conduct the
30 telephone call with the user of TD7, TD1 will be in handset mode and continue to be in handset mode when the call with TD7 is terminated and the user of TD1

puts the handset of TD1 back on-hook. However, when TD1 re-enters the multi-party audio teleconference call with TD2 to TD6, TD1 should be in speakerphone mode. Therefore, when telephone switch 104 establishes the third call leg between TD1 and ATS 100, telephone switch 104 instructs TD1 to
5 operate in speakerphone mode.

Further, to avoid the user of TD1 having to perform any further operations to re-enter the multi-party audio teleconference, other than putting the handset of TD1 back on-hook, telephone switch 104 applies modified signalling to the third call leg such that the third call leg will be automatically
10 answered by TD1.

When the user of TD1 initially dials the number associated with ATS 100, this results in an incoming multi-party audio teleconference call setup request call being received at telephone switch 104. In embodiments, the incoming multi-party audio teleconference call setup request comprises an
15 identifier associated with the multi-party audio teleconference call being conducted between TD2 and TD7 which allows telephone switch 104 to recognise which multi-party audio teleconference call the incoming multi-party audio teleconference call setup request call relates to, i.e. one provided by ATS 100. Telephone switch 104 therefore forwards the incoming multi-party audio
20 teleconference call setup request on to ATS 100 and establishes the first call between TD1 and ATS 100 on the basis of the received identifier.

In some embodiments, the received identifier comprises a telephone dialling number associated with TD1, for example detected from a calling line identifier (CLI) field of the incoming multi-party audio teleconference call setup
25 request.

In another embodiment, the received identifier comprises a telephone dialling number reserved for multi-party audio teleconference calls conducted via ATS 100, with telephone switch 104 being preconfigured to recognise multi-party audio teleconference call setup requests to the reserved dialling number
30 and forward them to ATS 100 accordingly.

In some embodiments, telephone switch 104 disables the first call leg between TD1 and ATS 100 by tearing down the first call leg between TD1 and ATS 100.

5 In other embodiments, telephone switch 104 disables the first call leg between TD1 and ATS 100 by maintaining the network resources reserved for establishing the first call leg, but not transmitting any audio data from TD1 to ATS 100 for the duration of the incoming call from TD7. In such embodiments, after the incoming call from TD7 is terminated, telephone switch 104 establishes the third call leg by utilising the maintained network resources. This avoids
10 having to establish the third call leg from scratch and can help speed up re-entry back into the multi-party audio teleconference call.

In some embodiments, if the user of TD1 chooses not to take the incoming call from TD7, or if the user of TD1 is unable to take the incoming call from TD7 such that TD1 remains in a ringing state for a predetermined time
15 period, telephone switch 104 will abort the handling of the incoming call and re-enable the first call leg between TD1 and ATS 100.

Figure 3 shows a flow diagram according to some embodiments. Figure 3 depicts the flow of signalling messages and audio data between various entities of Figure 1 when a multi-party audio teleconference call is interrupted
20 by an incoming call from a telephony device which is not participating in the multi-party audio teleconference call.

A multi-party audio teleconference call is currently being conducted between TD1, TD2, and TD3. Other telephone devices (not shown) may also participate in the multi-party audio teleconference call.

25 A call leg for transmittal of audio data associated with the multi-party audio teleconference call has been established between TD2 and ATS 100 as shown by step 3a. This call leg may be established via a telephone switch (not shown) which provides telephony services to TD2.

A call leg for transmittal of audio data associated with the multi-party
30 audio teleconference call has also been established between TD3 and ATS 100

as shown by step 3b. This call leg may be established via a telephone switch (not shown) which provides telephony services to TD3.

Telephone switch 104 establishes a call leg for transmittal of audio data associated with the multi-party audio teleconference call between TD1 and ATS 100 as shown by steps 3c and 3d. Telephone switch 104 stores state information indicating that the multi-party audio teleconference call between TD1 to TD3 is of the second call type and is thus a multi-party audio teleconference which can be interrupted by an incoming call to TD1.

During the multi-party audio teleconference call between TD1, TD2 and TD3, the user of TD7 dials a telephone dialling number associated with TD1 which results in an incoming call setup request from TD7 being received at telephone switch 104 as per step 3e. TD7 is not a participant in the multi-party audio teleconference call currently being conducted between TD1 to TD3. Telephone switch 104 inspects stored state information for the multi-party audio teleconference call and recognises that the multi-party audio teleconference is of the second call type and so can be interrupted with respect to TD1 by the incoming call from TD7. Telephone switch 104 therefore proceeds to handle the incoming telephone call setup request and the multi-party audio teleconference call in accordance with the stored state information indicating that the multi-party audio teleconference call is of the second call type.

Telephone switch 104 directs the incoming telephone call setup request to TD1 in step 3g which results in TD1 ringing, i.e. TD1 enters a ringing state in relation to the incoming telephone call from TD7. Step 3g also involves telephone switch 104 instructing TD1 to operate in handset mode.

In step 3f, telephone switch 104 disables the first call leg between TD1 and ATS 100. The audio teleconference call between TD2 and TD3 carries on as normal, but without TD1 participating.

In step 3h, the user of TD1 chooses to answer the incoming call from TD7, such action being notified to telephone switch 104 in step 3i. Telephone switch 104 establishes a second call leg for transmittal of audio data associated with the answered telephone call between TD1 and TD7 in steps 3j and 3k. The

users of TD1 and TD7 are thus able to conduct a telephone call to each other. During this telephone call, ATS 100 continues to provide a multi-party audio teleconference service to TD2 and TD3 and the audio teleconference call between TD2 and TD3 carries on as normal.

5 In step 3l, the user of TD7 terminates the telephone call between TD1 and TD7, such action being notified to telephone switch 104 in step 3m. Telephone switch 104 tears down the second call leg between TD1 and TD7.

Telephone 104 informs ATS 100 that TD1 is re-entering the multi-party audio teleconference call being conducted between TD2 and TD3 in step 3n.

10 In step 3o, telephone switch 104 instructs TD1 to operate in speakerphone mode. Telephone switch 104 establishes a third call leg for transmittal of audio data associated with the multi-party audio teleconference call between TD1 and ATS 100 in steps 3p and 3q. During establishment of the third call leg, telephone switch 104 applies modified signalling to the third call
15 leg such that the third call leg will be automatically answered by TD1, for example in conjunction with step 3o.

TD1 thus re-enters the multi-party audio teleconference call being conducted between TD2 and TD3, i.e. the audio teleconference call then has all three of TD1, TD2 and TD3 participating again.

20 Some embodiments relate to controlling a multi-party audio teleconference call in a telecommunications network. Such control is carried out by ATS 100 in relation to a multi-party audio teleconference call being conducted between at least three telephony devices in the network, for example telephony devices TD1, TD2, TD3, TD4, TD5 and TD6.

25 Various mixing modes are illustrated in Figures 4 to 6. During the multi-party audio teleconference, ATS 100 receives an audio data stream from each of telephony devices TD1, TD2, TD3, TD4, TD5 and TD6. Processor 200 of ATS 100 processes the received audio data streams and mixing module 202 of ATS 100 combines the received audio data streams into mixed audio data
30 streams. Different mixed audio streams are then transmitted to the telephony

devices TD1, TD2, TD3, TD4, TD5 and TD6 participating in the multi-party audio teleconference.

To avoid feedback or echoing, a mixed audio data stream transmitted to a given telephony device during the multi-party audio teleconference will not contain audio data received from that telephony device. For example, a mixed audio stream transmitted to TD6 will contain a mix of the audio data streams received from TD1 to TD5, but no audio data from the audio data stream received from TD6.

In each of the audio mixing modes shown in Figures 4 to 6, a mixed audio data stream is generated from a plurality of received audio data streams using an audio mixing mode. An audio mixing mode will generate a mixed audio data stream by combining the received audio data streams in the plurality according to respective volume adjustment levels defined for that audio mixing mode.

In accordance with embodiments, the mixing mode is normally set to a default audio mixing mode, in which each party provides a substantially equal contribution, and the volume adjustment levels for each incoming audio data stream included in any particular mixed audio data stream are equal. On receipt of an appropriate trigger, the ATS 100 switches to a non-default audio mixing mode, in which the volume adjustment levels for the respective streams are altered. On receipt of a further appropriate trigger, or timeout, the ATS 100 switches back to the default audio mixing mode.

During at least a part of the multi-party audio teleconference call, a first mixed audio data stream is generated using a first audio mixing mode, for example the default audio mixing mode, and is transmitted to TD6. The first mixed audio data stream is generated from a plurality of audio data streams in first respective volume adjustment levels, the plurality of data streams here being at least some of the audio data streams received from TD1 to TD5. During at least a different part of the multi-party audio teleconference call, a second mixed audio data stream is generated using a second audio mixing mode and is transmitted to TD6. The second mixed audio data stream is generated

from a plurality of audio data streams in second respective volume adjustment levels, the plurality of data streams here being at least some of the audio data streams received from TD1 to TD5.

Preferably, in each of the audio mixing modes shown in Figures 4 to 6, the ATS 100 determines mean input volumes for each data stream. The mean input volumes may be determined according to various known techniques, for example the mean input volume may be determined as an RMS (Root Mean Square) value of the incoming signal, to represent the power of the incoming signal. The mean input volume of each of the received audio data streams is preferably measured over a periodic or sliding window of between 0.5 to 5 seconds, preferably between 1 and 2 seconds.

Preferably, in each of the audio mixing modes shown in Figures 4 to 6, the ATS 100 determines volume adjustment levels for each data stream. These volume adjustments may be applied as a pre-amplification levels, applied in a pre-amplifier, before mixing of each of the audio data streams in equal proportions by an audio mixer, or alternatively may be applied as mixing levels, applied in differing proportions, in an audio mixer. The mixing levels may represent fixed, linear, amplification levels or variable, non-linear, amplification levels.

Preferably, the ATS 100 determines a normalised total output volume for each mixed data stream, and normalises the volume adjustment levels accordingly. This is shown in each of Figures 4 to 6, where the horizontal axis denotes time and the vertical axis denotes normalised total output volume of the mixed media data streams. The normalised total output volume of a mixed media stream may be normalised against various parameters, for example a predetermined acceptable total output volume, or against a measurement of the input volumes of the incoming data streams. For example, it may be normalised against an average of the mean input volumes of each of the received audio data streams. In the mixing modes shown in Figures 4 to 6, the normalised total output volume is normalised against an average of the mean input volumes of each of the received audio data streams, taken over a periodic

or sliding window of between 0.5 to 5 seconds, preferably between 1 and 2 seconds.

In some embodiments, the normalised total output volume is, in certain audio mixing modes, relatively low, for example in the region of 0.4 to 0.9, in this embodiment 0.6. However, due to the increase in the contribution from the audio data stream received from TD1 in the mixed audio data stream from time t1 onwards, the normalised total output volume of the mixed media stream can be seen to increase to a relatively high level, e.g. in the region of 0.9 to 1.5, in this embodiment 1.0. When the normalisation of the total output volume is, for example, conducted in relation to a predetermined acceptable total output volume, or against a measurement of the input volumes of the incoming data streams, this provision of a normalised output volume in default mode which is below the normalised output volume in a non-default mode provides the advantage that, when a recipient of the mixed audio stream receives the non-default mode mixed stream, they do not need to turn down the output volume on their telephone speaker since this is normally set to an acceptable level comparable to the normalised, relatively high level, of the non-default mode mixed stream.

In the mixing modes shown in Figures 4 to 6, the volume adjustment levels used for each of the received audio data streams T1 to T5 are shown schematically, to illustrate their respective sizes.

In the mixing modes shown in Figures 4 to 6, the first respective volume adjustment levels comprise substantially equal volume adjustment levels, whereby each input audio data stream in the plurality (i.e. TD1 to TD5) has a substantially equal volume adjustment level applied to generate the first mixed audio data stream transmitted to TD6.

Such an initial mixed audio data stream, transmitted to TD6, is depicted in each of Figures 4 to 6, in particular between time = 0 and time = t1. It can be seen that each of the audio data streams received from TD1 to TD5 has a substantially equal volume adjustment level (represented by equal area in Figures 4 to 6) applied in the first mixed audio data stream. The audio data

streams received from TD1 to TD5 make a non-zero, substantially equally adjusted, contribution, but there is no contribution from the audio data stream received from TD6.

5 In response to detecting a trigger during the multi-party audio teleconference call, a second mixed audio data stream generated using a second audio mixing mode is transmitted to TD6. The second mixed audio data stream is generated from the same plurality of data streams, but the respective volume adjustment levels are different, the plurality of data streams again being the audio data streams received from TD1 to TD5.

10 Similarly, mixed audio data streams comprising different mixes of received audio data streams will also be transmitted to the other telephony devices TD1 to TD5.

In the exemplary audio mixing mode shown in Figure 4, the second respective volume adjustment levels comprise a relatively high volume adjustment level applied to a first audio data stream received from TD1 and relatively low, substantially equal, volume adjustment levels applied to the other audio data streams in the plurality, i.e. the audio data streams received from TD2 to TD5. The volume adjustment level applied to TD1 may for example be at least double that applied to each of TD2 to TD5. Such a mixed audio data stream is depicted in Figure 4, in particular during the period between time = t1 and time = t2. It can be seen that each of the audio data streams received from TD2 to TD5 has a substantially equal volume adjustment, that there is a relatively higher contribution from the audio data stream received from TD1 and that there is no contribution from the audio data stream received from TD6. In this embodiment, all the audio data streams received from TD1 to TD5 make a non-zero contribution, but there is no contribution from the audio data stream received from TD6.

15
20
25

The volume adjustment levels applied to the audio data streams received from TD2 to TD5 during the period between time = t1 and time = t2 are lower than the volume adjustment levels applied to the audio data streams received from the same sources, TD2 to TD5, during the period between time = 0 and

30

time = t_1 . This is depicted in Figure 4 where the contributions to the mixed audio data stream of TD2 to TD5 during the period between time = 0 and time = t_1 can be seen to be greater (schematically represented as relatively large areas in Figure 4) than the contributions to the mixed audio data stream of TD2 to TD5 during the period between time = t_1 and time = t_2 (schematically represented as relatively small areas in Figure 4).

Figure 4 shows that during the period between time = 0 and time = t_1 , the total volume adjustment level of the mixed media stream is relatively low, for example in the region of 0.4 to 0.9, in this embodiment 0.6. However, due to the increase in the contribution from the audio data stream received from TD1 in the mixed audio data stream during the period between time = t_1 and time = t_2 , the total volume adjustment level of the mixed media stream can be seen to increase to a relatively high level, for example in the region of 0.9 to 1.5, in this embodiment 1.0.

In embodiments, the trigger comprises the mean input volume of the first received audio data stream rising above a first predetermined threshold. This could for example be due to the user of TD1 raising their voice to make an announcement to others in the room in which the user is located. Increasing the contribution due to the voice of the user of TD1 in the mixed audio data streams will enable the announcement to be more easily distinguished in the mixed audio data streams. Increasing the contribution due to the voice of the user of TD1 in the mixed audio data streams could continue for a predetermined time period, and/or after the detection of a predetermined period of silence (e.g. detected as a signal in which mean input volume is below a predetermined threshold) from the voice of the user of TD1, for example until time t_2 . This can be useful for example to ensure that the volume adjustment level for TD1 is increased for the entire announcement rather than just an initial part of the announcement. An audio mixing mode similar, or identical, to that applied prior to time = t_1 may thereafter be used again, as shown in Figure 4.

In the exemplary audio mixing mode shown in Figure 5, the second respective volume adjustment levels comprise relatively high, substantially

equal, volume adjustment levels applied to both a first audio data stream received from TD1 and a second audio data stream received from TD2 and relatively low, substantially equal, volume adjustment levels applied to the other audio data streams in the plurality, i.e. the audio data streams received from TD3 to TD5. The volume adjustment levels applied to TD1 and TD2 may for example be at least double that applied to each of TD3 to TD5. Such a mixed audio data stream is depicted in Figure 5, in particular during the period between time = t1 and time = t2. It can be seen that each of the audio data streams received from TD3 to TD5 has a substantially equal volume adjustment, that there is a relatively higher contribution from the audio data stream received from TD1 and TD2 and that there is no contribution from the audio data stream received from TD6. In this embodiment, all the audio data streams received from TD1 to TD5 make a non-zero contribution, but there is no contribution from the audio data stream received from TD6.

The volume adjustment levels applied to the audio data streams received from TD3 to TD5 during the period between time = t1 and time = t2 are lower than the volume adjustment levels applied to the audio data streams received from the same sources, TD3 to TD5, during the period between time = 0 and time = t1. This is depicted in Figure 5 where the contributions to the mixed audio data stream of TD3 to TD5 during the period between time = 0 and time = t1 can be seen to be greater (schematically represented as relatively large areas in Figure 5) than the contributions to the mixed audio data stream of TD3 to TD5 during the period between time = t1 and time = t2 (schematically represented as relatively small areas in Figure 5).

Figure 5 shows that during the period between time = 0 and time = t1, the total volume adjustment level of the mixed media stream is relatively low, for example in the region of 0.4 to 0.9, in this embodiment 0.6. However, due to the increase in the contribution from the audio data streams received from TD1 and TD2 in the mixed audio data stream during the period between time = t1 and time = t2, the total volume adjustment level of the mixed media stream can be

seen to increase to a relatively high level, for example in the region of 0.9 to 1.5, in this embodiment 1.0.

In these embodiments, the trigger is detected at time t_1 and could comprise the mean input volume of one or more of the first received audio data stream and the second received audio data stream rising above a second predetermined threshold. This could for example be due to the users of TD1 and TD2 having a relatively loud conversation with each other. Increasing the volume adjustment level of the voices of the users of TD1 and TD2 in the mixed audio data streams will enable their conversation to be more easily distinguished in the mixed audio data streams thus helping them to converse more easily above contributions to the mixed audio data streams due to background noise or chatter. Increasing the volume adjustment levels of the voice of the users of TD1 and TD2 in the mixed audio data streams could continue for a predetermined time period after the detection, for example until time t_2 . An audio mixing mode similar, or identical, to that applied prior to time t_1 may thereafter be used again, as shown in Figure 5.

Embodiments comprise storing data representative of one or more predetermined words and monitoring the received audio data streams for the utterance of any of the one or more predetermined words represented in the stored data. Monitoring module 204 of ATS 100 includes speech recognition capabilities, which may be embodied by any suitable speech recognition engine known in the art, such that when a user of a telephony device utters any of the one or more predetermined words, this can be detected in the audio data stream received from that user's telephony device. In such embodiments, the trigger comprises the monitoring detecting utterance of a given one or more predetermined words represented in the stored data in a received audio data stream.

In some embodiments, the given one or more predetermined words are uttered by a user associated with TD1 and are contained in the audio data stream received from TD1. The given one or more predetermined words comprise an identifier for a user associated with TD2. Such embodiments allow two users to

conduct a conversation with each other within the multi-party audio teleconference. The volume adjustment level of the conversation between the two users is increased in the resulting mixed audio data stream which allows their conversation to be distinguished more easily. The identifier could
5 comprise the name of one of the two users such that the conversation can be initiated by one of the users uttering the name of the other user.

In other embodiments, the given one or more predetermined words are uttered by a first user associated with TD1 and the given one or more words comprise an identifier for a second user associated with TD2. The given one or
10 more words further comprise an indication that the first user wishes to conduct a private telephone call with the second user separate to the multi-party audio teleconference call. The second mixed audio data stream is thus generated with second, different respective volume adjustment levels for the audio data streams in the plurality apart from audio data streams received from TD1 and TD2. In
15 the private telephone call, the audio data stream received from TD1 is transmitted to TD2 and the audio data stream received from the TD2 is transmitted to TD1, with substantially equal volume adjustment levels applied in each case – similar to a standard two-party telephone call.

The users of TD1 and TD2 are thus able to have a private telephone call
20 separate to the multi-party audio teleconference call and the multi-party audio teleconference call carries on between the remaining telephony devices (TD3 to TD5). The indication that the first user wishes to conduct a telephone call with the second user separate to the multi-party audio teleconference call may comprise the first user uttering one or more key words, or a phrase, such as
25 “private call” which are predetermined as being operable to trigger a private call, plus an identifier for the second user such as “Hey Joe”.

In the exemplary audio mixing mode shown in Figure 6, the media stream transmitted to TD6 during such a private call between TD1 and TD2 is shown between time = t1 and time = t2. The trigger is detected at time = t1. It
30 can be seen that each of the audio data streams received from TD3 to TD5 has a substantially equal volume adjustment level and contributes substantially

equally to the total volume of the first mixed audio data stream, and that there is no contribution from any of the audio data streams received from TD1, TD2 or TD6. Each of the audio data streams from TD1 and TD2 are thus cut from the mixed audio data stream sent to TD6. As such, the audio data streams received
5 from TD3 to TD5 make a non-zero contribution, but there is no contribution from the audio data stream received from TD1, TD2 or TD6, in the mixed audio data stream sent to TD6. On the other hand, the audio data stream from TD1 is sent, unmixed, to the other participant in the private conversation, TD2, and the audio data stream from TD2 is sent, unmixed, to TD2, during this period.

10 In the exemplary audio mixing mode shown in Figure 6, the second respective volume adjustment levels comprise relatively high, substantially equal, volume adjustment levels applied to each of the audio data streams received from TD3 to TD5 and the audio data streams received from TD1 and TD 2 are cut out in the mixed audio data stream sent to TD6. The respective
15 volume adjustment levels applied to each of the audio data streams received from TD3 to TD5 received during the period time = 0 to time = t1 comprise relatively low, substantially equal, volume adjustment levels. The volume adjustment levels applied to each of TD3 to TD5 during the period time = t1 to time = t2 may for example be at least 1.5 times that applied during the period
20 time = 0 to time = t1. Such a mixed audio data stream is depicted in Figure 6, in particular during the period between time = t1 and time = t2. It can be seen that each of the audio data streams received from TD3 to TD5 has a substantially equal volume adjustment, that there is no contribution from the audio data stream received from TD1 and TD2 and that there is no contribution from the
25 audio data stream received from TD6. In this audio mixing mode, all the audio data streams received from TD3 to TD5 make a non-zero contribution, but there is no contribution from the audio data stream received from TD1, TD2 and TD6.

The volume adjustment levels applied to the audio data streams received from TD3 to TD5 during the period between time = t1 and time = t2 are higher
30 than the volume adjustment levels applied to the audio data streams received from the same sources, TD3 to TD5, during the period between time = 0 and

time = t1. This is depicted in Figure 6 where the contributions to the mixed audio data stream of TD3 to TD5 between time = 0 and time = t1 can be seen to be lower (schematically represented as relatively small areas in Figure 6) than the contributions to the mixed audio data stream of TD3 to TD5 between time =
5 t1 and time = t2 (schematically represented as relatively large areas in Figure 6).

Figure 6 shows that between time = 0 and time = t1, the total volume adjustment level of the mixed media stream is relatively low, for example in the region of 0.4 to 0.9, in this embodiment 0.6. However, due to the cutting out of the contribution from the audio data streams received from TD1 and TD2 in the
10 mixed audio data stream during the period between time = t1 and time = t2, and the relative increase in the contributions from the audio data streams received from TD3 to TD5 in the mixed audio data stream during the period between time = t1 and time = t2, the mixed media stream can be seen to remain at a relatively low level, for example in the region of 0.4 to 0.9, in this embodiment
15 0.6.

The audio data streams from TD1 and TD2 may be re-introduced in the mixed audio data stream sent to TD6 in response to detecting an utterance of a word or phrase, for example “end private” during said private telephone call indicating that the user of TD1 wishes to end the private telephone call with the
20 user of TD2. An audio mixing mode similar, or identical, to that applied prior to time = t1 may thereafter be used again, as shown in Figure 6.

In some embodiments, the user of TD1 has access to computing device 114. In such embodiments, TD1 may for example be a desktop telephone and computing device 114 may be a desktop personal computer. ATS 100 can
25 provide information about a multi-party audio teleconference call being conducted between TD1 and other telephony devices, for example TD2 to TD6. ATS 100 provides a log-in web page via its web interface 210. The user of TD1 is able to log-in to the web page, for example by entering the telephone dialling number of TD1 or other appropriate identifier which ATS 100 will recognise as
30 being associated with the multi-party audio teleconference call being conducted between TD1 and TD2 to TD6.

In embodiments, the web interface 210 allows display of information associated with the multi-party audio teleconference such as visual indicators for each audio teleconference call participant. The web interface 210 allows a user to configure various settings such that the user can configure multi-party audio teleconference services provided via ATS 100.

Example of such configurable settings include the predetermined time period, the predetermined threshold volume adjustment level and the one or more predetermined words described above in relation to multi-party audio teleconference control embodiments.

In embodiments, the web interface 210 allows triggering for switching between mixing modes. For example, instead of a user having to raise their voice in order for the second audio mixing mode to be employed to generate a mixed audio data stream, the user may instead enter appropriate user input via the web interface to instruct ATS 100 manually. Further, a user may manually trigger a direct conversation with another user (with associated boosting of their associated audio data stream in the mixed audio data stream), for example by clicking on a visual indicator associated with that user in the web interface 210. Similarly, a user may manually trigger a private conversation with another user (with their respective audio data stream not being combined into mixed audio data streams transmitted to the remaining participants of the multi-party audio teleconference call), for example by clicking on a further visual indicator associated with that user in the web interface 210.

The above embodiments are to be understood as illustrative examples. Further embodiments are envisaged.

Whilst in the above embodiments, the multi-party teleconference call is between a total of six participants, it should be understood that any number of participants may be accommodated within a teleconference, and that the audio mixing modes shown may have any number of participants within a particular mixing mode. For example a mixing mode in which one participant is emphasized, and a mixing mode in which one or more other participants are de-emphasized (but still heard) may be provided, similar to the embodiment shown

in Figure 4. Further, for example a mixing mode in which two or more participants are emphasized, and a mixing mode in which one or more other participants are de-emphasized (but still heard) may be provided, similar to the embodiment shown in Figure 5. Further, for example a mixing mode in which
5 two or more participants having a private chat are cut out, and in which one or more other participants are relatively emphasized (compared to a default mixing mode) may be provided, similar to the embodiment shown in Figure 6.

For example, telephony device TD1 could comprise a mobile telephony device and telephone switch 104 could comprise a mobile switching centre in a
10 mobile telecommunications network connected to network 102. In such embodiments, instead of the display and user input capabilities of computing device 114 being employed to interface with the web server interface of ACS 100, corresponding display and user input functionality of the mobile telephony device could be employed. Embodiments could be implemented on a mobile
15 telephony device as an application installed on the mobile telephony device.

In further alternative embodiments, telephony device TD1 could comprise a smart desk phone. In such embodiments, instead of the display and user input capabilities of computing device 114 being employed to interface with the web server interface of ACS 100, corresponding display and user input
20 functionality of the smart desk phone could be employed.

Further alternative embodiments could be implemented using a softphone, for example installed on computing device 114; such embodiments need not require a separate telephony device such as TD1.

In the above embodiments, the techniques of embodiments are applied to
25 audio teleconference services; however it should be appreciated that embodiments may also be applied in relation to other teleconference services, such as video teleconference services.

It is to be understood that any feature described in relation to any some embodiments may be used alone, or in combination with other features
30 described, and may also be used in combination with one or more features of any other of the embodiments, or any combination of any other of the

embodiments. Furthermore, equivalents and modifications not described above may also be employed without departing from the scope of the invention, which is defined in the accompanying claims.

Claims

1. A method of handling an incoming telephone call of a first call type during a multi-party teleconference call which is of a second call type,
5 different from said first call type, said method comprising:

establishing a first call leg for transmittal of audio data associated with a multi-party teleconference call between a first telephony device and a multi-party teleconference service, said multi-party teleconference service being capable of connecting said first telephony device with at least two other
10 telephony devices during said multi-party teleconference call;

storing state information indicating that said multi-party teleconference call is of said second call type;

during said multi-party teleconference call, receiving an incoming telephone call setup request for said first telephony device, said incoming
15 telephone call setup request being associated with an incoming telephone call from a further telephony device which is not a participant in said multi-party teleconference call; and

handling said incoming telephone call setup request and the multi-party teleconference call in accordance with said stored state information indicating
20 that said multi-party teleconference call is of said second call type.

2. A method according to claim 1, wherein said handling comprises disabling said first call leg between said first telephony device and said multi-party teleconference service.
25

3. A method according to claim 2, wherein said disabling comprises tearing down said first call leg between said first telephony device and said multi-party teleconference service.

30 4. A method according to claim 2, wherein said disabling comprises maintaining the network resources reserved for establishing said first call leg,

but not transmitting any audio data from said first telephony device to said multi-party teleconference service.

5 5. A method according to any preceding claim, wherein said handling comprises directing said incoming telephone call setup request to said first telephony device, whereby said first telephony device enters a ringing state in relation to said incoming telephone call.

10 6. A method according to any preceding claim, comprising, in response to a user of said first telephony device answering said incoming call, establishing a second call leg for transmittal of audio data associated with said answered telephone call between said first telephony device and said further telephony device.

15 7. A method according to claim 6, comprising:
 in response to termination of said answered telephone call, tearing down said second call leg between said first telephony device and said further telephony device; and
 establishing a third call leg for transmittal of audio data associated with
20 said multi-party teleconference call between said first telephony device and said multi-party teleconference service.

 8. A method according to claim 7, wherein said establishing of said third call leg comprises applying modified signalling to said third call leg such
25 that said third call leg will be automatically answered by said first telephony device.

 9. A method according to claim 7 or 8, wherein said establishing of said third call leg comprises instructing said first telephony device to operate in
30 speakerphone mode.

10. A method according to any preceding claim, wherein said first telephony device is capable of operating either in handset mode or speakerphone mode,
wherein, during said multi-party teleconference call, said first telephony
5 device operates in speakerphone mode, and
wherein said handling comprises instructing said first telephony device to operate in handset mode.

11. A method according to any preceding claim, comprising
10 receiving an incoming multi-party teleconference call setup request from said first telephony device, said incoming multi-party teleconference call setup request comprising an identifier associated with said multi-party teleconference call,
wherein the multi-party teleconference call in respect of which said first
15 call leg is established is identified on the basis of said received identifier.

12. A method according to claim 11, wherein said received identifier comprises a telephone dialling number associated with said first telephony device.
20

13. A method according to claims 4 and 7, wherein establishing said third call leg comprises utilising said maintained network resources.

14. A method according to claim 5, comprising, in response to a user
25 of said first telephony device not answering said first telephony device in said ringing state for a predetermined time period, aborting said handling and re-enabling said first call leg between said first telephony device and said multi-party teleconference service.

15. A method of controlling multi-party teleconference calls in a
30 telecommunications network, said method comprising:

establishing a multi-party teleconference call between at least three telephony devices in said network;

receiving an audio stream from each of said at least three telephony devices in said network;

5 processing said received audio data streams such that the received audio data streams are combined into mixed audio data streams, the received audio data streams having an incoming volume and one or more mixing volume adjustment level settings which identify a volume adjustment level of the received audio data stream, relative to one or more other received audio data

10 streams, in said mixed audio data streams; and

 outputting said mixed audio data streams,

 wherein said method comprises:

 during at least a part of said multi-party teleconference call, transmitting to a first telephony device, of said at least three telephony devices, a first mixed

15 audio data stream generated using a first audio mixing mode, in which said first mixed audio data stream is generated from a plurality of received audio data streams, including a first received audio data stream and a second received audio data stream each having, in said first audio mixing mode, a respective volume adjustment level setting which results in a non-zero contribution to the first

20 mixed audio data stream, the volume adjustment level settings of said first and second received audio data streams having first respective levels in said first audio mixing mode; and

 in response to detecting a trigger during said multi-party teleconference call, transmitting to said first telephony device a second mixed audio data stream

25 generated using a second audio mixing mode, in which said second mixed audio data stream is generated from said plurality of received audio data streams, including said first received audio data stream and said second received audio data stream, said first and second received audio data streams each having, in said second audio mixing mode, a respective volume adjustment level setting

30 which results in a non-zero contribution to the second mixed audio data stream, the volume adjustment level settings of said first and second received audio data

streams having second respective levels in said second audio mixing mode, at least one of which is different to one of said first respective levels.

16. A method according to claim 15, wherein said first respective
5 levels comprise substantially equal volume adjustment levels, and each audio data stream in said plurality has a substantially equal volume adjustment level applied to generate said first mixed audio data stream.

17. A method according to claim 15 or 16, wherein said second
10 respective levels comprise a relatively high volume adjustment level for said first received audio data stream and a relatively low volume adjustment level for said second received audio data stream, in said second audio mixing mode.

18. A method according to claim 17, wherein said second respective
15 levels comprise a plurality of substantially equal volume adjustment levels, including said adjustment level for said first received audio data stream.

19. A method according to any of claims 15 to 18, wherein a total
20 volume adjustment level in said second mixing mode is higher than a total volume adjustment level in said first mixing mode.

20. A method according to any of claims 15 to 19, comprising
transmitting said second mixed audio data stream generated using said second
audio mixing mode to said first telephony device for a predetermined time
25 period after detection of said trigger.

21. A method according to any of claims 15 to 20, wherein said
trigger comprises a mean input volume of said first received audio data stream
rising above a predetermined threshold volume.

22. A method according to any of claims 15 to 20, comprising:
storing data representative of one or more predetermined words; and
monitoring said received audio data streams for the utterance of any of
said one or more predetermined words represented in said stored data,
5 wherein said trigger comprises said monitoring detecting utterance of a
given one or more predetermined words represented in said stored data.

23. A method according to claim 22, wherein said given one or more
predetermined words are uttered by a first user associated with a second
10 telephony device of said at least three telephony devices,
wherein said given one or more predetermined words comprise an
identifier for a second user associated with a third telephony device of said at
least three telephony devices, and
wherein said first audio data stream is received from said second
15 telephony device and said second audio data stream is received from said third
telephony device.

24. A method according to claim 23, wherein said given one or more
predetermined words are uttered by a first user associated with a second
20 telephony device of said at least three telephony devices,
wherein said given one or more words comprise an identifier for a
second user associated with a third telephony device of said at least three
telephony devices,
wherein said given one or more words further comprise an indication
25 that said first user wishes to conduct a telephone call with said second user
separate from said multi-party teleconference call,
wherein said second mixed audio data stream is generated without the
audio data streams received from said second and third telephony devices.

30 25. A method according to claim 24, said method comprising, in
response to detecting said trigger during said multi-party teleconference call,

transmitting said audio data stream received from said second telephony device to said third telephony device and transmitting said audio data stream received from said third telephony device to said second telephony device.

5 26. A method of controlling multi-party teleconference calls in a telecommunications network, said method comprising:

storing data representative of one or more predetermined words or phrases;

10 establishing a multi-party teleconference call between at least three telephony devices in said network;

receiving an audio stream from each of said at least three telephony devices in said network;

15 processing said received audio data streams such that the received audio data streams are combined into mixed audio data streams using one of a plurality of available mixing modes;

monitoring said received audio data streams for the utterance of at least one of said one or more predetermined words or phrases represented in said stored data; and

20 altering a mixing mode applied during the processing of said received audio data streams in response to said monitoring detecting the utterance of said one or more predetermined words represented in said stored data.

27. A method according to claim 26, wherein said altering comprises altering from a first mixing mode in which volume adjustment levels are applied to the received data streams in first respective levels to a second mixing mode in which the volume adjustment levels are applied in second, different, respective levels.

28. A method according to claim 26 or 27, wherein said altering comprises cutting at least a first participant from a mixed audio stream in response to detecting an utterance of a word or phrase indicating that said first

participant wishes to participate in a private telephone call with an identified other participant.

29. A method according to claim 28, comprising re-introducing at
5 least said first participant to the mixed audio stream in response to detecting an utterance of a word or phrase during said private telephone call indicating that said first participant wishes to end said private telephone call with said identified other participant.

10 30. Apparatus adapted to perform the method of any preceding claim.

31. Computer software adapted to perform the method of any of
claims 1 to 29.

15 32. A computer program product comprising a non-transitory computer-readable storage medium having computer readable instructions stored thereon, the computer readable instructions being executable by a computerized device to cause the computerized device to perform a method
20 according to any of claims 1 to 29.

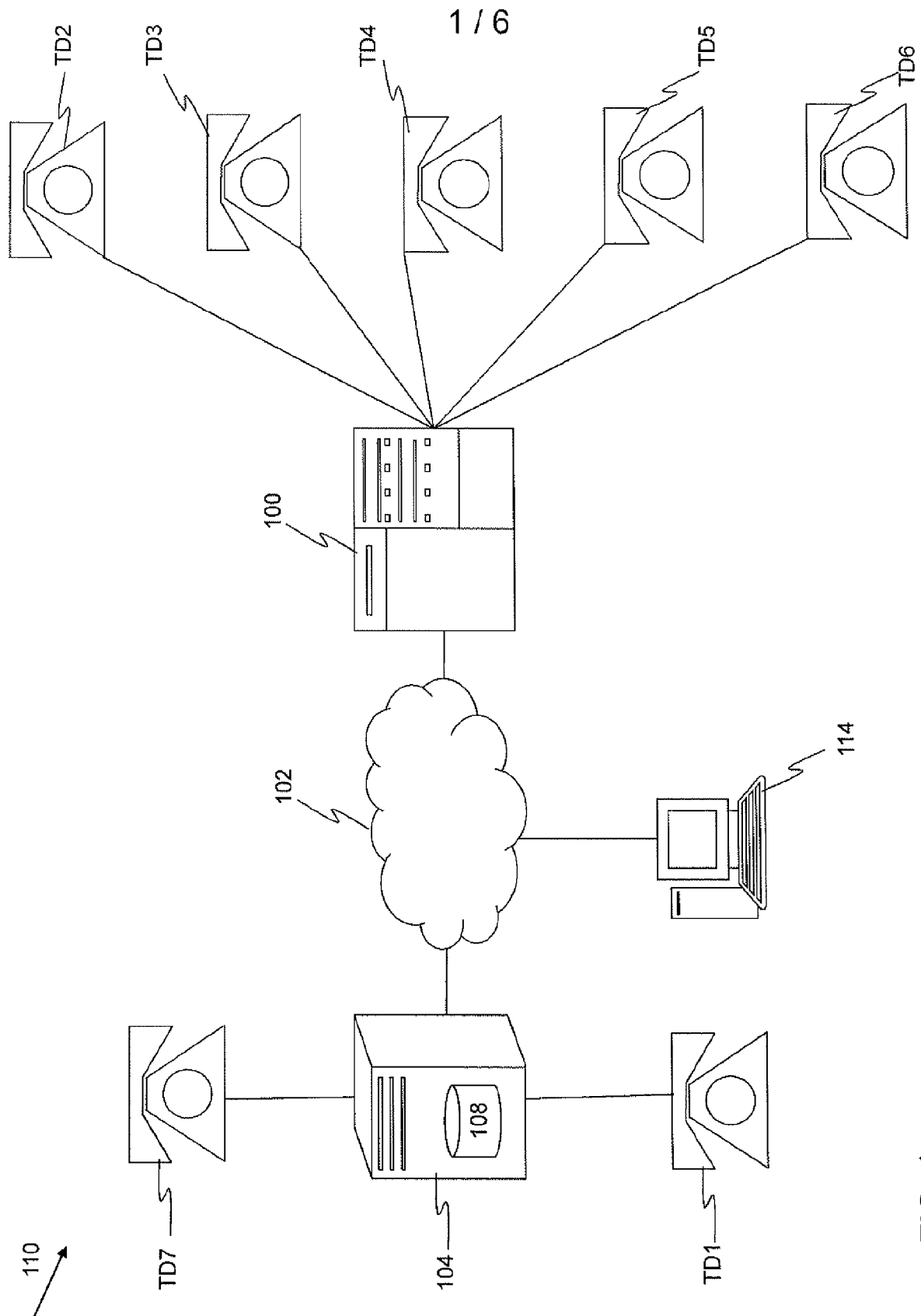


FIG. 1

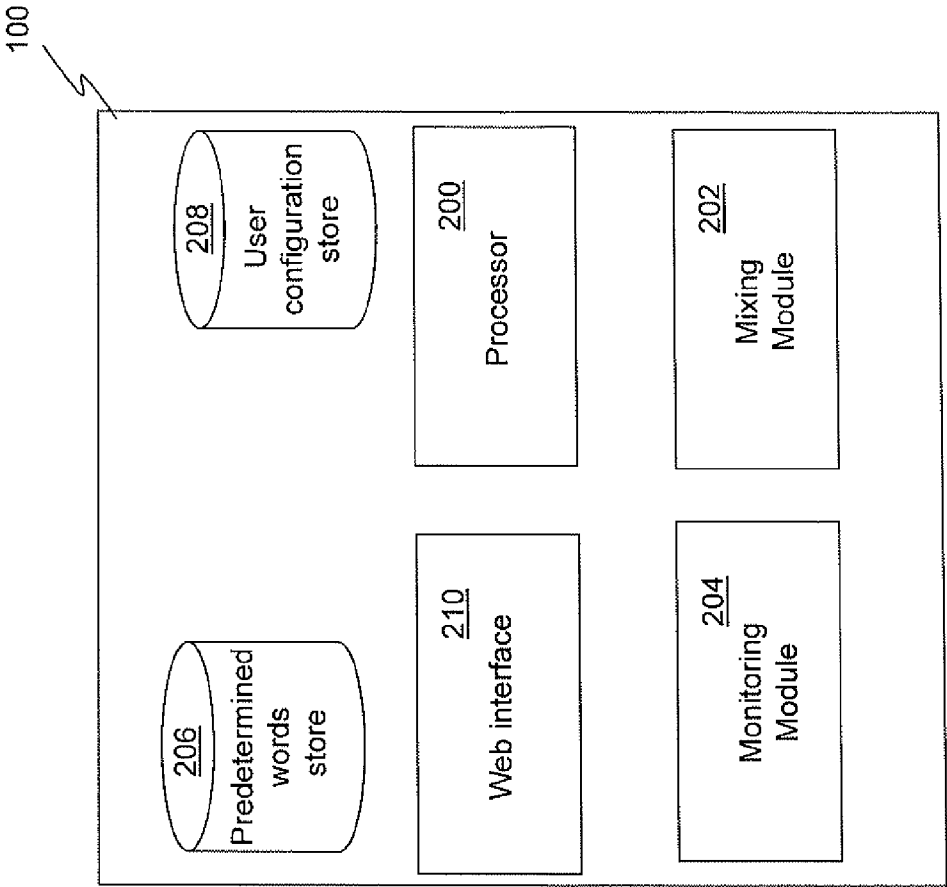


FIG. 2

3 / 6

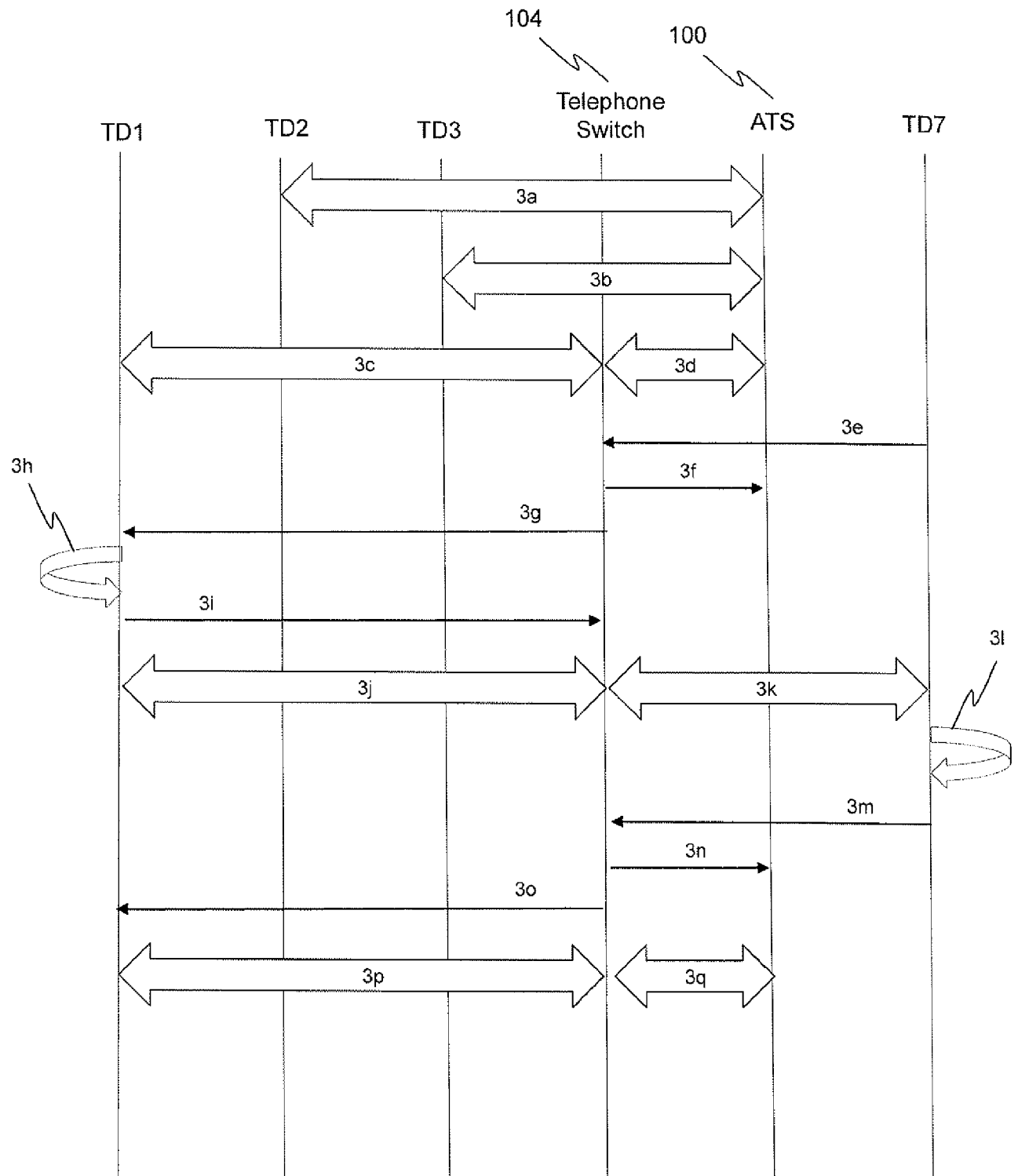


FIG. 3

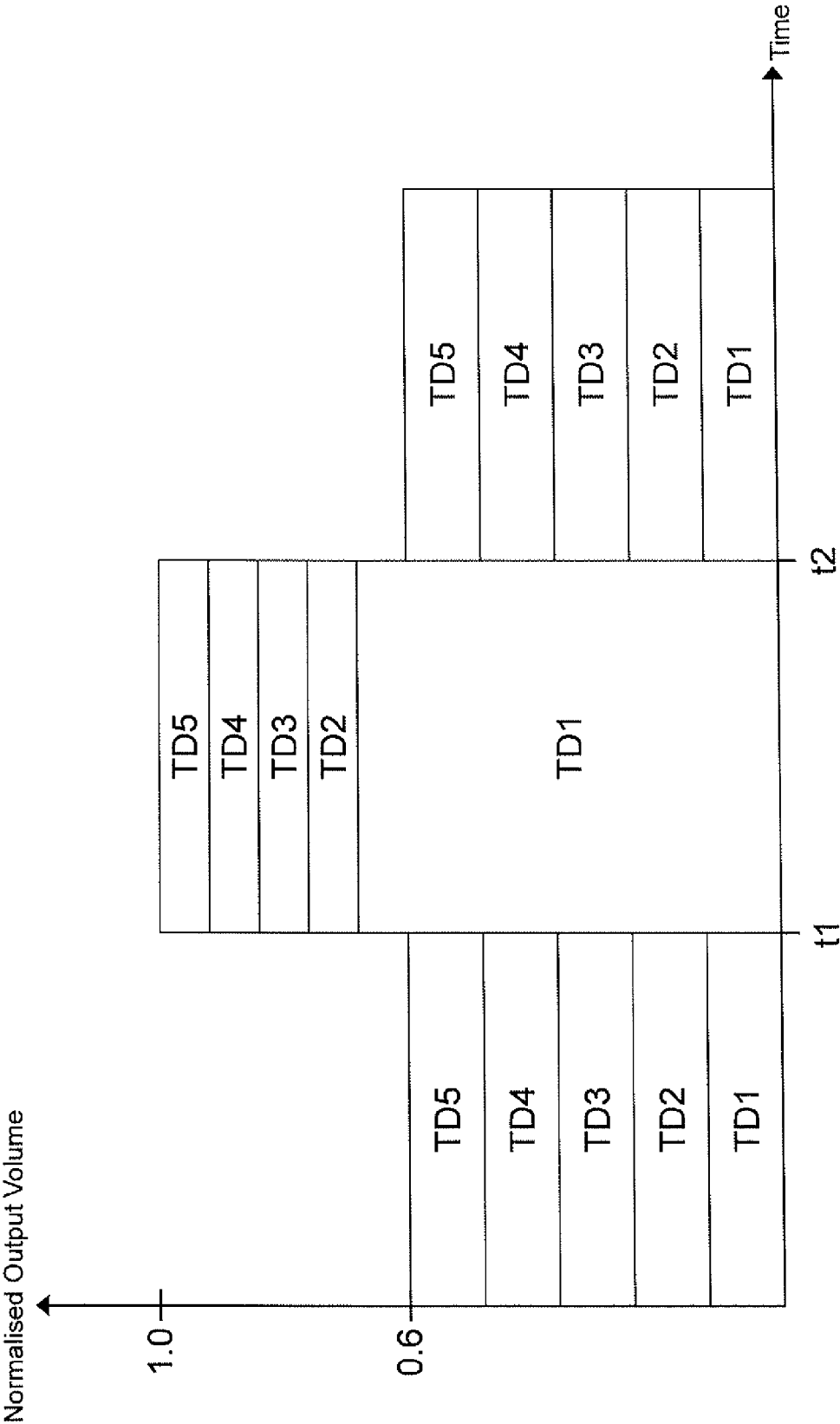


FIG. 4

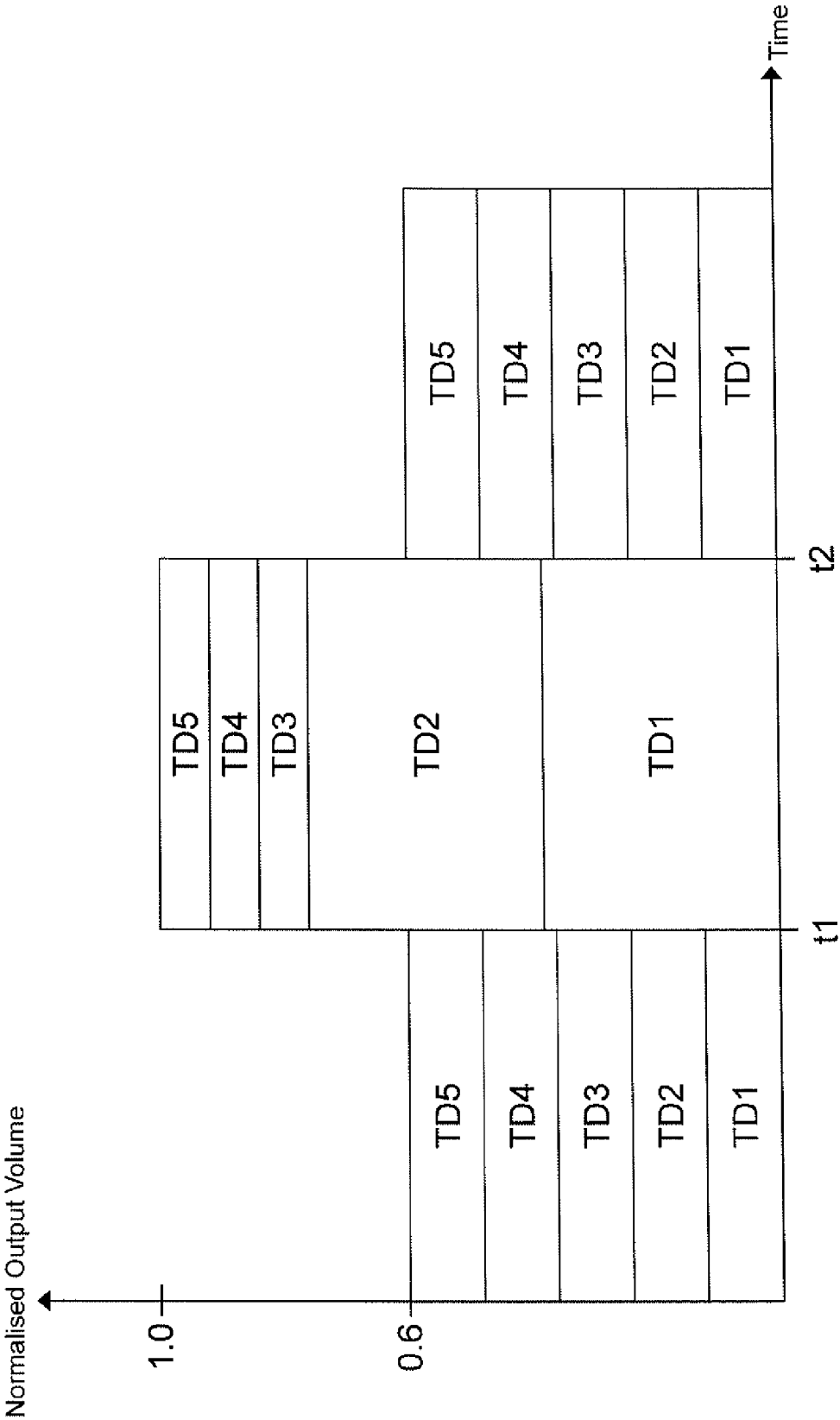


FIG. 5

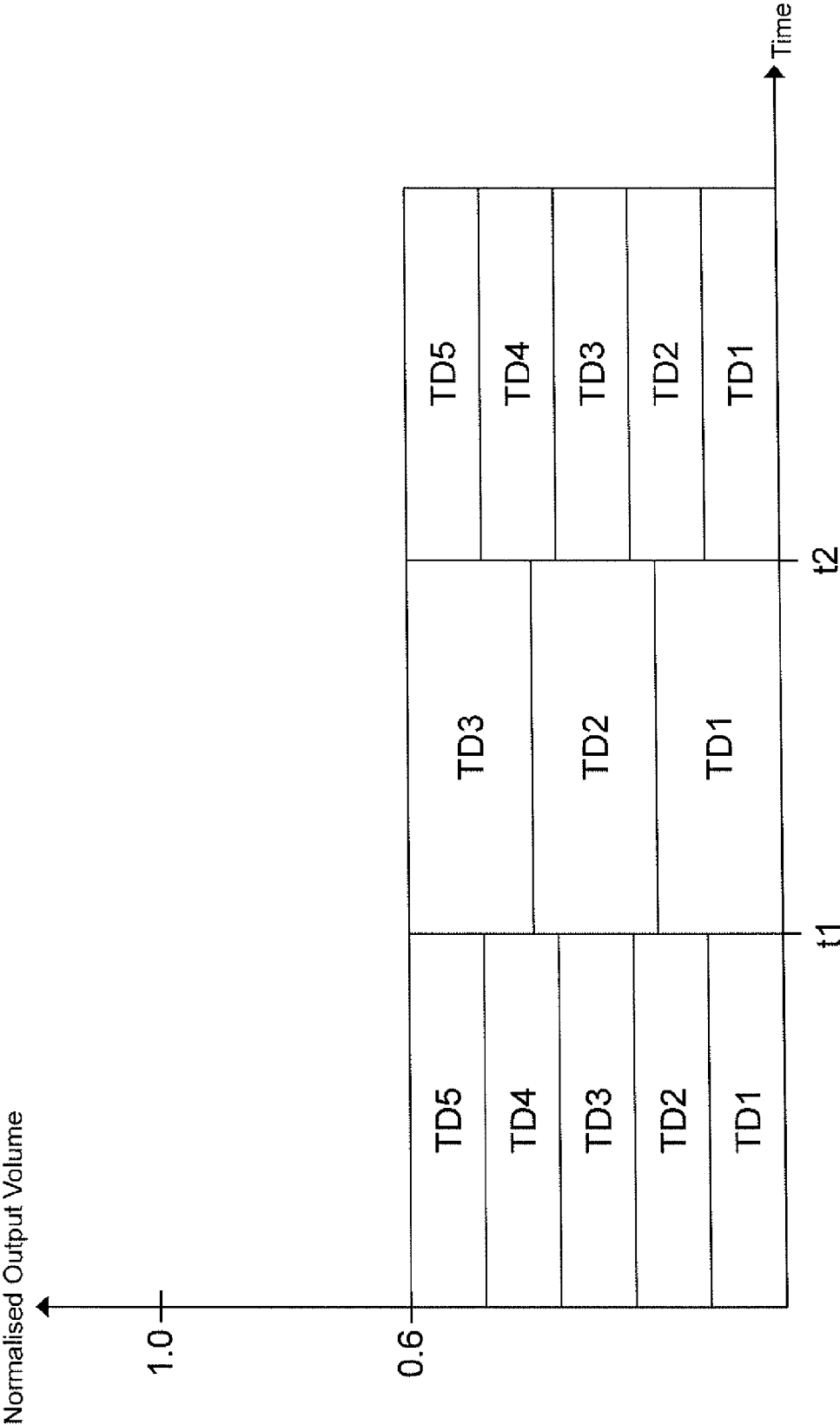


FIG. 6