

Mixing multi-party real-time text for presentation in conference-unaware user agents.

This specification describes media mixer procedures for a multi-party conference server to format real-time text from a number of participants into one single text stream to a participant with a terminal that has no features for multi-party text display. The procedures are intended for implementations using ITU-T T.140 [T.140] for the real-time text coding and presentation.

Contents

1	Short description	2
2	Functionality goals and drawbacks	2
3	Definitions	3
4	Presentation level procedures	4
4.1	Structure.....	4
4.2	Action on reception.....	5
4.2.1	Delete BOM	5
4.2.2	Insert in buffer.....	5
4.2.3	Maintain a display counter.....	5
4.2.4	Combination characters	6
4.2.5	Erasure.....	6
4.3	Initial action in the session	6
4.4	Maintaining a waiting time per participant.....	6
4.5	Storing time of reception for each character	6
4.6	Initial assignment of the Current Participant	6
4.7	Actions on assignment of a Current Participant.....	6
4.8	Actions on transmission and during the session.....	7
4.9	Actions when a Switch Reason has occurred	7
4.10	Handling of Control codes	7
4.10.1	Handling of Graphic Rendition SGR.....	8
5	Display examples	9
6	Summary of configurable parameters	9
7	Transport level considerations (informational).....	11
7.1	RTP transmission payload.....	11
7.2	Filtering of BOM characters	11
7.3	RTP Keep-alive.....	11

7.4	RTP and RTCP characteristics for multi-party session	11
8	References.....	12
9	Author.....	12
10	Acknowledgement.....	12
11	Document history.....	12

1 Short description

The media mixer procedures described here are intended to make real-time text from a number of call participants be coordinated into one text stream to a terminal originally intended for two-party calls.

A conference server is supposed to apply the procedures.

The procedures may also be applied on a terminal for display of multiple streams of real-time text in one area.

The intention is that text from each participant shall be displayed in suitable sections so that it is easy to read, and text from one active participant at a time is sent and displayed in real-time.

The receiving terminal is assumed to have one display area for received text.

The display is arranged by this procedure in a text chat style, with a name label in front of each text section where switch of source of the text has taken place.

When more than one participant transmits text at the same time, the text from only one of them is transmitted directly to the receiving terminals. Text from the other participants is stored in buffers in the conference server for transmission at a later time, when a suitable situation for switch of current transmitter can take place.

2 Functionality goals and drawbacks

The procedures are intended to make best efforts to present a multi-party text conversation on a terminal that has no awareness of multi-party calls. There are some obvious drawbacks, and a terminal designed with multi-party awareness will be able to present multi-party call contents in a more flexible way. Only two parties at a time will be allowed to display added text in real-time, while the other parties' produced text will need to be stored in the multi-party server for a moment awaiting a suitable occasion to be displayed. There are also some cases of erasure that will not be performed on the target text but only indicated in another way. Even with these drawbacks, the procedure provides an opportunity to display text from more than two parties in a smooth and readable way.

This specification does not introduce any new protocol element, and does not rely on anything else than basic two-party terminal functionality with presentation level according to ITU-T T.140 [T.140]. It is a description of a best current practice for mixing and presentation of the real-time text component in multi-party calls with terminals without multi-party awareness.

The procedures are applicable to scenarios, when the conference focus and a User Agent have not gone through any successfully completed negotiation about conference awareness for the real-time text medium neither on the transport level, nor on the presentation level.

3 Definitions

Active participant:	Any user sending text, or being in a pending period.
BOM	Byte-Order-Mark, the Unicode character FEFF in UCS-16.
Buffer:	A buffer intended for unsent text collected per participant.
Contributing participants:	The participants selected to contribute to the text stream sent to the recipients. By default all participants except the recipient are contributing participants for transmission to the recipient.
Current participant:	The participant for whom text currently is transmitted to the recipient in real time.
Current Recipients:	By default all participants.
Display Counter:	A counter for the number of displayable characters in a participant's buffer or in the current entry. Used for controlling how far erasure may be performed.
Erasure replacement	A character to be displayed when an erasure was done, but the text to erase is not reachable on the multi-party display. Default "X".
Message delimiter:	Character(s) forming the end of an imagined message. A configurable set of alternatives, consisting by default of: Line Separator, Paragraph Separator, CR, CRLF, LF.
Pending period:	A configurable time period of inactivity from a participant, by default set to 7 seconds after each reception of characters from that participant, evaluated as current time minus time stamp of latest entered character.
Sentence delimiter:	Characters forming end of sentence: A configurable set of alternatives, by default consisting of: dot ".", question mark "?" and exclamation mark "!" followed by a space.
Label:	A readable unique name for a participant, created by the server from a suitable source related to the participant, e.g. part of the SIP Display name, surrounded by the Label delimiters. The label should have a settable maximum length, with 12 being the default.
Label delimiters	A configurable set of characters at the edges of the Label, by default being a left bracket [at the leading edge and a closing bracket] followed by a space at the trailing edge.
Line Separator	Unicode UCS-16 2028. Used to request NewLine in Real-Time Text.
Maximum waiting time:	The maximum time any participant's text shall be allowed to wait for transmission, by default set to 20 seconds.
Recipient:	The terminal receiving the mixed text stream.
SGR	Select Graphic Rendition, a control code to specify colours etc.
Switch Reason:	A set of reasons to switch Current Participant, consisting of the following -Waiting time higher for any other participant than the current participant combined with any of the following states:

- A message delimiter was the latest transmitted item
- A sentence delimiter was the latest transmitted item
- A Pending Period has expired and still no text has been transmitted
- The Maximum Waiting time has expired followed by a Word Delimiter or an expired Time Extension.

Waiting time:	The time the first character in queue for transmission from a participant has been waiting in a buffer for transmission. The granularity shall be 0.3 Seconds or finer.
Word delimiter:	Character forming end of word: space
Time extension:	A configurable short extension time allowed after the Maximum waiting time during which a suitable moment for switching Current Participant is awaited, by default set to 7 seconds.

4 Presentation level procedures

The conference server applies these mixing procedures to text transmitted to all call participants who have not gone through a completed negotiation for conference awareness in real-time text presentation.

All the participants and the conference server use real-time text conversation presentation coding according to ITU-T T.140 [T.140]. A consequence is that real-time text transmissions are UTF-8 coded, with control codes selected from ISO 6429 [ISO 6429].

The description is from the conference server point of view.

4.1 Structure

The real-time text mixer structure described here is supposed to be placed in the media path so that it is implemented with one mixer per recipient. A mixer contains buffers for temporary storage of text intended for the recipient. Each mixer has one buffer for each contributing participant. A set of status variables is maintained per buffer and is used in the mixer actions. The mixer logic decides for each moment which participant's buffer content is to be sent on to the recipient. By default, the recipient does not contribute text to its own mixer. Text transmitted by a participant is usually displayed locally and will only cause confusion if it appears also in received text.

If there is a reason, own text can be configured to be transmitted also to the participants. That can enable a simplification of the mixer design to have only one common set of buffers instead of a set per recipient. That simplification will however hamper the flow of the conversation severely and is therefore NOT RECOMMENDED.

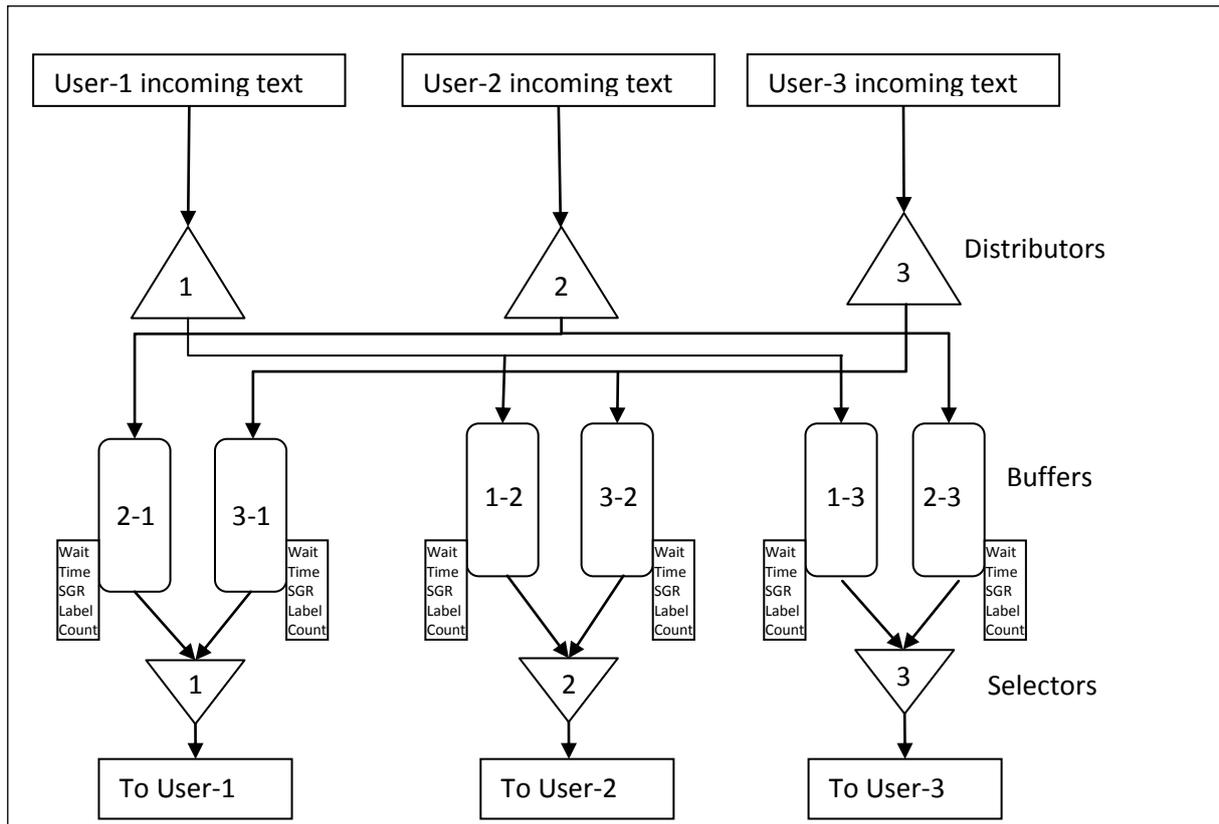


Figure 1: Structure of text mixers for a three-party call.

4.2 Action on reception

This description of the mixer is valid per recipient.

Text from each contributing participant is checked for a set of characteristics on reception.

4.2.1 Delete BOM

BOM characters are deleted.

4.2.2 Insert in buffer

Resulting text is put into the contributing participant's buffer in the receiving participant's mixer.

4.2.3 Maintain a display counter

For each text character that will take a position on the receiving display, a Display Counter for each participant is increased by one.

Note that there is one T.140 real-time text item that consists of two characters, but is regarded to be a unit and therefore increase the Display Counter with one only.

That is

CRLF.

Furthermore, the following control codes are regarded units that shall not take any position on the receiving display and shall therefore not increase the Display Counter:

0098 string 009C	(SOS-ST strings)
ESC 0061	(INT)
009B Ps 006D	(the SGR code, with special handling described below)
BEL	(Alert in session)

See the section on control codes below for details.

4.2.4 Combination characters

Also note that it is possible to use combination characters in Unicode. Such combination characters contain more than one character part. They shall only increase the Display Counter with one. The combination characters mainly have components in the series 0300 – 0361 and 20D0 – 20E1.

4.2.5 Erasure

If the control code for erasure, BS, is received, the following shall be done: If the Display Counter is 0, an Erasure Replacement character, by default being “X” is inserted in the buffer instead of the erasure, to mark that erasure was intended in earlier transmitted entries. (this matches traditional habits in real-time text when participants sometimes type XXX to indicate erasure they do not bother to make explicit). If the Display Counter is >0, then the counter is reduced by one, and the erasure control code BS put into the buffer.

4.3 Initial action in the session

BOM shall be sent initially to the recipients in the beginning of the session.

4.4 Maintaining a waiting time per participant

The time that text has been in the buffer is maintained as the waiting time for each buffer. A granularity of 0.3 seconds is sufficient.

4.5 Storing time of reception for each character

Each character that is stored in a buffer shall be assigned with a time stamp indicating its time of reception. A granularity of 0.3 seconds is sufficient. This time stamp is used for calculation of idle time and waiting time in the evaluation of switch reasons.

4.6 Initial assignment of the Current Participant

The first contributing participant to send text in the session is assigned to be the Current Participant.

4.7 Actions on assignment of a Current Participant

When a participant becomes the Current Participant, the following initial actions shall be performed:

1. Scanning transmissions and timers for a Switch Reason is inactivated.
2. The Current Recipients are set so that all transmissions go to the new set of Current Recipients (See definition).
3. A Line Separator is transmitted if the switch reason was any other than a message delimiter.
4. The Label is transmitted
5. Any stored SGR code is transmitted
6. Scanning transmissions and timers for a Switch Reason is activated.
7. Text in the buffer is transmitted, recalculating and setting the waiting time for each transmitted character based on the time of reception of next character in the buffer.

If a switch occurs during transmission from the buffer, the remaining buffer contents is maintained and transmission can continue next time this transmitter becomes the current participant. Any text

entered into the buffer for the current participant is after that sent to the recipient until a Switch Reason occurs

4.8 Actions on transmission and during the session

Transmissions are checked for control codes to act on at transmission as described below in the section about handling of control codes and such actions are performed.

When the scanning of transmission and timers for a Switch Reason is active, the timers and the transmission to the recipient is analyzed for detection if a Switch Reason has occurred. See the definition of Switch Reasons for details.

4.9 Actions when a Switch Reason has occurred

If a Switch Reason has occurred, then the following actions shall be performed:

1. The Display Counter of the Current Participant is set to zero
2. If there is an SGR code stored for the Current Participant, a reset of SGR shall be sent by the sequence SGR 0 [009B 0000 006D].
3. A participant with the longest waiting time is assigned to be the Current Participant, and the procedure for assignment of a Current Participant described above is performed.

4.10 Handling of Control codes

The following control codes are specified by ITU-T T.140. Some of them require consideration in the conference server. Note that the codes presented here are expressed in UCS-16, while transmission is made in UTF-8 transform of these codes. Other sections specify procedures for handling of specific control codes in the conference server.

Name	Code	Usage
BEL	0007	Bell, provides for alerting during an active session.
BS	0008	Back Space, erases the last entered character.
NEW LINE	2028	Line separator.
CR LF	000D 000A	A supported, but not preferred way of requesting a new line.
INT	ESC 0061	Interrupt (used to initiate mode negotiation procedure).
SGR	009B Ps 006D	Select graphic rendition. Ps is rendition parameters specified in ISO 6429.
SOS	0098	Start of string, used as a general protocol element introducer, followed by a maximum 256 bytes string.
ST	009C	String terminator, end of SOS string.
ESC	001B	Escape - used in control strings.
Byte order mark	FEFF	Zero width, no break space, used for synchronization.
Missing text mark	FFFD	Replacement character, marks place in stream of possible text loss.

Code for message border, useful, but not mentioned in T.140:

New Message	2029	Paragraph separator
-------------	------	---------------------

4.10.1 Handling of Graphic Rendition SGR.

The following procedure shall be followed in order to let the participants control the graphic rendition of their entries without disturbing other participants' graphic rendition.

The text stream sent to a recipient shall be monitored for the SGR sequence. The latest conveyed SGR sequence is also stored as a status variable for the recipient. If the SGR 0 code initiated from the current participant is transmitted, the SGR storage shall be cleared.

5 Display examples

The following pictures are examples of the view on a participant's display.

Conference	Alice
	I will arrive by TGV.
[Bob]:My flight is to Orly.	Convenient to the main station.
[Eve]:Hi all, can we plan for the seminar.	
[Bob]:Eve, will you do your presentation on Friday?	
[Eve]:Yes, Friday at 10.	
[Bob]: Fine, wo	We need to meet befo

Figure 2 : Alice who has a conference-unaware client is receiving the multi-party real-time text in a single-stream. This figure shows how a coordinated column view MAY be presented on Alice's device.

[Alice] Hi, Alice here.	^
[Bob] Bob as well.	
[Eve] Hi, this is Eve, calling from Paris. I thought you should be here.	
[Alice] I am coming on Thursday, my performance is not until Friday morning.	
[Bob] And I on Wednesday evening.	
[Eve] we can have dinner and then take a walk	
<Eve-typing> But I need to be back to the hotel by 11 because I need	-
	-
of course, I underst	v

Figure 3 shows a conference view with real-time text preview. Bob's text is buffering until a Current switch reason.

6 Summary of configurable parameters

A number of configurable parameters are described in this specification. This table provides a summary of the parameters on presentation level. A service provider implementing a multi-party service may want to set specific values on these parameters to adapt the characteristics of the service. It is possible to control them per recipient, if desired.

Parameter	Purpose	Possible values	Default value	Comment
Current Recipients	Control if participant shall get their own text.	Exclude or Include Current Participant	Exclude	Own transmissions are usually displayed sufficiently locally
Erasur replacement	Character to show erasure, when erasure cannot be done	Character	X	May need to have other value for other than Latin script.
Message delimiter	Detection of suitable place in text for switching Current Participant	List of Unicode editing codes	Line Separator, Paragraph Separator, CR, CRLF, LF	Other than Latin based scripts may have other conventions
Pending period	Inactivity timer for detection of time to Switch Current Participant	Time in seconds	7	Longer times may cause inefficient transmission. Shorter time may cause unwanted switching cutting lines of thought inconveniently
Sentence delimiter	Characters forming end of sentence	List of delimiters.	. or ? or ! followed by a space	Used for deciding on a position in the text to switch Current Participant according to configured logic.
Label length	Length of label put in front of or above entry.	Number of characters	12	Includes any surrounding characters
Label delimiters	Set of characters at the edges of the label	Two strings. One in the beginning, one after.	[] followed by a space	It may be valid to include a Line Separator instead of the space
Maximum waiting time	The maximum time any participant's text shall be allowed to wait for transmission	Seconds	20	After this time a Switch will be forced within the Time Extension
Word delimiter	Delimiter for words	List of characters	Space	Used for detection of suitable switch position if Maximum Waiting time has passed.

Parameter	Purpose	Possible values	Default value	Comment
Time extension	Time for maximum further waiting for a Switch Reason	Time in seconds	7	After this time a Switch is forced.

7 Transport level considerations (informational)

This chapter is purely informational.

It contains advice on how to handle the transport level of the real-time text transmission, for the case that uses RTP [RTP] with the IETF RFC 4103 [RFC 4103] payload format for real-time text.

This transport is the most common transport protocol for real-time text in SIP, and some aspects of it is therefore briefly described here.

7.1 RTP transmission payload

Transport of real-time text can be made by RTP, with payload as specified in IETF RFC 4103 [RFC 4103]. Redundancy is preferred and negotiated in SDP.

Transmission in RTP is then done according to the recommendation in RFC 4103 in 300 ms samples.

On reception of text from the participants, any actions for recovering from packet loss, or marking loss in the text stream is made. This is a transport level function that may be provided by the RTP transport for real-time-text.

Text from only one participant is included as primary data in each RTP packet.

7.2 Filtering of BOM characters

Any BOM characters received from the participants are deleted from the text on the transport level.

7.3 RTP Keep-alive

A valid keep-alive transmission is made to the participants who need it during periods of no text transmission. By default all participants are regarded to need it. Valid methods for keep-alive are a BOM character in UTF-8 transform, an empty T.140 block, or an RTCP packet if RTCP and RTP share the same port [keep-alive]. The interval between keep-alive shall be configurable and set to 25 seconds by default. This is a lower level function that may be handled in the RTP transport for real-time-text. The RTCP mechanism is optional to implement, but preferred. If it is implemented it is used if the participant supports it, otherwise the empty T.140 block is used.

7.4 RTP and RTCP characteristics for multi-party session

The procedure described in this specification is an RTP mixer in the terms of the Real Time Protocol [RTP]. Therefore, the CSRC parameter in RTP packets includes the SSRC of the participant providing the primary data in that packet.

RTCP reports are also transmitted, with SDP packets, containing the CNAME and NAME parameters. If NAME is provided, the Label could be derived from the NAME.

The Label according to the presentation level specification above is to be sent with RTP Packets separate from the buffer contents and is marked in the RTP packets with a CSRC for the conference server.

The receiver MAY use these parameters to enhance the multi-party display.

8 References

- [T.140] ITU-T T.140 Application protocol, text conversation (including amendment 1.)
- [RFC 4103] IETF RFC 4103 RTP Payload for text conversation
- [RTP] IETF RFC 3550 RTP: A Transport Protocol for Real-Time Applications.
- [RFC 4579] IETF RFC 4579 SIP Call Control – Conferencing for user agents.
- [ISO 6429] ISO 6429 Control functions for coded character sets.
- [UTF-8] IETF RFC 3629 UTF-8, a transformation format of ISO 10646
- [Unicode] The Unicode Consortium, "The Unicode Standard – Version 4.0"
- [ISO 10 646-1] ISO 10 646 Universal multiple-octet coded character set (UCS)
- [keep-alive] IETF draft-ietf-avt-app-rtp-keepalive
- [UCS-16] See ISO 10 646-1

9 Author

Gunnar Hellström
Omnitor
Stockholm, Sweden
www.omnitor.se
gunnar.hellstrom@omnitor.se

10 Acknowledgement

This resource was developed with funding in part from the National Institute on Disability and Rehabilitation Research, U.S. Department of Education, RERC on Telecommunications Access, grant # H133E090001 . However, the contents do not necessarily represent the policy of the Department of Education, and you should not assume endorsement by the Federal Government.

11 Document history

Date	Main change
29 April 2011	Modification to enable a switch to occur during transmission from buffer contents.
17 April 2011	The reasons to switch current participant was changed to being described as states instead of events in order to let incoming text cause switching if the state for switching is already valid. A superfluous New Line transmission after Message delimiter was also deleted in the logic.
6 March 2011	Transport chapter made informational and

	separate
4 March 2011	First publication