Technical Disclosure Commons

Defensive Publications Series

May 29, 2019

Text Data Embedded into a Voice Call

Jim Sun

Aaron Lee

Gary Jian

Text Data Embedded into a Voice Call

Abstract:

This publication describes systems and techniques directed to embedding text data into a voice call. A first wireless-communication device, such as a smartphone, includes a text-voice manager application. The first wireless-communication device performs the voice call with a second wireless-communication device that includes a peer text-voice manager application. The first wireless-communication device, by executing instructions of the text-voice manager application, performs operations that include receiving text data for a text message, converting the text data to a voice message, detecting silent periods in a stream of a voice transmission, and embedding the voice message into the detected silent periods in the stream of the voice transmission. The second wireless-communication device receives the stream of the voice transmission, detects a pattern in the stream of the voice transmission that indicates a presence of the voice message, de-converts the voice message to the text data, and embeds the text data into a text message displayed on the second wireless-communication device.

Keywords:

embedded text message, embedded SMS message, voice call, text-to-speech, automatic speech recognition, voice channel, text channel, Physical Uplink Shared Channel (PUSCH), Physical Downlink Shared Channel (PDSCH)

Background:

Today, it is common for a person to perform a voice call and transmit text data from his smartphone to another person's smartphone using a wireless network such as a Fourth Generation Long Term Evolution (4G LTE) or a Fifth Generation New Radio (5G NR) wireless network. The

wireless network typically allocates resources of wireless communication channels such as a Physical Uplink Shared Channel (PUSCH) and a Physical Downlink Shared Channel (PDSCH) for the communication to perform the voice call and transmit the text data.

In some instances, while a person is having a conversation with the other person during the voice call, it may be beneficial for him to transmit information to the other person as part of a dedicated text message (*e.g.*, a dedicated short message service (SMS) text message) so that the other person isn't burdened with remembering or recording the information from the conversation. Examples of such instances include an instance in which the user wants to share an address or a phone number for social reasons or an instance in which the user wants to share medical information during an emergency. In such instances, transmitting the information as part of the dedicated text message can provide clarity and conciseness over the information being provided through conversation.

Transmitting the information as part of the dedicated text message, however, poses several challenges. A first challenge is that the user may not have access to a provider that offers text message communication services (and if the user does have access to such a provider, there may be fees or expenses associated with using the text message communication services). A second challenge is that the transmitting the information as part of the dedicated text message may require using allocated resources of the PUSCH and the PDSCH, which may compromise efficient use of resources available to the wireless network.

Description:

This publication describes systems and techniques directed to embedding text data into a voice call. A first wireless-communication device, such as a smartphone, includes a text-voice manager application. The first wireless-communication device performs the voice call with a

second wireless-communication device that includes a peer text-voice manager application. The first wireless-communication device, by executing instructions of the text-voice manager application, performs operations that include receiving text data for a text message, converting the text data to a voice message, detecting silent periods in a stream of a voice transmission, and embedding the voice message into the detected silent periods in the stream of the voice transmission. The second wireless-communication device receives the stream of the voice transmission, detects a pattern in the stream of the voice transmission that indicates a presence of the voice message, de-converts the voice message to the text data, and embeds the text data into a text message displayed on the second wireless-communication device.

Fig. 1, below, illustrates an example wireless-communication device and elements of the wireless-communication device that supports text data embedded into a voice call.

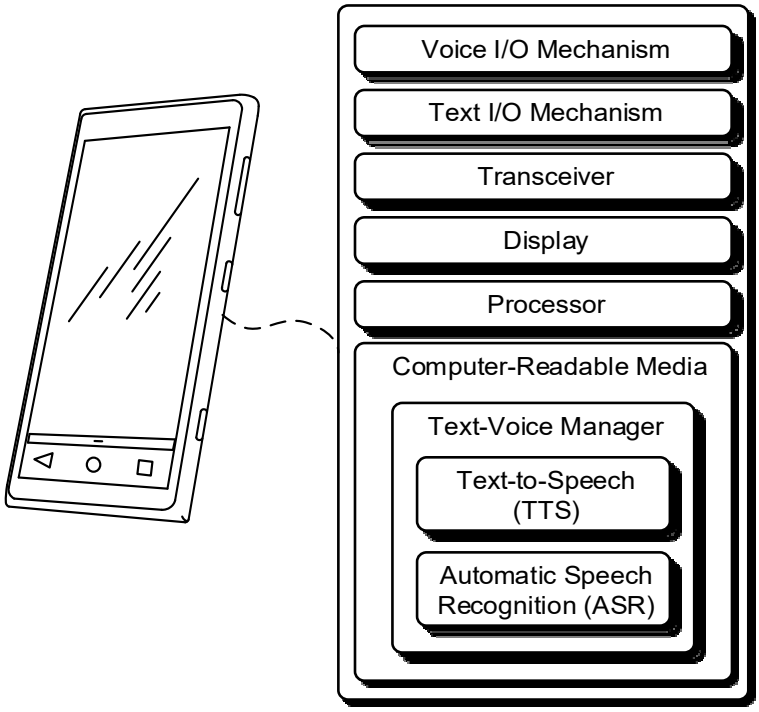

Fig. 1

As illustrated, the wireless-communication device is a smartphone. However, other wireless-communication devices (*e.g.*, a tablet, a laptop computer, a wearable device, or the like) can also support text data embedded into a voice call. The wireless-communication device includes a voice input/output mechanism (*e.g.*, a speaker), a text input/output mechanism (*e.g.*, a text keyboard that may be a touchscreen keyboard), a transceiver (*e.g.*, a 4G LTE or 5G NR transceiver for transmitting data to, and receiving data from, access points of a wireless network), a display (*e.g.*, a Light Emitting Diode (LED) display or Liquid Crystal Display (LCD)), and a processor. The processor may be a single-core processor or a multiple-core processor composed of a variety of materials, such as silicon, polysilicon, high-K dielectric, copper, and so on.

The wireless-communication device also includes a computer-readable medium (CRM) storing a text-voice manager application. The CRM may include any suitable memory or storage device such as random-access memory (RAM), static RAM (SRAM), dynamic RAM (DRAM), non-volatile RAM (NVRAM), read-only memory (ROM), or Flash memory. The text-voice manager application includes a text-to-speech (TTS) module and an automatic speech recognition (ASR) module that, when executed by the processor of the wireless-communication device, cause the wireless-communication device to perform operations described within this document. In some instances, the text-voice manager application may be offered by a service provider and downloaded to the wireless-communication device.

Fig. 2 illustrates an example technique associated with embedding text data into a voice call. The technique is performed using elements of a first wireless-communication device (*e.g.*, the wireless-communication device of Fig. 1) performing a voice call through a wireless network (*e.g.*, a 4G LTE or 5G NR wireless network) with a second wireless-communication device (*e.g.*,

a peer wireless-communication device configured similarly to the wireless-communication device of Fig. 1).

Fig. 2

In the example technique of Fig. 2, a first user (*e.g.*, the user depicted in the left side of Fig. 2) who is using the first wireless-communication device executes a voice-call application to perform the voice call with the second user (*e.g.*, the other user depicted in the right side of Fig. 2) who is using the second wireless-communication device (*e.g.*, a wireless-communication device that is a peer to the first wireless-communication device). The first user determines that he would like to transmit, to the second user, text data that includes an address and a phone number. As illustrated, the first user inputs text data for a text message into the first wireless-communication device. The user can input the text data into the first wireless-communication device using a text input mechanism of the first wireless-communication device (*e.g.*, the text I/O mechanism of Fig. 1). As one example, the text input mechanism may include a touchscreen keyboard that the first wireless-communication device, under direction of the voice-text manager application, displays in a window that is associated with the voice-call application and presented on a display of the first wireless-communication device (*e.g.*, the display of Fig. 1). As another example, the text input mechanism may include a selectable menu linked to a contact list or a pre-written

message that the first wireless-communication device, under direction of the voice-text manager application, displays in another window that is associated with the voice-call application.

The technique continues, where the first wireless-communication device performs operations that include (i) converting the text data to a voice message (*e.g.*, through the processor of Fig. 1 executing the TTS module of Fig. 1) and (ii) detecting silent periods in a stream of a voice transmission originating from the first wireless-communication device and being transmitted, via a voice channel of the wireless-communication network as part of the voice call, to the second wireless-communication device. The voice channel may use resources (*e.g.*, resource blocks that span frequency and time domains) associated with a Physical Uplink Shared Channel (PUSCH) and a Physical Downlink Shared Channel (PDSCH) as scheduled and allocated by the wireless-communication network.

In some instances, the silent periods may correspond to one or more predetermined periods of time during which the voice input/output mechanism of the first wireless-communication device detects no audible input from the first user. In other instances, the silent periods may be dynamic and vary with the start and stop of audible inputs received from the first user by the voice I/O mechanism.

The first wireless-communication device, under direction of the voice-text manager application, then embeds the voice message (*e.g.*, the voice message of the converted text data) into the detected silent periods in the stream of the voice transmission. To do so, the first wireless-communication device may use a variety of techniques that include introducing lag-times or buffer-times to the stream of the voice transmission to accommodate embedding the voice message into the stream of the voice transmission and avoid contention amongst resources of the voice channel.

Continuing with the technique, the second wireless-communication device, under direction of a peer voice-text manager application (*e.g.*, corresponding to the voice-text manager application of Fig. 1) receives the stream of the voice transmission. As the second wireless-communication device receives the stream of the voice transmission, it detects a pattern that indicates a presence of the voice message. In some instances, detecting the pattern may include the second wireless-communication device, under the direction of the peer voice-text manager application, recognizing digital markers inserted into the stream of the voice transmission by the first wireless-communication device. In other instances, detecting the pattern may include the second wireless-communication device, under the direction of the ARS module included in the peer voice-text manager application, recognizing keywords or phrases in the stream of the voice transmission.

Under direction of the ARS module included in the peer voice-text manager application, the second wireless-communication device then de-converts the voice message to the text data (*e.g.*, the text data that the first user of the first wireless-communication device determined he would like to send to the second user). De-converting the voice message to the text data, in some instances, can include "stitching" the text data together with other text data that is de-converted from other voice messages embedded in the stream of the voice transmission.

The second wireless-communication device then embeds the text data into a text message and displays the text message on a display of the second wireless-communication device (*e.g.*, a peer display corresponding to the display of Fig. 1). In some instances, displaying the text message may include displaying the text message using a text messaging application on the second wireless-communication device, while in other instances displaying the text message may include displaying the text message in a window generated by the peer voice-text manager application

(*e.g.*, a window inserted into a peer voice call application being displayed on the second wireless-communication device).

Fig. 3, below, illustrates operations of the technique as performed with respect to the first wireless-communication device. Under certain conditions, the operations, or portions of the operations, as illustrated, may be re-ordered or performed by entities other than the first wireless-communication device (*e.g.*, a provider of a cloud-based service or the wireless network may offer and execute the operations of the text-voice manager application of the first wireless-communication device).

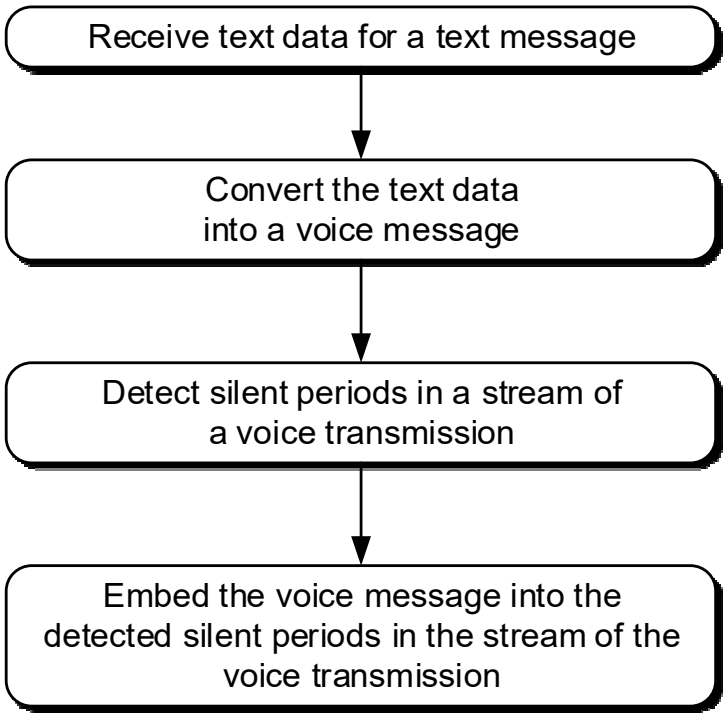

Fig. 3

Fig. 4, below, illustrates operations of the technique as performed with respect to the second wireless-communication device. Under certain conditions, the operations, or portions of the operations, as illustrated, may be re-ordered or performed by entities other than the second

wireless-communication device (*e.g.*, a provider of a cloud-based service or the wireless network may offer and execute the operations of the text-voice manager application of the first wireless-communication device).

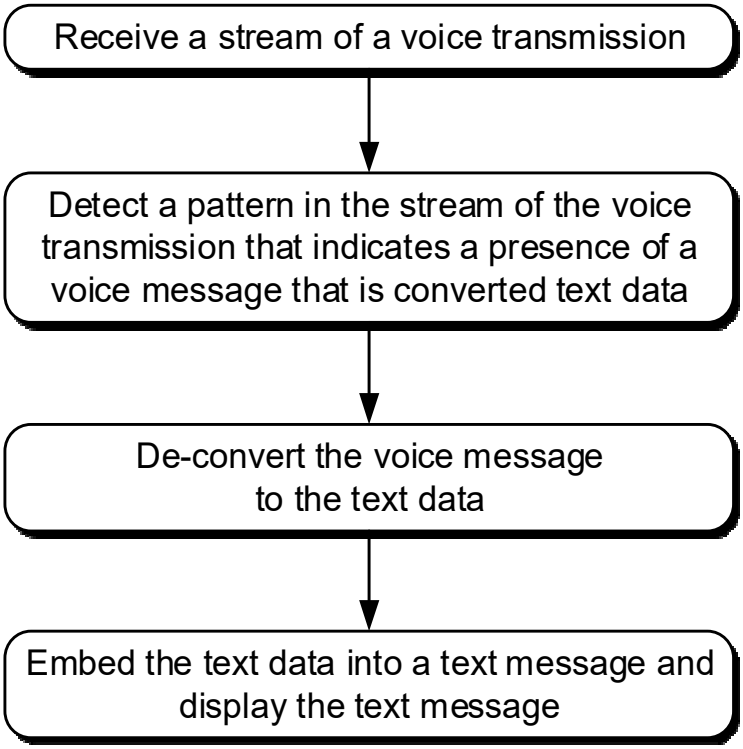

Fig. 4

Further to the systems and the techniques directed to text data embedded into a voice call, as described above, a user may be provided with controls allowing the user to make an election as to both if and when systems, programs or features described herein may enable collection of user information (*e.g.*, information about a user's social network, social actions or activities, profession, preferences, or current location), and if the user is sent content or communications from a server. In addition, certain data may be treated in one or more ways before it is stored or used, so that personally identifiable information is removed. For example, a user's identity may be treated so

that no personally identifiable information can be determined for the user, or a user's geographic location may be generalized where location information is obtained (such as to a city, ZIP code, or state level), so that a particular location of a user cannot be determined. Thus, the user may have control over what information is collected about the user, how that information is used, and what information is provided to the user.

In conclusion, text data embedded into a voice call, using techniques and systems as described above, overcomes challenges that may be associated with transmitting information as part of a dedicated text message.

References:

[1] Tang, Weiqi, Kazim A. Das, Yi He, and Ling Jin. Method and Apparatus for Live Call Text-to-Speech. US Patent 9311911, filed July 30, 2014 and issued April 12, 2016.

[2] Donoghue, Karen, and William C. Barnett. Method and Apparatus for Augmenting Voice Data on a Mobile Device Call. US Pub. 2006/0193448, filed December 2, 2005 and published August 31, 2006.