CNG 495

Cloud Computing

Fall - 2023

Term Project Proposal

Autospeech

Muhammad Somaan 2528404

Mohamed Eldeeb 2643120

Topic

Android automatic voice translator app using Google Cloud services.

Project Explanation

The aim behind this project is to provide an easy and simple method of communicating with people using different languages. The app will allow capturing voice input, translating it, and playing the translated voice output. It is also planned to have a live speech mode, which would auto-translate in real time without user input (similar to transcription). The uses of this project are many, such as providing a communication method for tourists in a foreign country, or acting as an educational tool for second language learners.

Continuous speech might sometimes heavily affect the already translated words of a sentence. The live speech mode offers a solution for this by allowing automatic translation and transcribing of the voice input, which would result in a more comprehensible speech output. Here is an example of this feature:

Transcribed sentence	Input voice (EN)	Output text (TR)
t1	the quick fox jumps…	hızlı tilki atlar…
t2	the quick fox jumps over…	hızlı tilki üzerinden atlıyor…
t3	the quick fox jumps over the lazy dog.	hızlı tilki tembel köpeğin üzerinden atlar.

As can be seen, the translated speech differs greatly after each word is added. Because of this, when the output voice is spoken, a text transcription of the translated sentence will be displayed, and if there was any text that needed retranslation (e.g. t1 to t2, or t2 to t3), it will be re-spoken separately of the already spoken words. After the input sentence is finished (or is structurally complete), the fully translated voice output can be re-spoken.

This feature tries to mimic the behaviour of human translators, who would often wait until the sentence they are hearing is finished, or they would repeat part of it to re-explain it better.

Cloud Delivery Models

The following Google Cloud SaaS (Software as a Service) components will be used in this project:

- Speech-to-Text API: Converts audio to text by applying powerful neural network models.
- Translation API: Integrates text translation into a website or application.
- Text-to-Speech API: Synthesizes natural-sounding speech by applying powerful neural network models.

Diagrams

The following diagram (figure 1) describes the data flow between the application and cloud services.

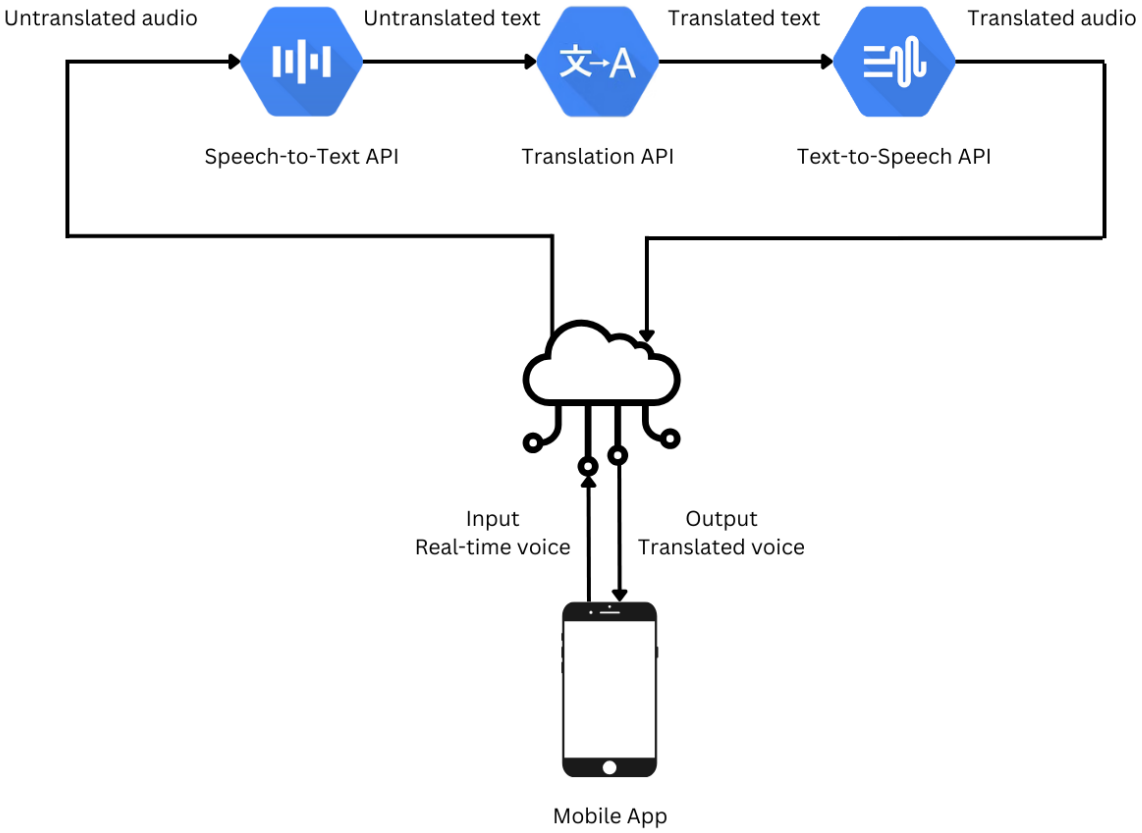

Figure 1: Data flow between application and cloud services

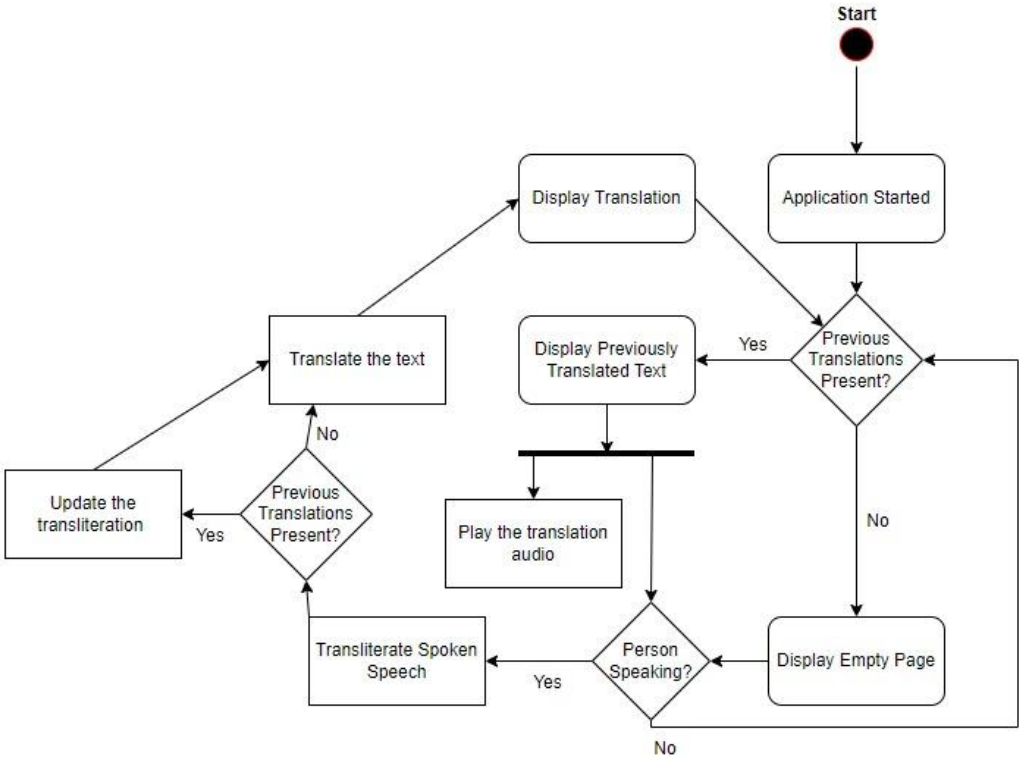

Figure 2: Process of transliterating and translating speech input

Figure 2 is a computation diagram that shows the process of transliterating and translating speech input.

Expected Contribution of Team Members

Here is the expected contribution of each team member:

Team member 1:

Creating the UI of the application as well as interactive features.

Team member 2:

Communicating with the cloud services and formatting data, and implementing the live speech mode.

References

Google. (n.d.). Cloud Speech-to-Text API. https://console.cloud.google.com/marketplace/product/google/speech.googleapis.com

Google. (n.d.). Cloud Translation API. https://console.cloud.google.com/marketplace/product/google/translate.googleapis.com

Google. (n.d.). Cloud Text-to-Speech API. https://console.cloud.google.com/marketplace/product/google/texttospeech.googleapis.com