

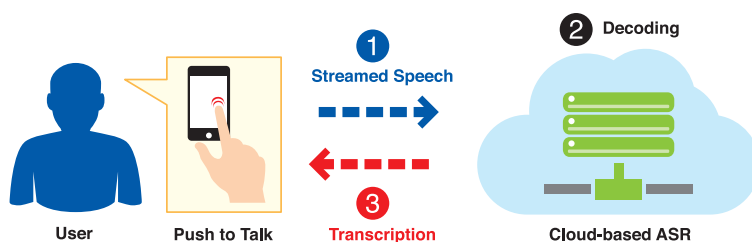
IMAGINATION TO REALITY

Cloud-based Automatic Speech Recognition



Technology Overview

Over the past few decades, automatic speech recognition technology has become increasingly mature and accurate, especially for commonly spoken languages such as English and Mandarin. The rapid expansion of the mobile device market is driving new demand for compact, small form factor speech transcription, with applications to voice command and control, voice dialing, and dictation. The focus of the technology is therefore to offer high accuracy general-purpose speech recognition for common languages with high market demand, as well as specialize for local and regional languages which large commercial vendors do not support.



- The above diagram shows the overall architecture of the cloud-based automatic speech recognition (ASR) system.
- An end user/client connects to the designated server and streams continuous speech for recognition.
- The server sends back the best (or n-best) decoding transcription instantly after each utterance is finished streaming.
- The server will support multiple clients at the same time, depending on the available processing power and memory as well as the configured target language.

Technology Features

- Server-based automatic speech recognition platform providing high accuracy transcription over the network.
- Support for Unix, Windows, Android and iOS platforms.
- Simple to use Applications Programming Interface with available bindings and Software Development Kit in multiple languages.
- Secure and reliable transcription communication.
- Support for English, Mandarin and Malay languages, with future plans to scale to many more Southeast Asian languages.

Potential Applications

- Voice search services for mobile devices.
- Automatic speech-to-text transcription services.
- Easy integration for applications requiring an automatic speech recognition component.

Benefits

A fast and highly accurate speech decoding algorithm is made possible using state-of-the-art weighted finite state transducers. Server-client architecture allows a continual improvement on back-end algorithms and models, without affecting the installed user-base, leading to increased accuracy. Multi-threaded implementation also allows a large scale deployment with appropriate scaling.

Contact Us

Industry Development Department
Institute for Infocomm Research (I2R)
1 Fusionopolis Way, #21-01 Connexis (South Tower), Singapore 138632
Tel: (65) 6408 2000 Fax: (65) 6776 1378 Email: inddev@i2r.a-star.edu.sg
www.i2r.a-star.edu.sg