

Pengembangan model investigasi keterlambatan dan kerangka kerja untuk mendapatkan kinerja ketepatan waktu aplikasi pengenalan wicara berbasis TCP = Development of delay investigation model and framework for real time performance of TCP based speech recognition application

Asril Jarin, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=20453965&lokasi=lokal>

Abstrak

ABSTRAK

Implementasi sistem pengenalan wicara berbasis jaringan, seperti: Internet, akan mengalami degradasi yang disebabkan oleh kehilangan dan keterlambatan data. Sebagian aplikasi pengenalan wicara lebih memilih keterlambatan data demi ketersediaan seluruh data wicara secara kalimat-per-kalimat. Ketersediaan seluruh data akan membantu sistem pengenalan wicara menjaga kinerja akurasi yang semestinya. Akan tetapi, pengguna biasanya lebih menghendaki batas keterlambatan yang wajar sebagai syarat dari kinerja memuaskan aplikasi. Dalam disertasi ini, sebuah model analitik dikembangkan untuk menginvestigasi batas waktu-tunda wajar sebuah skema aplikasi pengenalan wicara berbasis TCP yang menempatkan sebuah pemanggil data wicara di klien. Batas waktu-tunda wajar didefinisikan sebagai keterlambatan maksimal yang diperkenankan dalam pengiriman seluruh data setiap kalimat wicara via TCP. Pengembangan model dilakukan melalui analisis transien berdasarkan kajian model discrete-time Markov dari multi-media streaming via TCP. Selanjutnya, sebuah metode perhitungan dari model distribusi keterlambatan paket aliran TCP pada kondisi steady-state diuji dengan membandingkan hasil-hasil perhitungannya dengan hasil investigasi dari model berbasis analisis transien. Hasil perbandingan menunjukkan bahwa analisis transien adalah metode investigasi yang lebih tepat. Pada target penelitian berikutnya, sebuah kerangka kerja menggunakan protokol HTTP/2 plus Server Sent Event SSE diajukan sebagai solusi ketepatan waktu aplikasi pengenalan wicara berbasis TCP. Kerangka kerja ini dibangun berdasarkan pada kerangka kerja pengenalan wicara full-duplex yang dikembangkan dengan menggunakan teknologi WebSocket. Berdasarkan pada hasil percobaan, aplikasi menggunakan HTTP/2 plus SSE memiliki angka perbandingan kinerja latensi sebesar 3,6 lebih baik daripada aplikasi menggunakan WebSocket. Walaupun angka ini masih lebih kecil daripada angka kualitatif perbandingan ketepatan waktu yang lebih baik, yakni sebesar 5, ada beberapa alasan dikemukakan yang berasal dari keunggulan-keunggulan fitur-fitur HTTP/2 dalam mengurangi latensi aplikasi dan juga dari kelemahan WebSocket bila ditempatkan dalam jaringan dengan proxy server, untuk menyimpulkan bahwa kerangka kerja aplikasi menggunakan HTTP/2 plus SSE dapat menjadi alternatif lebih baik daripada kerangka kerja aplikasi dengan WebSocket.

<hr />

ABSTRACT

Implementation of network based speech recognition, such as Internet, will suffer degradation due to packet loss and delays. Most of network speech recognition applications prefer to tolerate delay in order to receive all speech data completely that is delivered sentence by sentence. The availability of all speech data helps the application to save the expected accuracy of recognition in case of no packet loss. However, users practically require an acceptable delay to have satisfactory performance of the application. In this research,

an analytical model is developed to investigate the acceptable delay of TCP based speech recognition that employs a speech segmenter at the client. The acceptable delay is defined as a maximum allowable delay in sending all data for each speech sentence via TCP. For the purpose of model development, there are two analytical methods, i.e., transient analysis and steady state analysis. In the transient analysis, the investigation model is developed based on the discrete time Markov model of multimedia streaming via TCP, whereas in the steady state analysis, the investigation uses a calculation method of packet delay distribution model. Furthermore, the results of transient analysis experiment are compared with the calculation of packet delay distribution model at the steady state. The comparison shows that the transient analysis is more appropriate method of investigation. Next work, a framework using HTTP 2 protocol plus Server Sent Event SSE is proposed as a real time solution for TCP based speech recognition applications. This framework is developed on the basis of a full duplex speech recognition framework using WebSocket. Based on the experimentation results, the application of HTTP 2 plus SSE has a comparison factor of latency performance in amount of 3.6 better than the application of WebSocket. Although this factor is still smaller than a qualitative factor 5 that can state a better latency performance, there are some reason from the advantages of HTTP 2 features in reducing latency as well as from the limitation of WebSocket in a network with proxy server, to conclude that the framework of HTTP 2 plus SSE is a better alternative than the framework using WebSocket.