A network gateway is a mechanism which provides protocol translation and/or validation of network traffic using the metadata contained in network packets. For media applications such as Voice-over-IP, the portion of the packets containing speech data cannot be verified and can provide a means of maliciously transporting code or sensitive data undetected. One solution to this problem is through Voice Activity Detection (VAD). Many VAD's rely on time-domain features and simple thresholds for efficient speech detection however this doesn’t tell us much about the signal being passed. More sophisticated methods employ machine learning algorithms, but train on specific noises intended for a target environment. Validating speech under a variety of unknown conditions must be possible; as well differentiating between speech and non-speech data embedded within the packets. A real-time speech detection method is proposed that relies only on a clean speech model for detection. Through the use of Gammatone filter bank processing, the Cepstrum and several frequency domain features are used to train a One-Class Support Vector Machine which provides a clean-speech model irrespective of environmental noise. A Wiener filter is used to provide improved operation for harsh noise environments. Greater than 90% detection accuracy is achieved for clean speech with approximately 70% accuracy for SNR as low as 0dB.

Major: Master of Science

Educational Career:
Bachelor’s of Bachelor of Science in Electrical Engineering, BS, 2009, University of Central Florida

Committee in Charge:
Wasfy Mikhael, Chair, Electrical Engineering and Computer Science
Parveen Wahid, Electrical Engineering and Computer Science
Aman Behal, Electrical Engineering and Computer Science

Approved for distribution by Wasfy Mikhael, Committee Chair, on March 1, 2013.

The public is welcome to attend.