

# REAL-TIME SPEAKER-INDEPENDENT LARGE VOCABULARY CONTINUOUS SPEECH RECOGNITION

Xialong Li

Dr. Yunxin Zhao, Dissertation Supervisor

## ABSTRACT

In this dissertation, a real-time decoding engine for speaker-independent large vocabulary continuous speech recognition (LVCSR) is presented. Three indispensable and correlated performance measurements -- accuracy, speed, and memory cost, are carefully considered in the system design. A novel algorithm, Order-Preserving Language Model Context Pre-computing (OPCP) is proposed for fast Language Model (LM) lookup, resulting in significant improvement in both overall decoding time and memory space without any decrease of recognition accuracy. The time and memory savings in LM lookup by using OPCP became more pronounced with the increase of LM size.

By using the OPCP method and other optimizations, our one-pass LVCSR decoding engine, named TigerEngine, reached real-time speed in both tasks of Wall Street Journal 20K and Switchboard 33K, on the platform of a Dell workstation with one 3.2 GHz Xeon CPU. TigerEngine is to be used in automatic captioning for Telehealth.