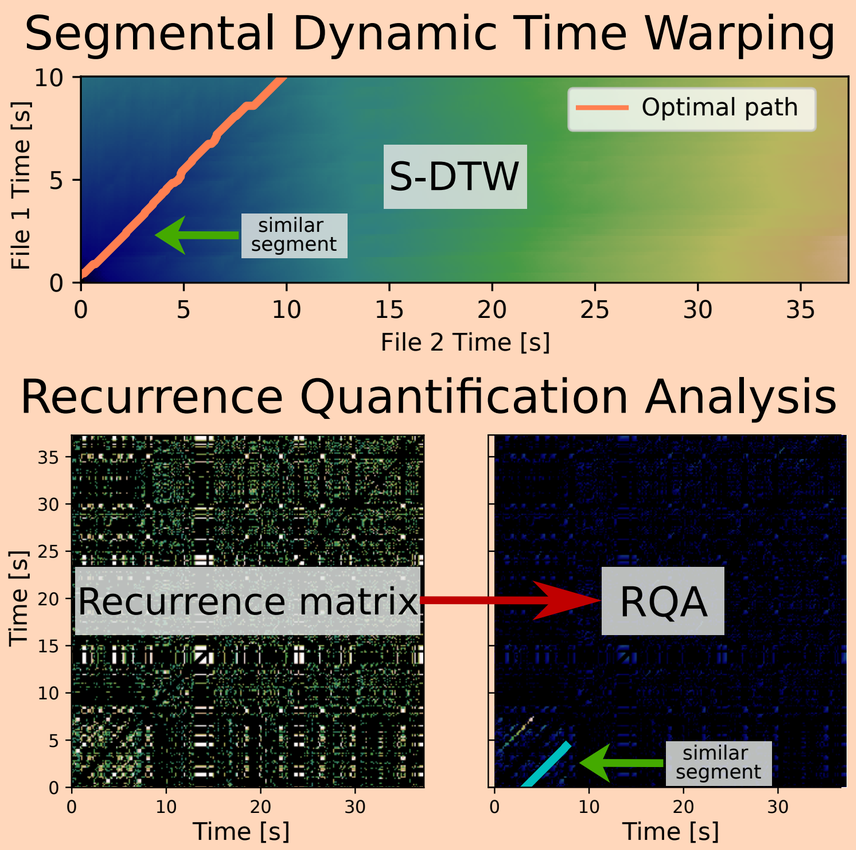
Detection of Pre-Recorded Messages in Speech

9

Author: Dominik Boboš
Supervisor: prof. Dr. Ing. Jan Černocký

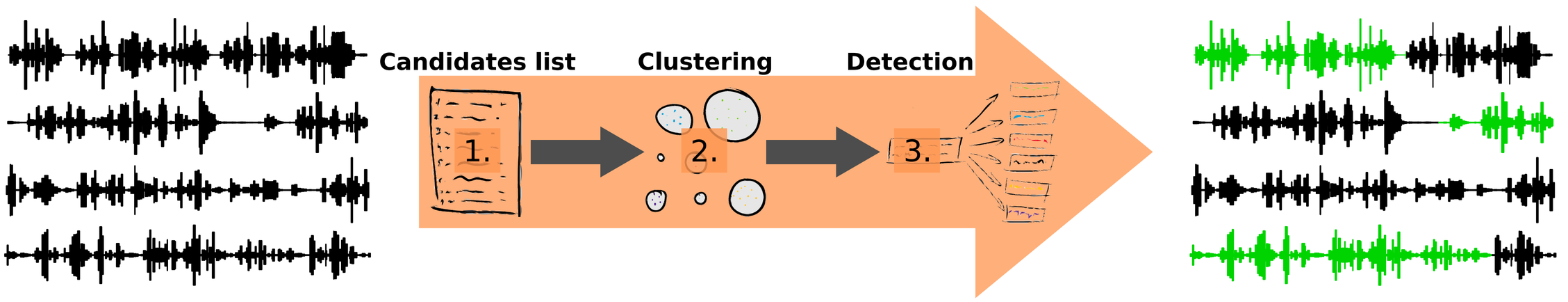
Motivation

- Analysis of a large dataset of telephone conversations is a difficult. task. It requires a lot of space in memory and takes loads of time.
- More information in calls is redundant and repetitive such as telephone operator pre-recorded messages (*"Thank you for calling, please leave a message.").*
- Detection of these pre-recorded messages increases productivity for Law Enforcement Agencies and call centres.



Proposed sulution

- Created simulated dataset by mixing real telephone operator pre-recorded messages into recordings with calls.
- Standard techniques like Segmental Dynamic Time Warping (S-DTW), Recurrence Quantification Analysis (RQA) and Fuzzy Phoneme String Matching (FSM) were combined together in the following order:
  1. First-pass analysis by applying RQA
  2. S-DTW clustering of the candidates from 1.
  3. Detection of the pre-recorded messages by FSM



Results

- Average time of processing one recording is **2.81** seconds.
- Created system is **64 times** faster than manual processing (average recording is 3 minutes long).
- The system is completely unsupervised with an accuracy of **97.9 %** (EER 2.1 %).