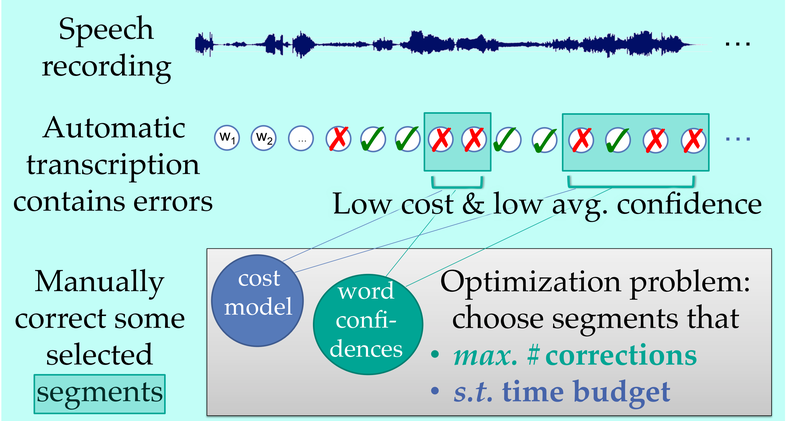
Contact: matthias.sperber@kit.edu

On-the-fly User Modeling for Cost-sensitive Correction of Speech Transcripts

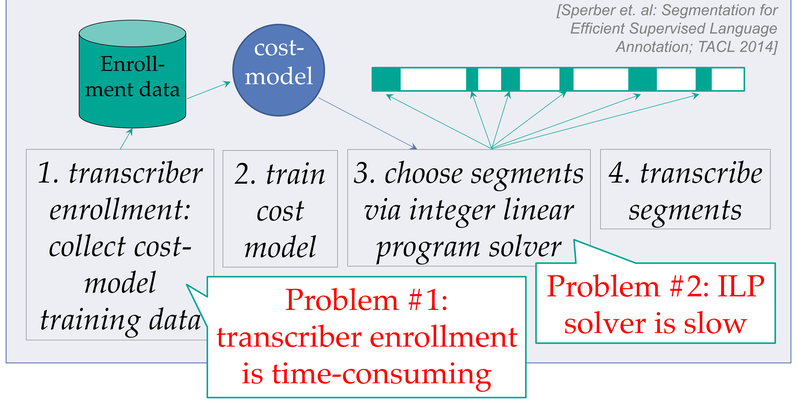
Matthias Sperber[1], Graham Neubig[2], Satoshi Nakamura[2], Alex Waibel[1]

[1] Karlsruhe Institute of Technology, Germany; [2] Nara Institute of Science and Technology, Japan

What to do when automatic speech recognition produces too many errors, and manual transcription is too expensive?

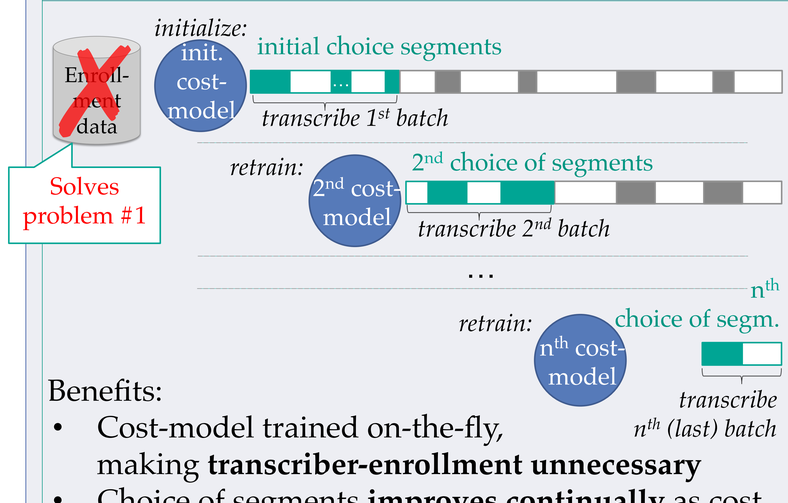


Baseline



Contributions

1. On-the-fly updating of cost model & segmentation



Benefits:

- Cost-model trained on-the-fly, making **transcriber-enrollment unnecessary**
- Choice of segments **improves continually** as cost model improves
- **Recover from time-prediction errors** → e.g. remove segments when running out of time

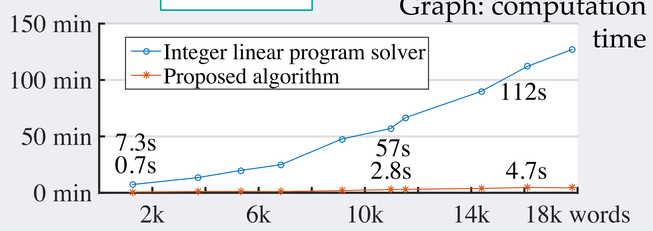
On-the-fly Cost Modeling:

- Gaussian Process regression: supports a prior!
- Use crude prior: cost ≈ 2sec + 1sec per word.
- Unrealistic, but captures 2 key properties:
  1. Longer segments need more time.
  2. Cognitive overhead for switching segments.
- After each update, predictions get more accurate.

2. Fast segmentation algorithm

Solves problem #2

- Important: don't want the transcriber to wait while updating.
- Optimal choice of segments is NP-hard! (→ competing goals of min. cost vs. max. #corrections)
- Optimize combined objective: λ#*corrections* - *(1-*λ*)cost* via dynamic programming.
- Binary search to find λ that comes closest to time budget.

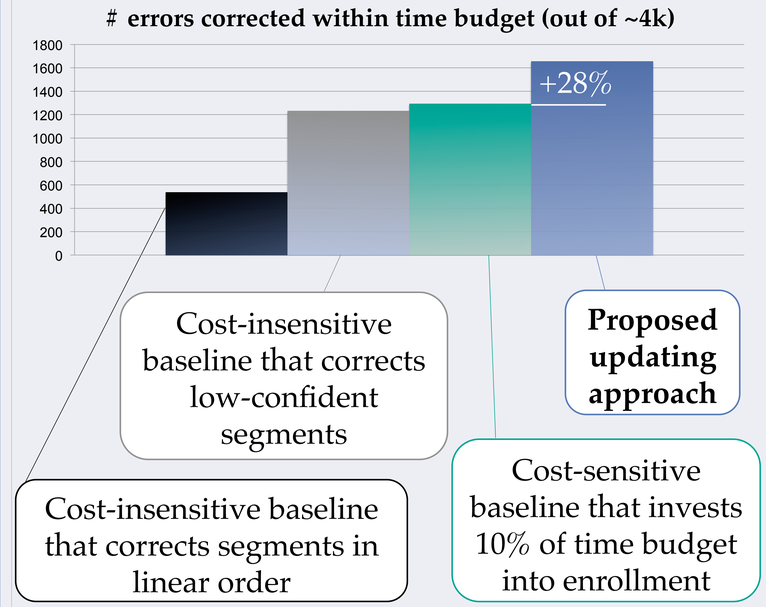


Experiments

Setup

- Task: Correct automatic transcripts of 10 TED talks (104 minutes of audio).
- Available time budget: 100 minutes.
- Simulating a realistic transcriber:
  - Train a cost model on correction times observed from a real transcriber.
  - Apply noise to output times.
  - Assume that transcriber behaves like this oracle cost-model (and perfectly corrects all errors).

Simulation Results



Wrap-Up

- Updating approach improves correction efficiency by 28% over prior cost-sensitive approach.
- Get rid of time-consuming transcriber enrollment.
- Fast updates remove transcriber waiting time.

KIT – University of the State of Baden-Wuerttemberg and National Research Center of the Helmholtz Association

Nara Institute of Science and Technology NAIST

Karlsruhe Institute of Technology