D6: StenoPhone

Cambrea Earley, Ellen Seeser, and Mitchell Yang

Application Area

• Application area: distributed team meetings

• Previous solution: third party transcription plugged into meeting software, transcription capabilities without speaker identification

• New solution: transcription with speaker identification integrated in meeting hardware and software

Solution Approach



Complete Solution



| < | → App | s 2 | 2 (| A Calen | No dar | t sec | ure C | ec2 | -18-2 | 221· | -2-198 Can | 8 .us - vas | eas P | t-2.co Piaz | omp za | ute. | ama ECE | zona Capst | ws.cc | om:80 | 100/n | neet Gra | ings/ | 57/ | • | 18-335 / | 18-732 | 3 | 18330 | 6 | 15-4 |
|------|----------|------|-------|-------------------|-----------|-------|----------|-------|-------|-------|---------------|-----------------------|----------|----------------|-----------|------|------------|---------------|-------|-------|-------|-------------|-------|-----|---|----------|--------|---|-------|---|------|
| M | 6 | >ti | n | a | 5 | 7 | | | | | | | | | | | | | | 50 | | | | | | | | | | | |
| Le | eave | Mee | eting | 9 | | Enc | l Me | eting | | | Add N | licro | phoi | ne | | D | ownl | oad | | | | | | | | | | | | | |
| Tra | an | scr | ip | t: | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| [p2] | 0 | kay | 50 | now | we | have | a | lot | of da | ata | and | I | | | | | | | | | | | | | | • | | | | | |
| [p2] | I | nee | ed t | o f | ix t | the s | peal | ker | okay | why | don | 't y | ou | expl | ain | it | | | | | | | | | | | | | | | |
| [p2] | 1 | ike | it | 5 S | uppo | sed | to | | | | | | | | | | | | | | | | | | | | | | | | |
| [p2] | f | or e | ever | y t | ime | the | aud | io c | oming | g tł | nroug | h bu | t i | t se | ems | to | be | like | only | у | | | | | | | | | | | |
| [p2] | G | oogl | e y | eah | Ιt | hink | \$ 50 | too | | | | | | | | | | | | | | | | | | | | | | | |
| | | | 21 | | | | | | a | 12.10 | 112 | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Setup Demo



Meeting Demo





| Requirement | Metric | Test | Results and Tradeoffs |
|----------------------------------|--|---|---|
| Audio Transmission Latency | Mouth-to-Ear Latency (ms) < 150 ms | Route audio packets through server, to and from same microphone to find RTT | Processing vs inter packet delay <u>Latency <= 30 ms</u> |
| Audio Quality | Dropped packets (%) < 5% | Count original and final number of packets after transmitting an audio stream | Dropped packets <= .3% |



| Requirement | Metric | Test | Results and Tradeoffs |
|------------------------|--------------------------------|--|---|
| Transcript Latency | Average Word Delay (s) < 3s | Capture timestamp of audio captured by mic and timestamp of packet arrival in browser. 80 samples taken. | <u>Average word delay <= 1.8s</u> |
| Transcript Accuracy | Word Error Rate (%) < 25% | Check transcript for word error (substitution, deletion, and insertion) after speaking a known text.* 543 samples taken from 4 trials. | Paid services vs open source packages <u>Average WER: 18%</u> |

*100+ words. Well-formed sentences featuring common English words. Constant over multiple tests.

| Speaker Identification Accuracy | Speaker Identification Error (%) < 25% | Check transcript for identification error after conducting a conversation with known contents and speaker switches. 437 samples taken from 3 trials. | Paid services vs open source packages DOA assumptions vs relying on ML Speaker ID Error 12% |
|--|---|--|---|
| Formatting Accuracy (chronology and speaker ID tags) | Formatting Error Rate (%) < 5% | Check transcript for formatting error instances after conducting a conversation with known contents and microphone switches. 400 samples taken from 5 trials. | Formatting error = 0% |

StenoPhone Result

oh did you manage to get a response from her [P1] ves did she said she would [sp 3] a delivery willing to make other [sp 3] want [sp 3] in the City to bring us the parts we were great did she indicate a price for the delivery [P2] she is going to charge 50 dollars [sp 3] with all right I'll speak the department to ask for \$50 for the delivery fee [P2]

Sample Script (P1, P2 in one room; sp 3 in another room)

[P1] Oh, did you manage to get a response from her?

[sp 3] Yes, I did. She said she would be willing to make a delivery in the city to bring us the parts we want.

[P2] That's great. Did she indicate a price for the delivery?

[sp 3] She is going to charge fifty dollars.

[P2] All right. I'll speak with the department to ask for fifty dollars for the delivery fee.

Project Management

| | | | | Week 3(Feb 15-21) | | v | Week 4(Feb 22-28) | | | ek 5(Marc | x 5(March 1-7) Week 6 | | | ch 8-14) | Week 7(March 15-21) | | | Wee | ek 8(March | 22-28) | Week 9 | 9(March 2 | 9-April 4) | Wee | ek 10(Apr | 15-11) | Wee | Week 11(April 12-18) | | | ek 12(Apr | l 19-25) | Week | 13(April 2 | .6-May 2) | We | ek 14(May | 3-5) |
|---|-------------|----------|-------------------------|-------------------|-----|-------|-------------------|----------|----------|-----------|-----------------------|-----------|--------|-----------|---------------------|----------|----------|----------|------------|---------|-----------|-----------|------------|-----|-----------|--------|-----|----------------------|-----|-----|-----------|----------|------|------------|-----------|-----|-----------|-----------|
| Task | Team Member | Status | м | w | F | м | w | F | м | W | F | м | w | F | м | W | F | м | w | F | M | w | F | м | W | F | м | w | F | м | w | F | м | w | F | м | w | |
| Phase 0: Design | all | done | - | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Initial research | all | done | all | all | all | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Project Proposal Presentation | all | done | all | all | all | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Phase 1: System Setup | all | done | - | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Network Design: RPi to AWS | cne | done | * | | | cne | cne | cne | 1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Network Design: AWS to rpi | cne | done | - | | | | | cne | cne | cne | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Hardware Setup | cne | done | • | | | | | cne | cne | cne | cne | _ | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Respeaker Config on Raspberry Pi | cne | done | * | | | | | | | cne | cne | cne | 1 | | | | | | | | | | | | | | | | | | | | | | | | | |
| AWS Database + Django Initialization | mfy | done | - | | | mfy | mfy | mfy | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| AWS EC2 Setup | mfy | done | * | | | mfy | mfy | mfy | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Speech to Text ML Initial Version | eseeser | done | - | | | esees | ser eseese | r eseese | r | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Design Presentation & Report (presentation due 7th, report 17th) | all | done | - | | | all | all | all | all | all | all | all | all | all | | | | | | | | | | | | | | | | | | | | | | | | |
| Backend Pre-ML Processing | eseeser | done | - | | | | | eseese | r eseese | r eseese | r | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Test & Slack Time | all | done | * | | | | | | _ | | | all | | | | | _ | | | | | _ | | _ | _ | | | | | _ | | | _ | | | _ | | |
| Phase 2: Backend | all | done | - | | | | | - | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Database Creation and Integration | mfy | done | * | | | | | mfy | mfy | mfy | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Website Creation, Including Backend | mfy | done | - | | | | | mfy | | | mfy | mfy | | | mfy | mfy | mfy | | | | | | | | | | | | | | | | | | | | | |
| Non-Moving Speaker ID (no ML) | eseeser | done | • | | | | | | | eseese | r eseese | r eseeser | eseese | ər | | | | | | | | | | | | | | | | | | | | | | | | |
| Server connection at AWS | cne | done | * | | | | | | | | | cne | cne | cne | | | | | | | | | | | | | | | | | | | | | | | | |
| Microphone IO on RaspberrryPi | cne | done | • | | | | | | | | | | cne | cne | cne | cne | cne | | | | | | | | | | | | | | | | | | | | | |
| Client Code on Raspberry PI | cne | done | * | | | | | | | | | | | | | cne | cne | cne | | | | | | | | | | | | | | | | | | | | |
| Speech to Text ML Integration (first pass at meeting manager) | eseeser | done | - | | | | | | | | | _ | eseese | er eseese | r eseesei | r | | | | | | | | | | | | | | | | | | | | | | |
| Audacity Filter Design + Audio Process | mfy | done | - | | | | | | | | mfy | mfy | mfy | mfy | | | | | | | | _ | | | | | | | | | | | | | | | | |
| 2 Audio Device Audio Streaming | cne | done | - | | | | | | | | | | | | | | | - | cne | cne | cne | cne | | | | | | | | | | | | | | | | |
| Mic Initialization Backend | eseeser | done | - | | | | | | | | | | | | | | | eseese | er eseesei | eseeser | r | | | | | | | | | | | | | | | | | |
| Testing & Slack Time | all | done | * | | | | | | | | | | | | | | | all | all | all | | | | | _ | | | | | | | | | | | | | |
| Phase 3: Frontend | all | done | * | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Create housing | mfy | done | * | | | | | | | | | | | | | | | mfy | mfy | | | | | | | | | | | | | | | | | | | |
| Transcript Support for Multi-mic Meetin | eseeser | done | - | | | | | | | | | | | | eseese | r eseese | r eseese | r | | | | | | | | | | | | | | | | | | | | |
| Full On-Website Meeting Setup | eseeser | done | - | | | | | | | | | | | | | eseese | r eseese | r eseese | er: | | | | | | | | | | | | | | | | | | | |
| Moving Speaker ID (ML) | eseeser | done | * | | | | | | | | | | | | | | | | eseese | eseese | r eseeser | r eseeser | eseeser | r | | - | | | | | | | | | | | | · · · · · |
| Transcript Web Streaming | mfy | done | - | | | | | | | | | | | | | | | | | mfy | mfy | mfy | mfy | mfy | mfy | | | | | | | | | | | | | |
| Integration of RPi + Transcript + ML | all | done | * | | | | | | | | | | | | | | | | | | all | all | all | all | all | all | | | | | | | | | | | | |
| Complete System Debugging / Testing | all | done | - | | | | | | | | | | | | | | | | | | | | | | all | all | all | all | | | | | | | | | | |
| Phase 4: Touchup | all | progress | * | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Web CSS | all | done | - | | | | | | | | | | | | | | | | | | | | | | | all | all | all | all | | | | | | | | | |
| AWS Migration | all | done | - | | | | | | | | | | | | | | | | | | | | | | | | | | all | all | | | | | | | | 1 |
| Testing | all | done | - | | | | | | | | | | | | | | | | | | | | | | | | | | | all | all | all | all | all | all | | | |
| Slack Time | all | progress | * | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | all | all | | |
| Final Report, Presentation, Final Video | all | progress | - | | | | | | | | | | | | | | | | | | | | | | | | | all | all | all | all | all | all | all | all | all | | |

Takeaways / Conclusion

• Tradeoffs of python development

• Challenges of audio streaming

• Software as a service vs. open source software packages

• Rewards of project design

