

Санкт-Петербургский государственный электротехнический университет им.
В.И. Ульянова (Ленина)

Разработка сервиса электронного протоколирования совещаний и защит

Выполнил:	Петров Сергей Андреевич, гр. 7303
Руководитель:	Шевская Наталья Владимировна, ассистент каф. МОЭВМ

Санкт-Петербург, 2022

Актуальность и план

Актуальность: электронные созвоны и совещания часто необходимо протоколировать для извлечения ключевых моментов/вопросов, разработанное решение позволит сотрудникам вузов или компаний решить проблему и автоматизировать процесс протоколирования рабочих совещаний (например, очных защит дипломов студентов и аспирантов).

План на текущий семестр:

1. Устранение сильной зависимости качества распознавания от качества звука (микрофона)
2. Отображение плохо распознаваемых слов в клиентской части решения
3. Распознавание английских терминов в русской речи
4. Устранение зависимости качества распознавания от скорости и качества речи

Очистка входного звука с микрофона от шума

<https://github.com/timsainb/noisereduce>

Sainburg T., Thielk M., Gentner T. Q. Finding, visualizing, and quantifying latent structure across diverse animal vocal repertoires //PLoS computational biology. – 2020. – Т. 16. – №. 10. – С. e1008228.

Эксперимент/Метрика	spacy	jaccard_raw	jaccard_process	euclidean	cosine
result_only_medfilt	0.99116094	0.69473684	0.76190476	0.71636781	0.99143697
result_no_medfilt_no_sample_nfft4096_timemask64	0.99173723	0.75268817	0.79761904	0.72951242	0.99186392
result_no_medfilt_with_sample_nfft4096_timemask64	0.99147399	0.72631578	0.78823529	0.72528363	0.99156084
result_no_medfilt_no_sample_default	0.99127441	0.70833333	0.77380952	0.71748315	0.99123657
result_no_medfilt_with_sample_default	0.99109030	0.70833333	0.77380952	0.71499794	0.9912101546229997
result_raw	0.99173723	0.75268817	0.79761904	0.72951242	0.99186392

Результат: подбор оптимальных параметров для очистки шума

Улучшение отображения слов на клиентской части решения

Использование параметра `conf` для выделения слова на клиенте:

```
{  
  "conf": 1,  
  "end": 12.81,  
  "start": 11.73,  
  "word": "протоколирования"  
}
```

Результат:

(24.51) sergei_test: **как** же они уже отображает

(27.84) sergei_test: выделенные красным слова

(32.37) sergei_test: что означает

(35.19) sergei_test: что **приложения** не **уверена** в данных моментах

План на следующий семестр/ссылки

План дальнейшей работы:

- исследование подходов для улучшения распознавания вопросов спикеров, расширение/улучшение набора правил для распознавания вопросов;
- добавление возможности для редактирования получившейся транскрипции/расшифровки совещания/конференции, предложение альтернатив для распознанных слов с низким коэффициентом распознавания;
- оптимизация разработанного решения, в частности взаимодействия между основным сервером и сервером распознавания речи;

Ссылки:

Репозиторий - <https://github.com/moevm/zoom-transcriber>

Результаты экспериментов по очистке шума - [google drive](#)