

# Разработка алгоритма диаризации

Завадский Илья  
Программная Инженерия,  
Математико-Механический факультет,  
СПбГУ, 2020

Научный руководитель: к. ф.-м. н., ст. Преп. Луцив Д.В.  
Консультант: рук. проекта, ООО "ЦРТ" Тимченко М.С.



# Введение

- **Диаризация дикторов** (или разделение дикторов) — это процесс разделения входного аудиопотока на сегменты в соответствии с идентификацией личности диктора.
- **Детектор речевой активности** (англ. Voice Activity Detection, VAD) — это алгоритм, классифицирующий исходные участки фонограммы как речь или не речь.



## Цель

- Разработать, реализовать, протестировать и внедрить алгоритм диаризации дикторов, позволяющий уменьшить накладные расходы на вычислительные ресурсы и память



# Задачи

- Рассмотреть существующие решения;
- Изучить формат хранения данных;
- Разработать, реализовать и внедрить алгоритм diarизации дикторов;
- Протестировать на реальных данных;



# Технологии

1. RedHat 7
2. WinSCP + PuTTY + Postman
3. TeamCity
4. Mercurial Hg
5. Audacity
6. Java 1.8

# Реализация

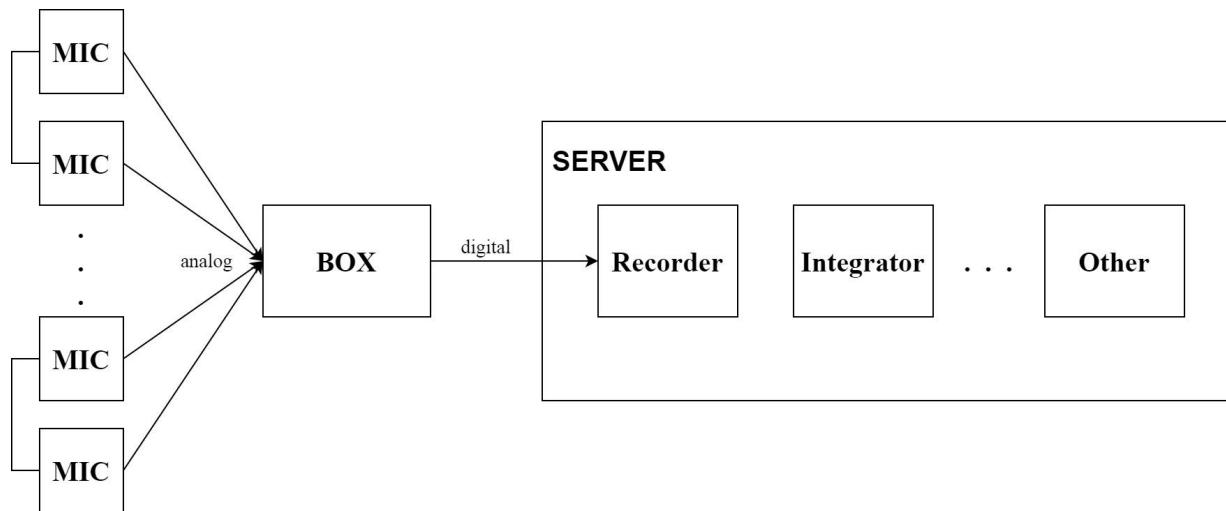


Рис. 1: Схема обработки аудиосигнала

# Реализация: VAD

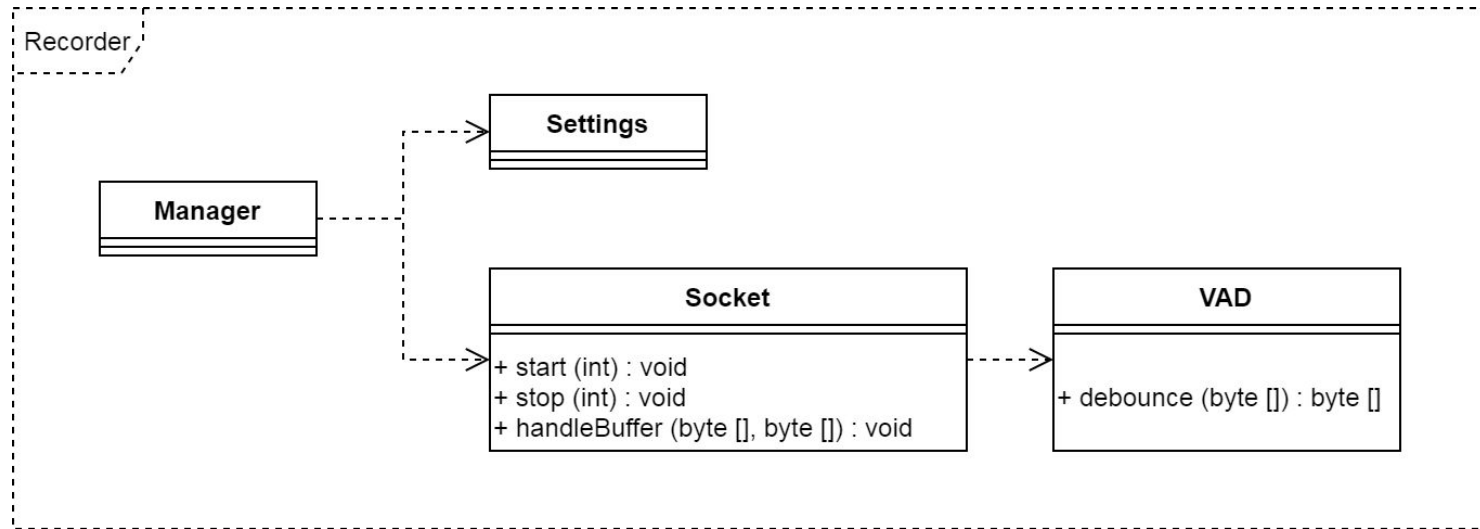


Рис. 2: Диаграмма Recorder

# Реализация: VAD

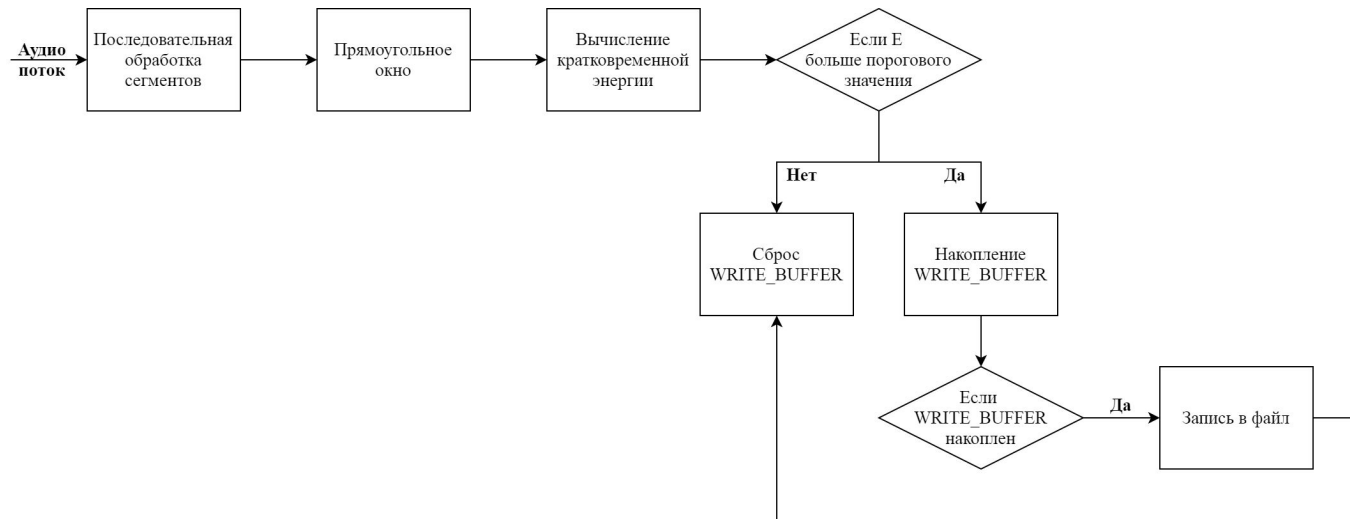


Рис. 3: Блок-схема работы VAD-алгоритма



# Реализация: Diarizer

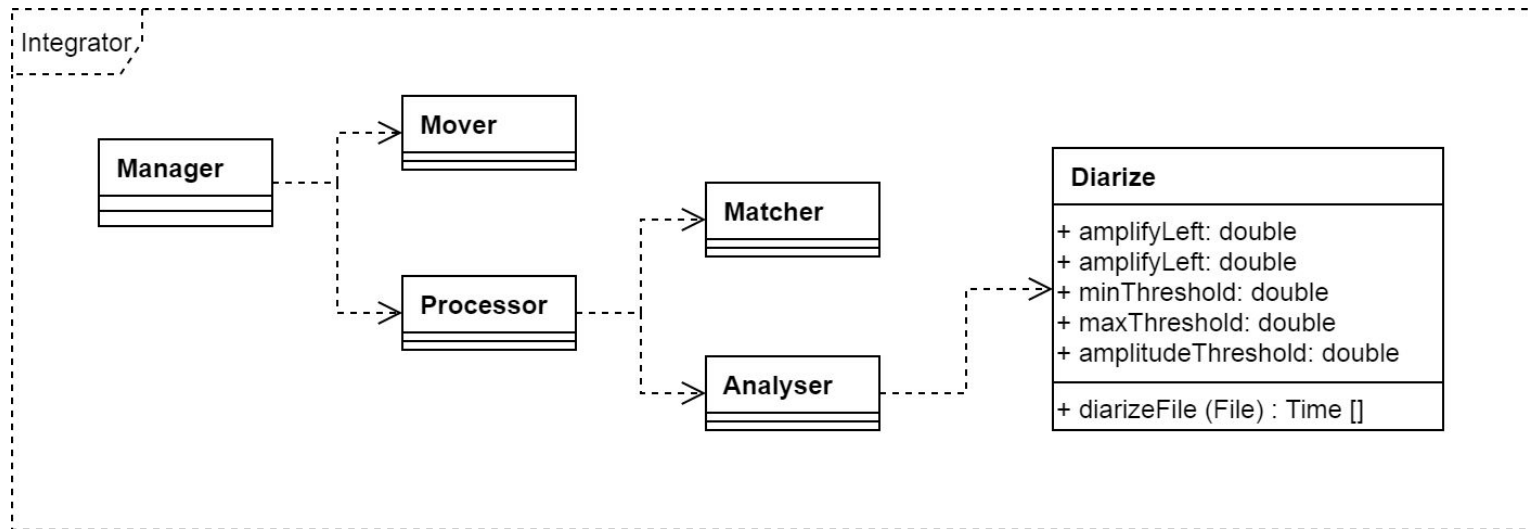


Рис. 4: Диаграмма Integrator

# Реализация: Diarizer



Рис. 5: Блок-схема работы алгоритма диаризации



# Тестирование

Память, которая требуется для хранения аудиозаписи/аудиозаписей за весь день	2.57 ГБ / 1.44 ГБ
Отношение времени диаризации аудиофайла до и после внедрения детектора речевой активности	1
Отношение времени полной обработки аудиофайла до и после внедрения детектора речевой активности и алгоритма диаризации	0.80
Отношение времени полной работы сервиса Integrator до и после внедрения детектора речевой активности и алгоритма диаризации	0.80

Таблица 1: Результаты тестирования



# Результаты

**В рамках работы были выполнены следующие задачи:**

- Рассмотреть существующие решения;
- Изучить формат хранения данных;
- Разработать, реализовать и внедрить алгоритм диаризации дикторов;
- Протестировать на реальных данных;

**Код алгоритма и примеры:** [https://github.com/ilyazavadskij/Course\\_Work](https://github.com/ilyazavadskij/Course_Work)