

# МЕТОДЫ АНАЛИЗА РЕЧЕВОГО ПОТОКА

Жожикашвили В.А.

Петухова Н.В.

Фархадов М.П.

*Институт проблем управления им. В.А. Трапезникова РАН, г. Москва*

*[serena@ipu.ru](mailto:serena@ipu.ru), [mais@ipu.ru](mailto:mais@ipu.ru)*

Ключевые слова: автоматическое распознавание речи, анализ речи в реальном времени, пакетная обработка больших потоков речи, поиск ключевых слов в потоке речи.

## **Введение**

Успехи в развитии технологии автоматического распознавания речи открыли новые области ее применения, среди которых одно из ведущих мест занимает использование этой технологии для анализа речевых потоков. В таких приложениях распознаватель используется для прослушивания речи и реагирования на заранее определенные ключевые слова.

### **1. Основные подходы к анализу потока речи**

Имеется три основных подхода к реализации функции поиска ключевых слов.

Первый подход состоит в том, что анализируемый аудио поток вначале преобразуется в слова путем использования распознавателей с большим объемом словаря и способностью распознавания непрерывной речи (LVCSR). На второй фазе выполняется поиск ключевых слов в полученном тексте.

Второй подход также имеет две фазы и состоит в том, что вначале речь преобразуется в последовательность фонем, а на второй фазе производится поиск фонетически транскрибированных ключевых слов в этом сформированном потоке фонем.

Третий подход реализуется в одну фазу и состоит в том, что распознаватель, работающий на основе фонемных транскрипций, ищет в речевом потоке последовательности фонем, описывающих ключевые слова.

### **2. Сравнение методов анализа потока речи**

Преимущества первого подхода:

- после выполнения преобразования речи в текст объем материала значительно сокращается, что важно для его хранения,
- не требуется предварительного описания ключевых слов,
- легко выполняется повторный поиск по тем же или другим ключевым словам, поскольку текстовая база уже создана.

Недостатки этого метода:

- как бы не был велик объем словаря, он все равно не может охватить все слова, которые может потребоваться искать, например иностранные имена, термины, сокращения. Добавление нового слова в словарь распознавателя – сложная процедура,
- распознаватели с большим объемом словаря являются чрезвычайно ресурсоемкими, поэтому их нельзя использовать в системах, где требуется малое время реакции.

Преимущества второго подхода те же самые, что и первого, т.е. список ключевых слов не должен быть определен заранее, до выполнения первой фазы, а также время поиска, выполняемого на второй фазе, значительно меньше, чем время поиска в речевом потоке

Недостатки второго подхода:

- первая фаза генерирует «зашумленный» результат, уровень ошибок может быть значительным, и это негативно сказывается на результатах поиска ключевых слов на второй фазе,
- поиск ключевых слов в их фонемном представлении является, как правило, более ресурсоемким, чем поиск слов в тексте при первом подходе,
- наличие двух фаз обработки требует значительных затрат процессорных ресурсов и это ограничивает его применение в реальном времени или в загруженных системах.

Преимущества третьего подхода: