

УДК 004.934

**УЛУЧШЕНИЕ АЛГОРИТМА ЛУЧЕВОГО ПОИСКА ГИПОТЕЗ ДЛЯ РЕЗУЛЬТАТОВ
РАСПОЗНАВАНИЯ РЕКУРРЕНТНОГО НЕЙРОСЕТЕВОГО ПРЕОБРАЗОВАТЕЛЯ.КУ**

Андрусенко А.Ю.

(Университет ИТМО)

Научный руководитель – к.т.н. Меденников И. П.

(Университет ИТМО)

Рекуррентный нейросетевой преобразователь – интегральный метод обучения системы автоматического распознавания речи. В данной работе рассматривается алгоритм лучевого поиска гипотез распознавания (декодирование) для данной системы и его улучшения с точки зрения скорости работы и качества пословной ошибки распознавания.

Наряду с классическими методами обучения систем автоматического распознавания речи были предложены более прямые подходы - так называемые интегральные методы, в которых нейронные архитектуры обучаются непосредственно моделировать последовательности признаков как символы алфавита. Прогнозирование независимых от контекста целей, таких как символы алфавита, с использованием единой архитектуры нейронной сети, вызвало большой интерес со стороны исследовательского сообщества, а также разработчиков, не являющихся экспертами. Это вызвано простотой создания полной системы распознавания речи без необходимости экспертных знаний.

Одним из популярных видов подобных интегральных методов является рекуррентный нейросетевой преобразователь (RNN-T). Его основным преимуществом является сочетание в одной системе акустической и языковой модели, обучаемых совместно на основании пословной ошибки распознавания речи. Данный подход уже успел себя зарекомендовать во многих задачах автоматического распознавания речи. Однако алгоритм декодирования (лучевой поиск гипотез, beam search), доступный в открытых исходных источниках кода, является далеким от совершенства, что значительно увеличивает время, затрачиваемое на получение результата распознавания подобных моделей. Для решения подобных проблем в классических системах распознавания речи существуют техники усечения луча поиска гипотез, которые способствуют значительному ускорению работы алгоритма декодирования без существенных потерь в качестве пословной ошибки распознавания речи.

В данной работе рассмотрены основные проблемы разрастания числа гипотез лучевого поиска и произведен замер времени работы данного алгоритма. Далее реализованы техники усечения луча поиска: ограничение ширины динамического диапазона накопленного веса гипотезы и ограничение на величину различия между временным и текущим лучом поиска.