

Н. М. Ч а в ч а в а д з е (Москва, ТВП). **Способ реализации технологии и распознавания «дикторов».**

В последние годы отмечается существенный рост интереса к автоматическим системам распознавания говорящего по голосу [6]. Потребность в разработке таких систем, главным образом, диктуется наличием широкого круга практических приложений, где требуется подтвердить или опознать определенную личность.

Область применения систем, определяющих личность говорящего по голосу, можно разделить на два обширных блока. Первый блок составляют задачи контроля прав доступа к различным системам, таких как контроль за физическим доступом в помещения, доступ к базам данных и вычислительным средствам, доступ к банковским счетам.

Второй блок связан с использованием систем анализа голоса говорящего при проведении криминалистических экспертиз и оперативно-розыскной деятельности. Здесь необходим анализ большого числа записей телефонных переговоров на предмет отождествления их с подозреваемыми.

В первом случае задача сравнения параметрического кода предъявляемого голоса с эталонным высказыванием заявленного лица, т. е. верификации «диктора», заключается в проверке простой гипотезы при простой альтернативе.

Во втором случае стоит более сложная задача выбора из большого количества телефонных переговоров конечного числа зарегистрированных лиц. Рассмотрим способ реализации технологии распознавания «дикторов» для этого случая более подробно.

Здесь предлагается осуществлять следующие действия с последовательным уменьшением потока обрабатываемых сообщений.

1. Удаление шумов и речевой звуковой информации в звуковых файлах. На данном этапе отбраковываются файлы с недостаточным количеством речи, в результате чего входящий поток уменьшается на 20–30%.

2. Идентификация «диктора» и определение его пола [1, 2, 3, 5]. Поток разбивается на подмножества файлов, привязанных к каждому из дикторов. Каждому файлу ставится в соответствие весовой коэффициент, характеризующий вероятность принадлежности файла «диктору». Для предотвращения пропуска истинных файлов необходимо найти оптимальное соотношение ошибок первого и второго рода на основании производительности системы, при котором для каждого «диктора» отбирается 10–20% от поступившего на идентификацию потока. Значение параметра «пол диктора» для отобранных файлов сохраняется и используется в качестве вспомогательного критерия на этапе анализа.

3. Верификация «диктора» [1–3]. Файлы из каждого подмножества верифицируются на основе карточки «диктора», к которому они привязались на предыдущем этапе. Вероятность принадлежности файла «диктору» характеризуется некоторым значением меры различимости. Программа верификации отбирает 5–10% файлов.

4. Определение языка речевого сообщения [1–4]. Полученное значение параметра «язык диктора» используется в качестве вспомогательного критерия в процессе анализа на следующем этапе.

5. Анализ полученных результатов. После идентификации и верификации «дикторов» для каждого интересующего абонента определяется набор звуковых файлов, которые требуется прослушать для принятия решения о принадлежности к абоненту. В качестве вспомогательных критериев отсева ложных вариантов, как отмечалось выше, используются значения параметров «пол диктора» и «язык диктора», определенные на 2-м и 4-м этапах. Дополнительными признаками отбора являются совпадение таких идентификационных признаков, как номер телефона, для нескольких файлов, а также объем речевого файла.

Таким образом, предлагается определять набор звуковых разговоров предполагаемых абонентов и принадлежащим им идентификационным признакам. Для того чтобы не пропустить истинные звуковые файлы, отбракованные на 1-м этапе и из-за

ошибок первого рода на 2–4-м этапах, рекомендуется повторно обработать исходный поток на предмет поиска их по найденной идентификационной информации.

СПИСОК ЛИТЕРАТУРЫ

1. *Benesty J., Sondhi M. M., Huang Y.* Springer Handbook of Speech processing. Berlin–Heidelberg: Springer-Verlag, 2008.
2. *Пикон Д.* Моделирование сигнала. / Под ред. Р. Попова. Кемерово, 2000.
3. *Рабинер Р. Л., Шафер Р. В.* Цифровая обработка речевых сигналов. / Под ред. М. В. Назарова и Ю. Н. Прохорова. М.: Радио и связь, 1981.
4. *Максимов А. В., Мельников С. Ю., Чавчавадзе Н. М.* Тенденции развития методов автоматической идентификации языка речевых и текстовых сообщений. — Обозрение прикл. и промышл. матем., 2009, т. 16, в. 2, с. 365–367.
5. *Котов М. А., Максимов А. В., Мельников С. Ю., Чавчавадзе Н. М., Федюкин М. В.* Способ реализации алгоритма идентификации дикторов с использованием ПЛИС. — Обозрение прикл. и промышл. матем., 2009, т. 16, в. 6, с. 1091–1092.
6. *Максимов А. В., Мельников С. Ю., Чавчавадзе Н. М., Федюкин М. В.* Развитие систем автоматической текстонезависимой идентификации дикторов. — Обозрение прикл. и промышл. матем., 2009, т. 16, в. 6, с. 1092–1094.