

# СИСТЕМЫ ВЕРИФИКАЦИИ СУБЪЕКТОВ ДОСТУПА НА ОСНОВЕ РЕЧЕВОГО СИГНАЛА

Крищеневич В.А., Захарьев В.А.

Кафедра защиты информации, Кафедра систем управления,  
Белорусский государственный университет информатики и радиоэлектроники  
Минск, Республика Беларусь  
E-mail: {krish, zahariev}@bsuir.by

*Доклад посвящен анализу предметной области связанной с построением систем верификации пользователей по биометрическим признакам (в частности, на основе характерных особенностей голоса) для управления доступом к персональной информации. Классификации подобных систем по типу используемой информации, анализу основных методов и алгоритмов, применяемых при построении современных систем данного типа.*

## ВВЕДЕНИЕ

В настоящее время, в связи с широким распространением персональной вычислительной техники в различных областях жизни, особо остро встает вопрос ограничения доступа пользователя к информации, хранящейся на данных устройствах. Для этих целей, широко применяются различные методы идентификации и верификации субъектов доступа. При этом все чаще стандартные парольные системы защиты, персональные идентификаторы и удостоверения личности заменяются или дополняются биометрическими системами идентификации и верификации пользователей. Это связано с тем, что наличие стандартных механизмов, зачастую является не достаточным для обеспечения необходимого уровня надежности защиты информации. Поэтому при создании современных систем контроля и управления доступом, применяются методы биометрического распознавания, основанные на вычислении и анализе индивидуальных характеристик субъекта доступа. На текущий момент в качестве измеряемых параметров используют различные человеческие черты, такие как отпечатки пальцев, изображения лица, записи речи диктора и т.д. Примерами таких систем могут выступать хорошо известные и зарекомендовавшие себя технологии и системы применяемые в персональных устройствах, такие как TouchID, FaceID и т.д.

### I. ВИДЫ ИНФОРМАЦИИ ПРИМЕНЯЕМЫЕ ДЛЯ ВЕРИФИКАЦИИ

В процессе распознавания речи могут использоваться лингвистические (непосредственно связанные с конкретным языком) и паралингвистические (не связанные с конкретным языком) особенности речи говорящего. К лингвистическим особенностям, относятся параметры характеризующие само сообщение, их определение позволяет ответить на следующие вопросы:

- О чем говорят?
- На каком языке говорят?

Паралингвистические особенности, отражают индивидуальные особенности и эмоционально-физиологические характеристики речи говорящего и в большой степени характеризуют источник речевого сообщения, т.е. диктора. Их обработка и анализ в системе позволяет ответить на следующие основные вопросы:

- Кто говорит?
- Пол говорящего?
- Возраст говорящего?
- Эмоциональное состояние говорящего?

Для верификации диктора используются признаки и информация второго вида, т.е. относящаяся к паралингвистической.

### II. КЛАССИФИКАЦИЯ СИСТЕМ

Для понимания особенностей функционирования необходимо дать классификацию основных типов систем, которые относятся к системам распознавания диктора, поскольку системы верификации диктора являются одним из подклассов подобных систем. Таким образом системы распознавания диктора, в зависимости от основных процессов происходящих в них, можно разделить на системы:

- Идентификации: определение личности неизвестного среди группы записанных голосов. Звукозапись неизвестного сравнивается с моделями всех зарегистрированных голосов и определяются вероятности их соответствия.
- Верификации: подтверждение заявленной личности человека с помощью речевого сигнала. Сравнивается предоставленный образец записи голоса с имеющимися образцами и на основании степени соответствия принимается решение принять или отклонить возможность выполнения дальнейших действий в системе.
- Диаризации: разделение входного аудиопотока на сегменты в соответствии с личностью диктора, чей голос звучит в данный момент времени.

### III. АЛГОРИТМ ВЕРИФИКАЦИИ ДИКТОРА

Верификация диктора направлена на определение того, соответствует ли личность диктора заявленной, и обычно требует одного сравнения. С другой стороны, идентификация дикторов среди N-размерной группы требует N число сравнений.



Рис. 1 – Обобщенный алгоритм верификации диктора

Основные этапы работы системы в процессе выполнения алгоритма, включают в себя следующие основные шаги:

- анализ и извлечение из речевого сигнала признаков характеризующих голос диктора (модуль обработки сигнала);
- классификация полученных признаков с помощью моделью диктора, хранящейся в системе (модуль обработки признаков);
- принятие решения об успешности верификации на основании результатов сравнения.

Для выполнения процесса классификации требуется следующих действий:

- Разработка универсальной модели: проанализировать сходства в параметрах голосов дикторов. Сама модель именуется универсальной фоновой моделью (УФМ), поскольку является инвариантной по отношению к конкретным фонемам языка. Она может быть определена с учетом пола, т.е. есть отдельная модель для мужских и женских голосов, но является общей для всех дикторов.
- Разработка модели диктора: проанализировать различия в параметрах голосов диктора. На этом этапе обычно создается одна модель для каждого рассматриваемого уникального диктора.
- Верификация: голос субъекта сравнивается с ранее зарегистрированными голосами, согласно алгоритму представленному выше.

Основные проблемы, возникающие при построении систем верификации:

- большое количество данных, необходимых для обучения моделей;
- несоответствие между данными на этапе обучения и этапе верификация (функционирования систем), поскольку могут измениться параметры акустического окружения (посторонние шумы).

Разница между вводными данными и проверочными данными может быть связана с:

- аппаратными средствами (микрофон);
- окружающей средой (шумы и помехи);
- каналом передачи (стационарный, VoIP);
- самим диктором (состояние здоровья).

Процесс принятия решения в системах верификации диктора осуществляется на основе определения вероятности того, что голос диктора соответствует уникальной модели диктора, и вероятности того, что характеристики голоса диктора относятся к универсальной фоновой модели.



Рис. 2 – Применение универсальной фоновой модели для верификации диктора

Далее определяется отношение данных вероятностей. Если соотношение выше порога принятия решений, то система принимает решение, что образец принадлежит конкретному диктору. В противном случае принимается, считается что диктор не определен. Стоит обратить внимание, что УФМ решение выносится с учетом особенностей голоса характерных для дикторов конкретного пола.

### ЗАКЛЮЧЕНИЕ

В докладе рассмотрены вопросы, связанные с системами верификации диктора на основе речевого сигнала: какие виды информации применяются для верификации, классификация подобных систем, обобщенный алгоритм функционирования систем, применение унифицированной фоновой модели для осуществления процесса верификации.

Технология биометрической верификации по голосу представляет особый интерес при её использовании в составе мобильных устройств, поскольку позволяет проводить процесс верификации без необходимости прямого тактильного или визуального взаимодействия с устройством, что в значительной степени улучшает показатели удобства использования подобных систем.

### СПИСОК ЛИТЕРАТУРЫ

1. Mak M.W., Jen-Tzung C. Machine Learning for Speaker Recognition. Cambridge University Press, 2020 – 336 p.
2. Fabien M., Basics of Speaker Verification. [Электронный ресурс]. Режим доступа: [https://maelfabien.github.io/machinelearning/basics\\_speech](https://maelfabien.github.io/machinelearning/basics_speech). Дата доступа: 15.10.2021.