

## ИНДЕНТИФИКАЦИЯ ПО ГОЛОСУ. СКРЫТЫЕ ВОЗМОЖНОСТИ

Белорусский государственный университет информатики и радиоэлектроники  
г. Минск, Республика Беларусь

Бледай А. С.

Гладкая В. С. – магистр техн. наук,  
ассистент каф. ИГиЭ

Цель работы – исследование идентификации по голосу и скрытые возможности. Среди систем биометрической идентификации наиболее широкое распространение (71% рынка) получили дактилоскопические системы, основанные на сравнении папиллярных узоров ладоней или пальцев рук. А вот возможность идентификации человека по рисунку радужной оболочки глаз, чертам лица (тем более по особенностям голоса) большинством граждан воспринимается как элемент научной фантастики.

Тем не менее, по сообщению биометрического портала, два зарубежных банка объявили о внедрении систем идентификации клиентов по голосу. Важно отметить, что в этих банках подавляющее большинство операций со счетами клиентов производятся с использованием телефона или сети Интернет.

Каковы же эти новые возможности? Рассмотрим их на примере технологий VoiceKey и VoiceNet, разработанных в компании "Центр речевых технологий".

Идентификация или верификация в технологиях VoiceKey и VoiceNet осуществляется по парольным фразам длительностью от 5 секунд или фрагментам спонтанной речи продолжительностью более 16 секунд. Обе технологии используются в качестве индивидуальных характеристик голоса, положения резонансных максимумов в спектре голоса, что обеспечивает надежную работу в условиях шумов и незначительных изменений эмоционального состояния человека. Но если VoiceKey ориентирован на сигналы микрофонного качества, то VoiceNet изначально адаптирован к работе с телефонными каналами.



Эти технологии дают возможность удаленной идентификации (сравнения с конкретным эталоном) или верификации (поиска в базе эталонов) клиентов; сложность или даже невозможность для злоумышленника имитировать голос с помощью магнитофона.

Во-первых, системы идентификации способны контролировать сразу несколько признаков, отличающихся от тех, что используются в речеслуховой системе, во-вторых, при воспроизведении записанной речи через миниатюрные громкоговорители в сигнал вносятся искажения, препятствующие идентификации говорящего. Невозможность идентификации человека, находящегося под угрозой насилия, поскольку эмоциональное состояние говорящего оказывает существенное влияние на характеристики голоса и речи. Возможность повышения надежности аутентификации за счет одновременного использования технологий идентификации по голосу и распознавания речи (произнесенного пароля).

Говоря о достоинствах технологий распознавания личности по голосу, нельзя обойти вниманием и некоторые их недостатки. Голос, в отличие от папиллярных узоров пальцев или ладоней, меняется с возрастом. Таким образом, клиентам придется периодически обновлять хранящийся в системе эталон речи. На голос оказывает влияние физическое и эмоциональное состояние человека в момент речевого акта. Так, например, система может не узнать по голосу человека, если он находится в состоянии алкогольного опьянения или у него во рту жевательная резинка, или он только что установил рекорд в беге.

Надежность работы системы в значительной степени зависит от качества канала передачи речевого сигнала к системе идентификации, в частности, от таких его характеристик, как частотный диапазон, уровень нелинейных искажений, отношение сигнал/шум, неравномерность частотной характеристики. Наивысшая надежность работы обеспечивается в том случае, когда эталон голоса клиента и его запрос поступают по одному и тому же каналу, например телефонному.

Список использованных источников:

1. [Журнал "Information Security/ Информационная безопасность" #3+4, 2006](#)
2. Speaker identification using mel frequency [Электронный ресурс]. – Режим доступа: [http://ethesis.nitrkl.ac.in/3745/1/final\\_yr\\_project\\_\\_thesis.pdf](http://ethesis.nitrkl.ac.in/3745/1/final_yr_project__thesis.pdf) (дата обращения 08.03.2018).