АНАЛИЗ НАПРАВЛЕНИЙ ПОВЫШЕНИЯ БЕЗОПАСНОСТИ ГОЛОСОВОЙ АУТЕНТИФИКАЦИИ ПОЛЬЗОВАТЕЛЕЙ СИСТЕМ ДОСТУПА

Пастушенко М.С., Красноженюк Я.О., Заіка М.В.

Кафедра «Інфокомунікаційної інженерії ім. В.В. Поповського», ХНУРЕ, Україна

E-mail: mykola.pastushenko@nure.ua
yana.krasnozheniuk@nure.ua
maksym.zaika@nure.ua

Abstract

The scientific problem of increasing the security of the voice authentication system is considered. To solve this problem, the main attention is currently paid to improving the quality of calculation of template elements based on the amplitude and frequency information of a voice signal. On their basis, solutions are formed using the Gaussian Mixture Model and Support Vector Machine, Hidden Markov Models or with the help of artificial neural networks. In this work, it is proposed to change the paradigm of digital processing of a voice signal when phase data is used as additional information to increase the security of the authentication system.

The results of experimental studies are given and it is shown that phase data make it possible to obtain adequate and reliable estimates of the elements of the template. The latter emphasizes the importance of phase information and its importance in solving other problems in the processing of voice signals.

Постановка задачи в общем виде

В настоящее время обостряется проблема обеспечения безопасности финансов, информации, услуг и ресурсов, доступ к которым осуществляется с помощью современных телекоммуникационных и компьютерных систем разного назначения. Об этом свидетельствуют многочисленные периодические сообщения в прессе о различных злоупотреблениях. Здесь же следует заметить, что западные финансовые учреждения пытаются не афишировать случаи хищения финансовых средств до 100 тысяч USD.

После драматических событий 11.09.2001 года странами G8 было принято решение, которое ориентировано на снижение рисков и повышение эффективности систем доступа различного назначения. Для этого предложено использовать в таких системах биометрические характеристики пользователя. В качестве основных характеристик рекомендовано использовать физиологические (статические) признаки пользователя, а именно, папиллярный узор пальцев, изображение лица и радужную оболочку глаза. Обусловлено это тем, что дактилоскопия, как и фото, очень широко и эффективно применяются в криминалистике для идентификации преступников. Более того, накоплены и интенсивно пополняются большие базы отпечатков, особенно в Западных странах. Здесь же заметим, что указанные статические биометрические признаки обладают ограниченной информативностью.

За последнее десятилетие биометрические технологии стали активно применяться во многих областях, связанных с обеспечением безопасности доступа к информации и материальным объектам, а также в задачах уникальной идентификации личности. Во многом этому способствовало распространение микропроцессорных технологий. Вместе с тем, предложенные биометрические признаки не позволили существенно повысить надежность систем доступа. Обусловлено это тем, что как в криминалистике, так и любой биометрической системе основными характеристиками являются два числа – FAR (False Acceptance Rate, ложный доступ в систему) и FRR (False Rejection Rate, ложный

отказ в доступе) [1]. Применительно к системе аутентификации первое число характеризует вероятность ложного совпадения биометрических характеристик двух людей. Второе – вероятность отказа доступа человеку, имеющего допуск. Система тем лучше, чем меньше значение FRR при одинаковых значениях FAR. В отличие от криминалистики, система аутентификации должна быть устойчива к подделке. Последнее не присуще криминалистике. Устойчивость к подделке – это эмпирическая характеристика, обобщающая то, насколько легко обмануть биометрическую систему.

Применительно к статическим биометрическим признакам можно констатировать, что они не удовлетворяют требованиям по устойчивости к подделке. Например, давно известны опыты японского криптографа Цутому Мацумото (Tsutomu Matsumoto), которые позволили вскрыть от 80 до 100% тестируемых дактилоскопических систем доступа. Из-за низкой устойчивости переходят от пространственных к трехмерным изображениям лица пользователя. Появились сообщения о подделке радужной оболочки глаза, которые формируют на основе нескольких фото с помощью современной фотоаппаратуры.

В связи с этим все больше внимания уделяется поведенческим (динамическим) признакам пользователя, а именно, подпись (форма букв, манера письма, нажим), голос, клавиатурный почерк и др. Основное преимущество указанных признаков – оперативное наращивание анализируемой последовательности по требованию системы. В общем случае можно утверждать, что указанные признаки имеют неограниченную информативность. Это существенно влияет на снижение величин FRR и FAR, а также повышает устойчивость к подделке.

Указанное преимущество особо ярко проявляются для систем голосовой аутентификации (СГА) [2]. Наряду с отмеченным выше, СГА обладают рядом дополнительных преимуществ, таких как: простота, компактность, дешевизна, возможность удаленной аутентификации с использованием телефонных каналов связи и др. [3]. Вопросам развития и внедрения этих систем сегодня посвящены многочисленные исследования и разработки, отдельные вопросы которых рассмотрены в работах Г. Фанта, Р.М. Болла, Г.С. Рамишвилли, Ф. Россе, В.Н. Сорокина, Г. Холлиена и др.

Здесь же следует отметить, что в современных СГА для идентификации используются преимущественно спектральные характеристики амплитуды речевого сигнала пользователя, а фаза материалов регистрации игнорируется [1]. При этом, давно известно, что фаза является более информативным параметром регистрируемого сигнала [4]. Поэтому в работе исследуется актуальная научно-техническая задача повышения безопасности голосовой аутентификации за счет изменения парадигмы цифровой обработки данных, которая ориентирована на использование фазовых данных материалов регистрации.

Выполненный анализ работ в области голосовой аутентификации показал, что открытым остается вопрос использования фазовых данных при оценке частоты основного тона, формантной информации, мел-частотных кепстральных коэффициентов (MFCC - Mel-frequency cepstrum coefficients) и др.. Поэтому цель данной работы – анализ использования фазовой информации голосового сигнала для оценки основных параметров голосовой аутентификации.

Методика и результаты исследований

При формировании признаков шаблона широкое применение нашли следующие признаки голосового сигнала: частота основного тона, формантная информация, спектральные и кепстральные коэффициенты.

Анализу подвергался речевой сигнал пользователя цифр от 0 до 9. Частота дискретизации сигнала составляла 64 кГц. Отношение сигнал/шум анализируемой последовательности составляло более 25 дБ.

При этом основное внимание будем уделять анализу диапазона спектра до 8 кГц, что обусловлено наличием отличительных признаков пользователя в его сигнале (в диапазоне от 0,1 кГц до 8 кГц).

Первый шаг методики проводимых исследований связан с формированием фазовых данных, которые, к сожалению, у голосового сигнала не регистрируются. В современных системах цифровой

обработки данных формирование фазовых данных базируется на использовании преобразования Гильберта. Преобразование Гильберта позволяет получить квадратурную составляющую голосового сигнала, на основе которой формировались фазовые данные.

Второй шаг связан с корректировкой фазовых данных, поскольку функция арктангенс формирует некорректный угол. Кроме этого, особенностью фазовых данных голосового сигнала – это априорно известные форма фазового сигнала и пределы его изменения. Этот факт дает основание использовать фазовые данные и для корректировки регистрируемой амплитудной и частотной информации голосового сигнала.

Третий шаг – использование фазовых данных для оценки элементов шаблона для системы голосовой аутентификации.

Последующие процедуры связаны с использованием соотношений расчета частоты основного тона, формантной информации и MFCC по амплитудно-частотной и фазовой информации голосового сигнала.

Последний шаг – сравнение результатов полученных в процессе анализа амплитудно-частотной и фазовой информации. В некоторых случаях для сравнения использовался нормированный коэффициент корреляции. При этом оценки полученные по амплитудно-частотных данных принимался в качестве эталона.

Обработка экспериментального голосового сигнала цифры «один» выполнялась в системе компьютерной математики MatLab.

В заключение представляются результаты оценки частоты основного тона, формантной информации и MFCC по амплитудно-частотной и фазовой информации голосового сигнала.

Для большинства случаев оценки полученные по амплитудно-частотной информации совпадают с оценками, рассчитанными по фазовым данным. Таким образом, исследуемые оценки шаблона системы голосовой аутентификации, полученные по фазовой информации, являются адекватными и достоверными.

Литература:

1. *Beigi H.*Fundamentals of Speaker Recognition. NY: Springer, 2011. 1029 p.

2. *Пастушенко Н.С., Педро В.Г., Файзулаева О.Н.* Исследование информативности фазовых данных голосового сигнала пользователя системы аутентификации / [Электронный ресурс] // Проблеми телекомунікацій. 2018. № 1 (22). С. 67 - 74. – Режим доступа к журн.: http://pt.nure.ua/2018/181_pastushenko_voice.pdf.

3. *Сорокин В.Н., Вьюгин В.В., Тананыкин А.А.* Распознавание личности по голосу: аналитический обзор // Информационные процессы. М.: РАН. 2012. Т. 12. № 1. С. 1–30.

4. *Oppenheim A.V., Lim J.S.* The Importance of Phase in Signals // Proceeding of the IEEE, 1981, t. 69(5), P. 529 - 541.