V.L. Evseev, Y.E. Kozlov

Modern Methods of Voice Authentication in Mobile Devices

Keywords: internet banking, voice biometrics, mobile authentication, multimodal access, evaluation of reliability.

Modern methods of voice authentication in mobile devices.The proposed evaluation of the probability errors of the first and second kind for multi-modal methods of voice authentication. The advantages of multimodal multivariate methods before, when authentication takes place in several stages – this is the one-stage, which means convenience for customers. Further development of multimodal methods of authentication will be based on the significantly increased computing power of mobile devices, the growing number and improved accuracy built-in mobile device sensors, as well as to improve the algorithms of signal processing.

В.Л. Евсеев, Ю.Е. Козлов

СОВРЕМЕННЫЕ МЕТОДЫ РЕЧЕВОЙ АУТЕНТИФИКАЦИИ В ПРИЛОЖЕНИЯХ МОБИЛЬНЫХ УСТРОЙСТВ

Речевая автоматическая идентификация и аутентификация личности бесконтактна и не требует от субъекта особых усилий. В связи с этим в России и за рубежом активно ведутся работы по развитию систем речевой аутентификации в мобильных приложениях. Кроме того, автоматическое распознавание клиентов – это один из важных фрагментов голосовых интерфейсов, чья роль в жизни людей непрерывно растет. Так, например, одним из распространённых сервисов является интернет-банкинг – предоставление клиенту доступа к своему счёту. В пользу использования технологии биометрической аутентификации по голосу для этого сервиса, есть несколько аргументов:

– достаточно легко создать банк образцов голоса, так как для этого не требуется дополнительного специального оборудования, голос записывается через микрофон телефона;

– запись голоса – это простая и понятная любому человеку процедура, она понятна клиенту.

В данной статье под термином «аутентификация» подразумевается проверка принадлежности субъекту доступа предъявленного им идентификатора, т.е. подтверждение его подлинности [1]. Таким образом, идентификация и аутентификация являются взаимосвязанными процессами распознавания и проверки подлинности пользователей.

Существует два основных подхода к архитектуре системы речевой аутентификации в мобильных приложениях:

– первый предполагает аутентификацию, не использующую сетевые соединения. Для такого подхода характерна «тонкая» первичная настройка на пользователя. Предполагается индивидуальное использование мобильного устройства, что вызвано ограничениями ресурсов мобильных устройств. Данный подход к аутентификации можно отнести к такому типу цифровых кодовых замков, где ключом к замку являться голос пользователя [2];

– второй подход осуществляет аутентификацию, используя клиент-серверную архитектуру. В данном случае речевые параметры передаются по радиоканалу на сервер для анализа. Этот подход является основным для банковских сервисов, предполагающих речевую аутентификацию с применением мобильных устройств, так как он не зависит от типа мобильного устройства, которое использует клиент.

Очевиден тот факт, что использование методов для аутентификации, включающих проверку знания контрольной фразы, дает дополнительные преимущества. В этом

случае проверяется знание клиентом самой фразы, производится тонкая настройка параметров речевой аутентификации за счет того, что необходимо анализировать речевые параметры короткого фрагмента с заранее известным содержанием. Поэтому такие методы речевой аутентификации получили наибольшее распространение. Важнейшим при аутентификации клиента является этап распознавания речи. В настоящее время над проблемой распознавания речи успешно работают множество фирм. И здесь главным в аутентификации клиентов является распознавание речи.

Системы распознавания речи, использующие мобильные устройства, можно классифицировать по множеству типов. Одна из классификаций представлена на рис. 1.

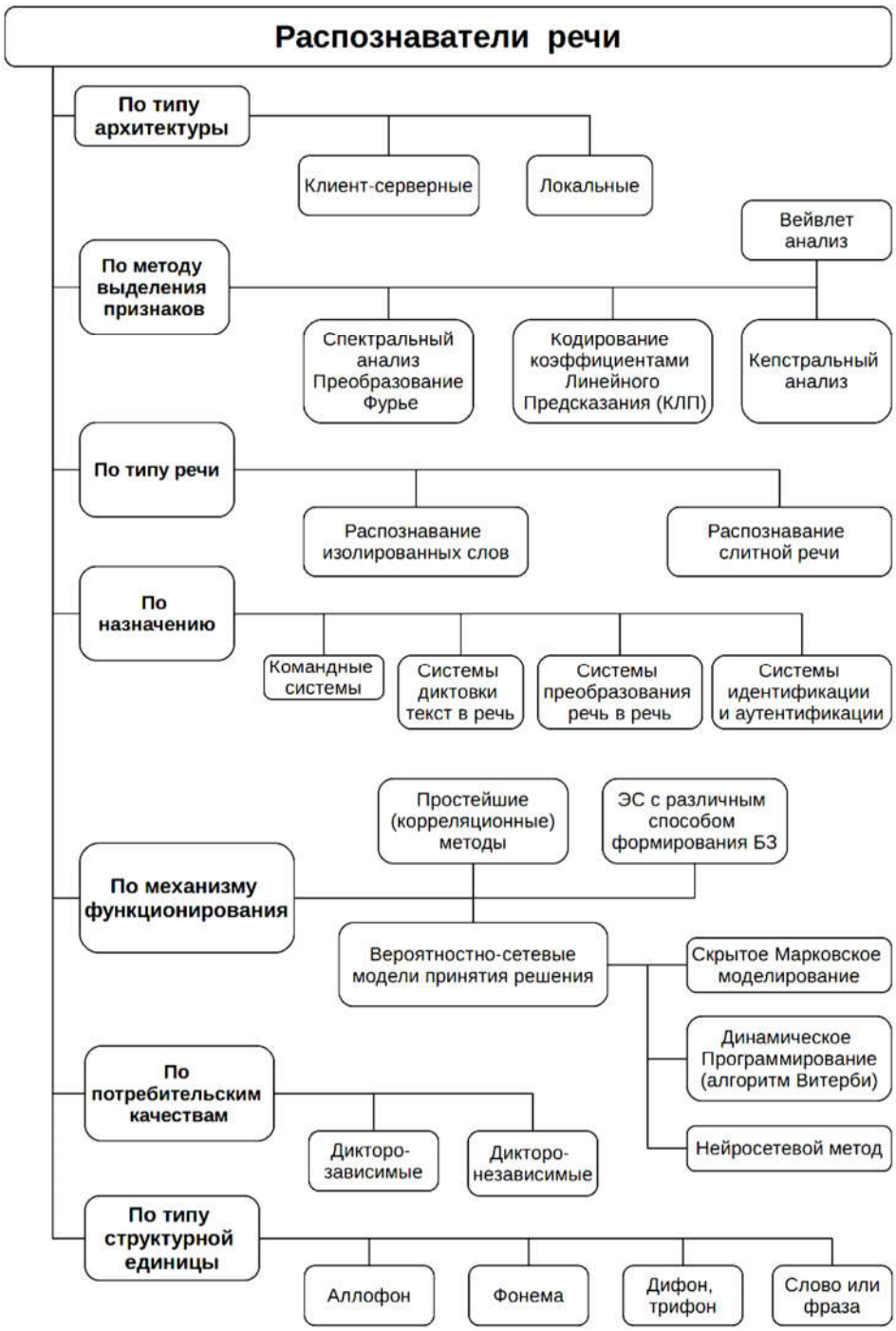

Рис. 1. Классификация распознавателей речи

При выполнении процедуры аутентификации в процессе распознавания речи клиентов, использующие мобильные устройства, могут иметь место ошибки двух типов:

1) ложное отклонение (FRR англ. FalseRejectionRate);

2) ложный допуск (FAR англ. FalseAcceptanceRate).

Ошибка ложного отклонения (ошибка 1-го рода), когда система верификации отвергает истинную идентичность «своего», характеризуется вероятностью ложной тревоги. А при допуске «чужого» имеет место ошибка ложного допуска (ошибка 2-го рода), характеризующаяся вероятностью пропуска «чужого». Тогда обобщенной характеристикой будет средняя вероятность ошибки, определяемая как полу сумма вероятностей ошибок 1-го и 2-го рода.

В результате каждая система может перестраиваться таким образом, что ошибки одного рода могут быть уменьшены за счет увеличения ошибок другого рода путем изменения порога принятия решения. Для оценки качества систем аутентификации можно использовать критерий ERR (Equal Error Rate), когда порог принятия решения выбирается так, чтобы обеспечить равенство обеих ошибок [3].

Алгоритм работы системы речевой аутентификации представлен на рис. 2.

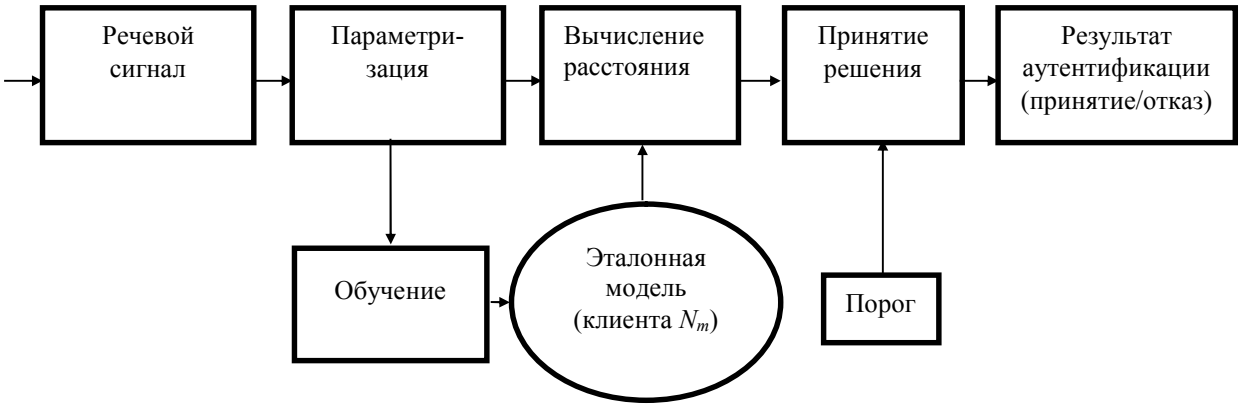

Рис. 2. Типовая схема аутентификации личности по голосу

Качество работы (точность) систем аутентификации и идентификации личности по голосу можно оценить следующим образом:

Пустьвероятности ошибок 1-го и 2-го рода будут соответственно равны FRR = α, FAR = β, тогда средняя вероятностью ошибки P_e можно рассчитать по формуле

$$P_e = \frac{1}{2}(\alpha + \beta). \quad (1)$$

Из (1) следует, что система аутентификации работает наилучшим образом при α и $\beta \to 0$ в результате $P_e \to 0$.

Тем не менее вероятности ошибок первого и второго рода для речевой аутентификации достаточно велики в банковских системах. Например, для системы «Голос» фирмы «ЦРТ»: FRR ≈ 4 %, а FAR ≈ 1 % [4]. В связи с этим на первый план выходят мультимодальные методы аутентификации. Такие системы используют несколько типов биометрических характеристик и позволяют соединять несколько типов биометрических технологий аутентификации в одну, что позволяет удовлетворять самым строгим требованиям к эффективности системы аутентификации.

Формулы для расчета вероятности средней ошибки аутентификации

Предположим, что имеются n независимых биометрических характеристик, для которых $P_{e1},\ldots,P_{en}$ – средние вероятности ошибок каждой из технологий биометриче-

ской аутентификации. Тогда средняя вероятность ошибки P_E при применении всех технологий одновременно может быть рассчитана по формуле

$$P_E = \frac{P_1 \times P_2 \times \ldots \times P_n}{P_1 \times P_2 \times \ldots \times P_n + (1-P_1) \times (1-P_2) \times \ldots \times (1-P_n)}. \quad (2)$$

Под независимостью биометрических характеристик будем понимать характеристики, регистрируемые разными датчиками с различных источников (например, голос, регистрируемый микрофоном, и изображение лица, регистрируемое видеокамерой).

Следует отметить, что ряд мультимодальных методов, использующих технологии для биометрической аутентификации, не являются независимыми. Это характерно для тех видов речевой аутентификации, где в качестве источника сигнала выступают речевые параметры, регистрируемые разными датчиками. Например, когда:

– аутентификация производится по нескольким речевым параметрам;

– аутентификация производится на основе речи, регистрируемой двумя микрофонами;

– для аутентификации, кроме речевой аутентификации, применяется аутентификация с использованием ларингофона, что доказано при контактно-разностном способе аутентификации [5].

– для аутентификации, наряду с речевой аутентификацией, применяют детектор присутствия по голосу [6].

Во всех вышеописанных случаях для расчета вероятности средней ошибки P_E требуется введение коэффициента корреляции σ, который, по сути, более точно охарактеризует систему аутентификации с точки зрения ее надежности (стойкости ко взлому), поскольку при достаточном техническом оснащении такие системы гипотетически могут быть взломаны с использованием одного поддельного источника речевого сигнала. Для этих систем расчет вероятности средней ошибки P_E будет иметь следующий вид:

$$P_E = \frac{P_1 \times P_2 \times \ldots \times P_n}{P_1 \times P_2 \times \ldots \times P_n + (1-P_1) \times (1-P_2) \times \ldots \times (1-P_n)} \text{x } \sigma, \quad (3)$$

где σ > 1 – коэффициент корреляции методов аутентификации.

Вывод

Мультимодальные методы имеют одно неоспоримое преимущество перед многофакторными: когда аутентификация проходит в несколько этапов, – это их одноэтапность, а значит, удобство для клиентов. Дальнейшее развитие методов мультимодальной аутентификации будет опираться на значительное увеличение вычислительных мощностей мобильных устройств, рост числа и повышение точности встроенных в мобильные устройства датчиков, а также на совершенствование алгоритмов обработки сигналов.

СПИСОК ЛИТЕРАТУРЫ:

1. Российская Федеральная государственная техническая комиссия. Защита от несанкционированного доступа к информации. Термины и определения. Руководящий документ. М., 1992.

2. Brunet Kevin, TaamKarim, Cherrier Estelle, Faye Ndiaga, Rosenberger Christopher. Solution, Speaker Recognition for Mobile User Authentication: An Android. https://hal.archives-ouvertes.fr/hal-00848318/ (дата обращения: 18.11.2015).

3. Голубинский А.Н., Булгаков О.М. Математические модели речевых сигналов для верификации и идентификации личности по голосу. Воронеж: Издательско-полиграфический центр Воронежского государственного университета. 2010.

4. Симончик К.К., Белевитин Д.О., Матвеев Ю.Н., Дырмовский Д.В. Доступ к интернет-банкингу на основе бимодальной биометрии // Мир измерений. 2014. Вып. № 3. С. 7.
5. Голубинский А.Н., Дворянкин С.В. К вопросу о параметризации результатов акустического зондирования тела человека при реализации контактно-разностного метода аудиоидентификации // Спецтехника и связь. 2011. Вып. № 2. С. 38–43.
6. Shchemelinin V., Simonchik K. ExaminingVulnerability of Voice VerificationSystems to Spoofing Attacks by Means of a TTSSystem // Speech and Computer. Springer International Publishing, 2013. – Pp. 132–137.

REFERENCES:

1. Federal Russian state technical Commission. Protection against unauthorized access to information. Terms and definitions. Guidance document. Moscow. 1992.
2. Brunet Kevin, TaamKarim, Cherrier Estelle, Faye Ndiaga, Rosenberger Christopher. Solution, Speaker Recognition for Mobile User Authentication: An Android. https://hal.archives-ouvertes.fr/hal-00848318/ (дата обращения: 18.11.2015).
3. Golubinsky A.N., Bulgakov O.M. Mathematical models of speech signals for verification and identification by voice. Voronezh. 2010.
4. Simonchik K.K., Belevitin O.D., Matveev Yu.N., Dyrmovsky D.V. Access Internet banking on the basis of bimodal biometrics // The world of measurement. 2014. No. 3.
5. Golubinsky A.N., Dvoryankin S.V. To the question of parameterization of results of acoustic sounding of the human body in the implementation of the contact-difference method autoidentification // Special equipment and communication. 2011. No. 2. Pp. 38–43.
6. Shchemelinin V., Simonchik K. ExaminingVulnerability of Voice VerificationSystems to Spoofing Attacks by Means of a TTS System // Speech and Computer. Springer International Publishing, 2013. Pp. 132-137.