УДК 612.087.1:57.087.1

АНАЛИЗ КАЧЕСТВЕННЫХ ПОКАЗАТЕЛЕЙ БИОМЕТРИЧЕСКИХ СИСТЕМ АУТЕНТИФИКАЦИИ ПОЛЬЗОВАТЕЛЕЙ

О.Н. ПАСТУШЕНКО,
И.Ш. НЕВЛЮДОВ
Харьковский национальный университет радиоэлектроники

Abstract *– The scientific optimization problem for decision-making in biometric systems from the perspective of the minimum average risk criterion is considered. Rapid development and rather wide use of biometric systems in various spheres of human activity, including in modern telecommunication systems, highlights the reliability of their performance. The work is based on the method of biometric systems theory of mathematical statistics, which is widely and effectively used in a number of modern engineering systems. For example, in modern radars conditional probability of false alarm is at the level and below. The requirements for errors of the 1st kind are not so stringent. It is this way is a key to select the requirements for the quality characteristics of biometric systems. A common approach for the equality of errors of the first and the second kind, which is widely used in present-day biometric systems, is not constructive, because it is valid only for the particular case, which is far from practical needs.*

In this article the weight optimization criterion of quality parameters for biometric identification systems (authentication), based on a more general criterion for minimum average risk of erroneous decisions, is obtained. Practical recommendations for the choice of their quality indicators are given. Further research will focus on the synthesis of decision rules for specific biometric systems.

Анотація *– Розглянуто наукове завдання оптимізації прийняття рішень у біометричних системах з позиції критерію «мінімуму середнього ризику». Зазначено практичні рекомендації з вибору їхніх якісних показників.*

Аннотация *– Рассмотрена научная задача оптимизации принимаемых решений в биометрических системах с позиции критерия «минимума среднего риска». Указаны практические рекомендации по выбору их качественных показателей.*

Введение

Биометрия появилась в конце XIX века как раздел науки, занимающейся количественными биологическими экспериментами с привлечением статистических методов. Полвека назад интерес к биометрии получил новый импульс в связи с появлением биометрических систем безопасности [1].

В классическом понимании биометрия – это наука, которая изучает методы идентификации (опознания, попросту говоря) конкретной личности на основе ее персональных физиологических или поведенческих характеристик [1]. До недавних пор основными потребителями биометрических методов идентификации были всевозможные правоохранительные структуры и спецслужбы. Однако в последние полтора десятка лет активнее всех других биометрическими методами идентификации стали интересоваться владельцы самых различных информационных систем [2-3]. Связано это, прежде всего, с активным проникновением компьютерных и телекоммуникационных технологий в бизнес и, как следствие, стремительным ростом ценности информации как таковой. А также, конечно, с появлением в глобальной сети

колоссальных объемов информации, доступ к которой должен быть ограничен. Международная Биометрическая Группа (International Biometric Group, IBG) отмечала увеличение доходов в индустрии биометрии почти в 7 раз – от 0,6 (в 2002 г.) до 4,04 (в 2007 г.) миллиардов долларов.

I. Общая характеристика биометрических систем

Собственно биометрическая идентификация личности как таковая не является, разумеется, самоцелью. Цепочка действий выглядит примерно так [1]:

1. Считываются биометрические данные пользователя.
2. Посредством обращения к локальной или внешней базе данных, заранее сформированных шаблонов с признаками пользователя, устанавливается его личность.
3. Опять же обращением к базе данных устанавливается список прав и обязанностей пользователя.
4. Принимается решение, зависящее от конкретной задачи.

Задачами, решаемыми с участием биометрических систем идентификации (аутентификации), могут служить [4, 5]:

- определение прав физического доступа – в охранных системах: от дверного замка или блокировки запуска автомобиля до пропуска на территорию предприятия и т.д.;

- определение прав виртуального доступа – в терминалах компьютерных или банковских сетей; системах удаленного доступа к ресурсам.

- учет и контроль – в государственных (например, системы контроля, охраны и допуска) или частных (например, системы маркетинговых исследований) организациях.

Основным преимуществом биометрических систем является интерфейсная простота их взаимодействия с клиентом.

Поэтому одной из самых популярных тем последних лет в области информационной безопасности стала биометрическая идентификация (аутентификация). Между тем, как все чаще выясняется, очень и очень многие люди (даже из числа специалистов) имеют о биометрии достаточно туманное представление.

Вероятно, основной проблемой биометрии является вопрос ее надежности. Понятие надежности, как правило, разделяют на три большие области [1]. Первую из них регулярно обсуждают сами производители биометрического оборудования. Речь идет о вероятностном характере производимой биометрическими устройствами идентификации. Поскольку условия сканирования каждый раз несколько отличаются, а сканируемые части тела или поведенческие рефлексы клиента также не вполне постоянны, можно говорить не о точном совпадении измерения с шаблоном (как это происходит, например, при сравнении с эталоном вводимого в компьютер пароля), а лишь о величине вероятностной меры правильного отождествления.

Поэтому все биометрические устройства характеризуются параметрами: «вероятность непризнания своего» (то есть вероятность не идентифицировать зарегистрированного пользователя системы) и «вероятность признания своим чужого» (то есть вероятность неверного отождествления постороннего с кем-то из легальных

пользователей). Именно эти характеристики биометрических систем и будут рассмотрены ниже в статье.

Второй аспект (не)надежности биометрических систем фирмами-производителями старательно замалчивается. Речь идет о защищенности систем от сознательного обмана, о способах симулировать объект биометрического сканирования. Известны способы обмана биометрических систем контроля доступа по отпечатку пальца. Например, японский криптограф Цутому Мацумото и группа его студентов в Университете Иокогамы (отнюдь не профессионалов-взломщиков) наглядно показали, как с помощью простейшего инвентаря и материалов можно обмануть практически любую из таких систем. Японские студенты проверили 12 коммерческих сканирующих устройств. И каждое из них смогли обмануть, в среднем в четырех случаях из пяти.

Специалистам в области биометрии все эти факты были давно известны, однако результаты подобных исследований сознательно замалчиваются. Выход из положения не прост, он требует привлечения более сложных в использовании и более дорогих методов биометрии (а лучше – многофункциональную аутентификацию), что сразу ставит под удар саму идею повсеместного распространения биометрических технологий. Приемлемого решения на данный момент можно добиться комбинированной проверкой – считыванием нескольких параметров, например отпечатка пальца и голоса, использованием биометрического контроля вместе со смарт-картами и т.п.

Наконец, третьим аспектом проблемы надежности является вопрос сохранности собранной биометрической информации. Большинство биометрических систем уязвимы для взлома посредством перехвата, сохранения и последующего воспроизведения данных. Насколько это осуществимо, зависит от метода передачи биометрической информации по сети. Однако это еще только полбеды.

Хуже то, что любой биокод, в отличие от безличного кода-пароля, практически всегда несет в себе гораздо больше информации, чем это нужно устройству для проверки доступа. Даже рисунок радужной оболочки глаза, не говоря уж о ДНК-коде, может сообщить специалисту важную информацию о состоянии индивидуума, его врожденных или приобретенных свойствах, в том числе болезнях. А эта информация, очевидно, является слишком интимной, чтобы давать доступ к ней не только своему лечащему врачу. Возможные злоупотребления очевидны каждому – от дискриминации при приеме на работу до прямого шантажа.

Вместе с тем, в современных биометрических системах для распознавания личности используются различные физиологические и поведенческие характеристики, такие как лицо, отпечатки пальцев, радужная и сетчатая оболочки глаза, голос, ручная подпись, геометрия руки, рисунок вен на руке и т. д.

В основу работы биометрических систем положена математическая статистика (а именно, проверка гипотез [6]), алгоритмы которой интенсивно используются в ряде современных технических систем, таких как: связь, радиолокация (различные радары), множестве байесовских систем. В качестве двух основных характеристик любой биометрической системы, построенной на основе статистической теории

проверки гипотез (тестов), можно принять ошибки первого и второго рода [6, 7]. В теории радиолокации их обычно называют «ложная тревога» и «пропуск цели», а в биометрии, наиболее устоявшиеся понятия – FAR (False Acceptance Rate, ложное распознавание) и FRR (False Rejection Rate, ложный отказ). Первая величина характеризует вероятность ложного совпадения биометрических характеристик двух пользователей (по существу, «допуск хакера на свою территорию»). Вторая величина – вероятность отказа доступа пользователю, который имеет допуск.

Ряд авторов [2-5] в качестве оптимального варианта выбора значений указанных ошибок предлагают использовать сравнительную характеристику EER (Equal Error Rate, равный коэффициент ошибок). Эта характеристика определяет точку, в которой величины FRR и FAR равны. Цель данной статьи – проанализировать справедливость данного утверждения.

Для достижения указанной цели рассмотрим научную задачу оптимизации принимаемых решений в биометрической системе по критерию минимума среднего риска, который широко используется в теории связи, радиолокации и других технических системах.

II. Качественные показатели биометрических систем

Для пояснения вводимых вероятностных характеристик биометрических систем, по аналогии с радиолокацией [7], рассмотрим две плотности вероятности, которые характеризуют шаблон пользователя и шаблон хакера (см. рис. 1). Естественно предположить, что эти шаблоны могут быть заданы в виде нормальных распределений. Поскольку шаблон пользователя получен в процессе обучения системы, то его основная характеристика (математическое значение) $q1$ имеет большое значение. Для пользователя-хакера, который пытается проникнуть в вычислительную систему, шаблон получен оперативно в процессе применения биометрической системы по назначению и имеет характеристику $q0$.

Введем и третью величину qh, которая может определять принимаемые решения. В результате оперативного анализа полученных биометрических характеристик и шаблонов, хранящихся в базе системы, принимается решение о характере допуска текущего пользователя: отнести его к пользователям системы или пользователям (хакерам), которые не допущены к ресурсам системы. В простейшем случае решение принимается при двух взаимно исключающих условиях:

- условие A_1 – биометрические характеристики принадлежат пользователю системы;

- условие A_0 – биометрические характеристики принадлежат пользователю-хакеру, не допущенному к ресурсам системы.

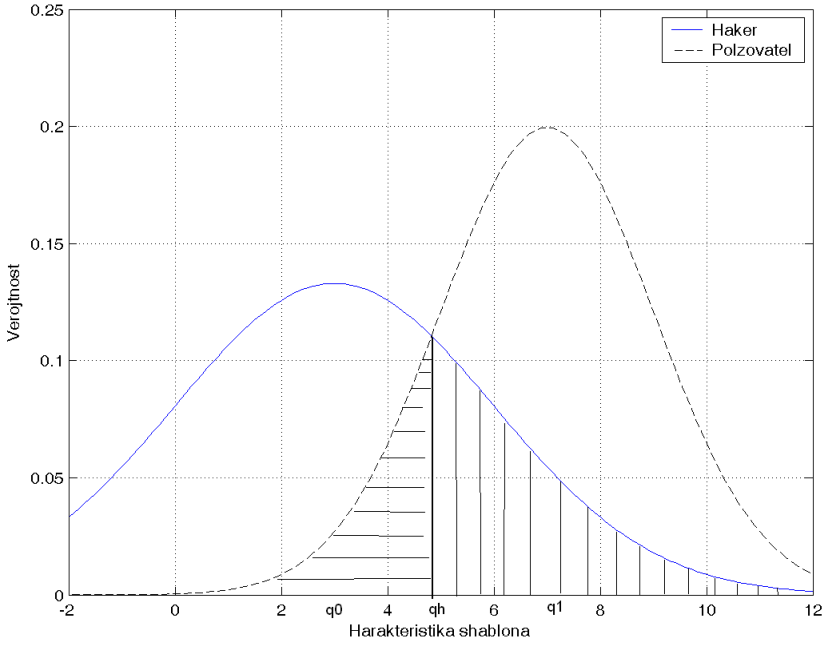

Рис. 1. Плотности вероятности анализируемых шаблонов

При автоматическом принятии решения в системе эти условия неизвестны.

В процессе анализа биометрических характеристик текущего пользователя и шаблона каждому условию могут соответствовать два вида решений:

- решение $\hat{A}_1$ – биометрические характеристики соответствуют шаблону, который хранится в базе системы;

- решение $\hat{A}_0$ – биометрические характеристики принадлежат не допущенному к ресурсам пользователю-хакеру.

При этом возможны четыре ситуации совмещения случайных событий «решения» и «условия»:

- $\hat{A}_1 A_1$ – правильный допуск пользователя к ресурсам системы;

- $\hat{A}_0 A_1$ – запрет на допуск пользователю (пользователь воспринят системой как хакер), ошибка 1-го рода (ложный отказ, False Rejection Rate, FRR);

- $\hat{A}_1 A_0$ – допуск к ресурсам системы пользователя-хакера, ошибка 2-го рода (ложное распознавание, False Acceptance Rate, FAR);

- $\hat{A}_0 A_0$ – запрет доступа к ресурсам пользователю-хакеру.

Перечисленным ситуациям соответствуют четыре вероятности совмещения событий, сумма которых равна единице:

$$P(\hat{A}_1 A_1) + P(\hat{A}_0 A_1) + P(\hat{A}_1 A_0) + P(\hat{A}_0 A_0) = 1. \qquad (1)$$

Как правило [7], каждому ошибочному решению ставят в соответствие некоторую плату – стоимость ошибки r_{ik} $(i = 0,1;\ k = 0,1)$. Для безошибочных решений эту стоимость можно считать равной нулю, т.е. $r_{11} = r_{00} = 0$. Тогда систему допуска

можно характеризовать средней стоимостью (математическим ожиданием стоимости) ошибочных решений

$$M\{r\} = \bar{r} = r_{01} \cdot P(\hat{A}_0 A_1) + r_{10} \cdot P(\hat{A}_1 A_0). \quad (2)$$

Лучшей из сравниваемых систем допуска следует считать ту систему, которая удовлетворяет минимуму стоимости $M\{r\}$ – критерию минимума среднего риска.

В связи с тем, что, как правило, отсутствует информация об априорных (доопытных) вероятностях $P(A_1)$ и $P(A_0)$, затруднителен и расчет вероятностей совмещения $P(\hat{A}_0 A_1)$ и $P(\hat{A}_1 A_0)$. Поэтому переходят к условным вероятностям, позволяющим получить качественные показатели исследуемых систем.

Качественными показателями в условиях аутентификации пользователя системы являются условные вероятности: правильного допуска к ресурсам

$$D = P(\hat{A}_1 \mid A_1) = P(\hat{A}_1 A_1) / P(A_1) \quad (3)$$

и ложного отказа (запрет на допуск пользователю, горизонтальная штриховка на рис. 1)

$$\hat{D} = P(\hat{A}_0 \mid A_1) = P(\hat{A}_0 A_1) / P(A_1). \quad (4)$$

Поскольку соответствующие одному и тому же условию A_1 решения $\hat{A}_1$ и $\hat{A}_0$ взаимоисключающие, то

$$D + \hat{D} = 1. \quad (5)$$

Качественными показателями принятия решения при аутентификации пользователя-хакера являются условные вероятности: ложного распознавания (допуск хакера к ресурсам системы, вертикальная штриховка на рис. 1)

$$F = P(\hat{A}_1 \mid A_0) = P(\hat{A}_1 A_0) / P(A_0) \quad (6)$$

и правильного запрета аутентификации (доступа)

$$\hat{F} = P(\hat{A}_0 \mid A_0) = P(\hat{A}_0 A_0) / P(A_0), \quad (7)$$

причем

$$F + \hat{F} = 1.$$

Используя приведенные соотношения (3)–(7), выражение (2) для средней стоимости ошибочных решений можно представить в виде

$$\bar{r} = r_{01} \cdot \hat{D} \cdot P(A_1) + r_{10} \cdot F \cdot P(A_0)$$

или, после замены $\hat{D} = 1 - D$ и простых преобразований,

$$\bar{r} = r_{01} \cdot P(A_1) \cdot [1 - (D - l_0 \cdot F)], \quad (8)$$

где

$$l_0 = \frac{r_{10} \cdot P(A_0)}{r_{01.} \cdot P(A_1)}. \quad (9)$$

При этом критерий оптимизации аутентификации (допуска) по минимуму среднего риска сводится к так называемому весовому критерию

$$D - l_0 \cdot F = \mathbf{max}. \quad (10)$$

Этот критерий показывает, что по совокупности требований повышения условной вероятности правильного допуска к ресурсам D и понижения условной вероятности ложного распознавания F следует стремиться к увеличению «взвешенной» разности $D - l_0 \cdot F$.

Множитель l_0, называемый весовым множителем, зависит от соотношения стоимостей ошибочных решений каждого вида и величин априорных вероятностей, рассматриваемых условий принятия решения A_1 и A_0. Заметим, что критерий оптимизации (10) является следствием более общего критерия минимума среднего риска.

Обратим внимание на следующее. При $P(A_1) = P(A_0)$ и $r_{01} = r_{10}$, величина множителя $l_0 = 1$. В этом случае критерий (10) преобразуется к виду

$$D - F = \mathbf{max},$$

или, что эквивалентно – $F = \hat{D}$. В ряде работ [2-5] этот режим работы биометрической системы считается оптимальным. Для проверки достоверности этого утверждения, проанализируем его более подробно.

Равенство $P(A_0) = P(A_1)$ свидетельствует о том, что априорные вероятности появления пользователя и хакера на входе биометрической системы одинаковы и равны 0,5. В действительности $P(A_1) >> P(A_0)$. Более того, не равны и стоимости ошибочных решений. Естественно предположить, что стоимость ошибочного запрета доступа зарегистрированному пользователю r_{01} значительно меньше стоимости r_{10}, которая характеризует условную вероятность ложного распознавания («допуск хакера на охраняемую территорию»). В этих условиях, как правило, $l_0 > 1$, а значит и $\hat{D} > F$. Заметим, что ошибки 1-го рода (ложный отказ, False Rejection Rate, FRR) имеют менее тяжелые последствия (требуют повторной регистрации пользователя), в отличие от ошибок 2-го рода (ложное распознавание, False Acceptance Rate, FAR), которые приводят к допуску хакера к ресурсам и услугам, например, телекоммуникационной системы.

Выводы

Бурное развитие и достаточно широкое использование биометрических систем в различных сферах человеческой деятельности, в том числе и в современных телекоммуникационных системах, выдвигает на первый план надежность их функционирования. В основу работы биометрических систем положены методы теории

математической статистики, которая очень широко и эффективно используется в ряде современных технических систем. Например, в современных радарах условная вероятность ложной тревоги находиться на уровне 10^{-12} и ниже. При этом требования к ошибкам 1-го рода не такие жесткие. Именно этот путь является основным для выбора требований к качественным характеристикам биометрических систем.

Распространенный подход, широко используемый в современных биометрических системах, о равенстве ошибок первого и второго рода не является конструктивным, поскольку он справедлив только для частного случая, который далек от потребностей практики. В статье получен весовой критерий оптимизации качественных показателей биометрических систем идентификации (аутентификации), который базируется на более общем критерии минимума среднего риска принятия ошибочных решений. Дальнейшие исследования будут направлены на синтез решающих правил для конкретных биометрических систем.

Список литературы:

1. *Болл Р.М.* Руководство по биометрии: пер. с англ. Н.Е. Агаповой. – М.: Техносфера, 2007. – 368 с.

2. Распознавание личности по голосу: аналитический обзор / В.Н. Сорокин, В.В. Вьюгин, А.А. Тананыкин // Информационные процессы. – 2012. – Том 12, №1. – С. 1-30. Режим доступа: http://www.jip.ru/2012/1-30-2012.pdf.

3. Традиционные методы биометрической аутентификации и идентификации / В.М. Колешко, Е.А. Воробей, П.М. Азизов, А.А. Худницкий, С.А. Снигерев. – Минск: БНТУ, 2009. 107 с.

4. *Сущенок О.А.* Оценка эффективности работы биометрических систем // Системи обробки інформації. – 2011. – Вип. 4 (94). – С. 79-81. Режим доступу: http://www.nbuv.gov.ua/portal/natural/soi/2011_4/sushen.pdf.

5. *Лысак А.Б.* Идентификация и аутентификация личности: обзор основных биометрических методов проверки подлинности пользователя компьютерных систем // Математические структуры и моделирование. – 2012. – Вып. 26. – С. 124-134.

6. *Бронштейн И.Н., Семендяев К.А.* Справочник по математике для инженеров и учащихся втузов. – М.: Наука, Гл. ред. физ.-мат. лит., 1986. – 544 с.

7. Теоретические основы радиолокации / Под ред. Ширмана Я.Д. – М.: Сов. радио. 1970. – 560 с.