ОСНОВНЫЕ ПАРАМЕТРЫ БИОМЕТРИЧЕСКИХ СИСТЕМ

Михайлов Алексей Алексеевич
начальник сектора отдела ФКУ НИЦ «Охрана» МВД России, подполковник полиции,
Колосков Алексей Анатольевич
старший научный сотрудник ФКУ НИЦ «Охрана» МВД России, подполковник,
Дронов Юрий Иванович
старший научный сотрудник ФКУ НИЦ «Охрана» МВД России

ВСТУПЛЕНИЕ

В настоящее время наблюдается бурное развитие биометрических систем контроля и допуска (далее биометрии) как за рубежом, так и в России. Действительно, использование биометрии для целей охраны чрезвычайно привлекательно. Любой ключ, таблетку – Touchmemory, Proxy-карту или другой материальный идентификатор можно украсть, сделать дубликат и таким образом получить доступ к объекту охраны.

Цифровой пин-код (вводится человеком с помощью клавиатуры) можно зафиксировать с помощью банальной видеокамеры, и потом есть возможность шантажа человека или угрозы физического воздействия на него с целью получения значения кода. Редко кто из читателей, на собственном опыте или на опыте своих знакомых, не сталкивался с таким способом мошенничества. Появился даже термин, обозначающий данный способ изъятия честно заработанных денег у граждан, – скимминг (от англ. skim – снимать сливки).

Биометрический идентификатор невозможно украсть или получить путем шантажа, что делает в перспективе его очень привлекательным для целей охраны и доступа. Правда, можно попытаться создать имитатор биологического признака человека, но тут должна проявить себя в полной мере биометрическая система и отвергнуть подделку.

Вопрос «обхода» биометрических систем – это большая и отдельная тема, и в рамках этой статьи мы не будем ее затрагивать, да и создать имитатор биологического признака человека – непростая задача.

Особенно отрадно отметить активное развитие данного направления охранной техники в России. Например, «Русское общество содействия развитию биометрических технологий, систем и коммуникаций» существует с 2002 года.

Существует и технический комитет по стандартизации ТК 098 «Биометрия и биомониторинг», который работает достаточно плодотворно (выпущено более 30 ГОСТ, см.: http://www.rusbiometrics.com/), но нас, как пользователей, больше всего интересует ГОСТ Р ИСО/МЭК19795-1-2007 «Автоматическая идентификация. Идентификация биометрическая. Эксплуатационные испытания и протоколы испытаний в биометрии. Часть 1. Принципы и структура».

ТЕРМИНЫ И ОПРЕДЕЛЕНИЯ

Для того чтобы понимать, о чем пишут в нормативных документах, необходимо определиться в терминах и определениях. Чаще всего по своему физическому принципу пишут об одном и том же, но называют совершенно иначе. Итак, о наиболее значимых параметрах в биометрии:

VERIFICATION (верификация) – процесс, при котором происходит сравнение представленного пользователем образца с шаблоном, зарегистрированным в базе данных (ГОСТ Р ИСО/МЭК19795-1-2007). Здесь принципиальным является, что один образец сравнивается с одним шаблоном (сравнение один к одному с биометрическим шаблоном), поэтому любая биометрическая система будет иметь лучшие показатели для верификации по сравнению с идентификацией.

IDENTIFICATION (идентификация) – процесс, при котором осуществляется поиск в регистрационной базе данных и предоставляется список кандидатов, содержащих от нуля до одного или более идентификаторов (ГОСТ Р ИСО/МЭК19795-1-2007). Здесь принципиальным является, что один образец сравнивается со многими шаблонами (сравнение один ко многим), и ошибка системы многократно возрастает. Идентификация становится наиболее критичным параметром для систем биометрии, основанной на распознавании характерных черт лица человека. Для машины лица людей практически идентичны.

FAR (False Acceptance Rate) – вероятность несанкционированного допуска (ошибка первого рода), выраженное в процентах число допусков системой неавторизованных лиц (имеется в виду верификация). Вероятностные параметры

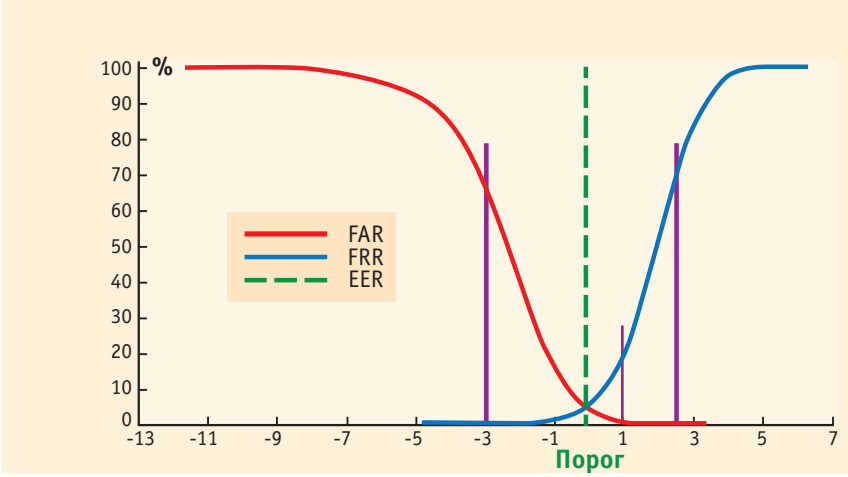


Рис. 1. Графики FAR и FRR

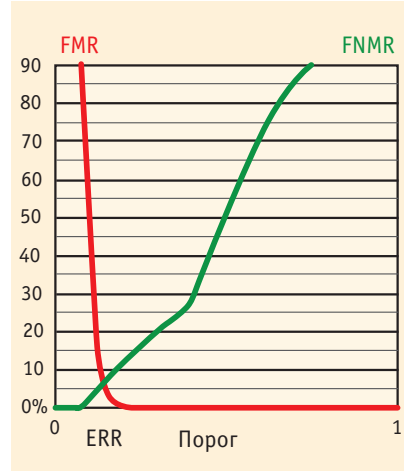


Рис. 2. Графики FMR и FNMR

выражаются или в абсолютных величинах (10-5), для параметра FAR это означает, что 1 человек из 100 тыс. будет несанкционированно допущен, в процентах данное значение будет (0,001%).

ВЛД – вероятность ложного допуска (FAR), (ГОСТ Р ИСО/МЭК19795-1-2007).

FRR (False Rejection Rate) – вероятность ложного задержания (ошибка второго рода), выраженное в процентах число отказов в допуске системой авторизованных лиц (имеется в виду верификация).

ВЛНД – вероятность ложного недопуска (FRR), (ГОСТ Р ИСО/МЭК19795-1-2007).

FMR (False Match Rate) – вероятность ложного совпадения параметров. Где-то мы это уже читали, см. FAR, но в данном случае один образец сравнивается со многими шаблонами, заложенными в базу данных, т.е. происходит идентификация.

ВЛС – вероятность ложного совпадения (FMR), (ГОСТ Р ИСО/МЭК19795-1-2007).

FNMR (False Non-Match Rate) – вероятность ложного несовпадения параметров, в данном случае один образец сравнивается со многими шаблонами, заложенными в базу данных, т.е. происходит идентификация.

ВЛНС – вероятность ложного несовпадения (FNMR), (ГОСТ Р ИСО/МЭК19795-1-2007).

Параметры (как и остальные перечисленные выше) взаимосвязаны *(рис. 1)*. Меняя порог FAR и FRR – «чувствительности» биометрической системы, мы одновременно изменяем их, выбирая требуемое соотношение. Действительно, можно так настроить биометрическую систему, что она с большой долей вероятности будет пропускать зарегистрированных пользователей, но и с достаточной долей вероятности будет пропускать и незарегистрированных пользователей. Поэтому данные параметры должны быть указаны одновременно для биометрической системы.

Если указывается только один параметр, то вас, как пользователя, это должно насторожить, поскольку таким образом очень легко завысить параметры в сравнении с конкурентом. Утрируя, можно сказать, что самый низкий коэффициент FAR будет иметь неработающая система, уж точно она никого несанкционированно не допустит.

Более или менее объективным параметром биометрической системы является коэффициент EER.

Коэффициент EER (равный уровень ошибок) – это коэффициент, при котором обе ошибки (ошибка приема и ошибка отклонения) эквивалентны. Чем ниже коэффициент EER, тем выше точность биометрической системы.

Для параметров FMR и FNMR строят аналогичный график *(рис. 2)*. Обратите внимание, что этот график всегда должен иметь привязку к объему базы данных (обычно числа выбирают с шагом 100, 1000, 10000 шаблонов и т.д.).

КОО – кривая компромиссного определения ошибки (англ. DET – detection error trade-off curve; DET curve). Модифицированная кривая рабочей характеристики, по осям которой отложены вероятности ошибки (ложноположительная – по оси X и ложноотрицательная – по оси У), (ГОСТ Р ИСО/МЭК19795-1-2007).

Кривую КОО (DET) используют для построения графика вероятностей ошибок сравнения (ВЛНС (FNMR) в зависимости от ВЛС (FMR)), вероятностей ошибок принятия решения (ВЛНД (FRR) в зависимости от ВЛД (FAR)) *(рис. 3-4)* и вероятностей идентификации на открытом множестве (ВЛОИ в зависимости от ВЛПИ), (ГОСТ Р ИСО/МЭК19795-1-2007).

Графики, отображающие качество работы биометрических систем, достаточно многочисленны, иногда создается впечатление, что их назначение – запутать доверчивого пользователя. Существуют еще РХ – кривая рабочей характеристики (англ.

Рис. 3. График DET

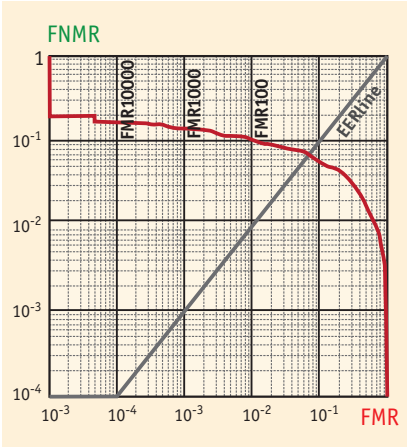
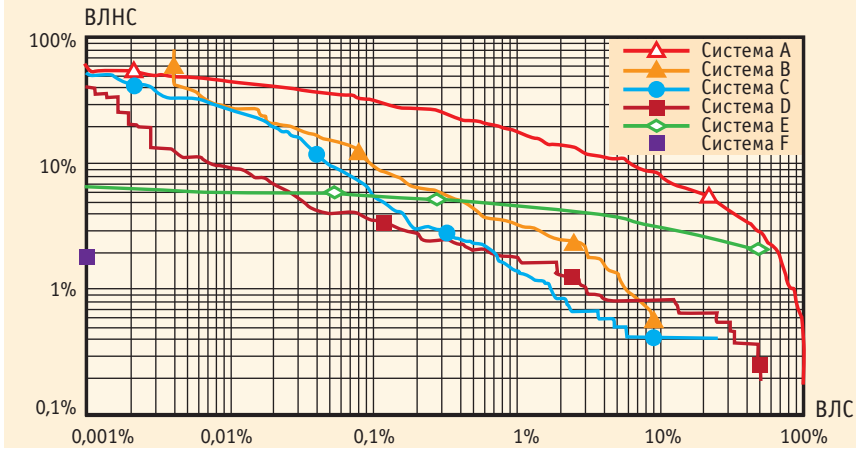


Рис. 4. Пример кривых КОО (ГОСТ Р ИСО/МЭК19795-1-2007)



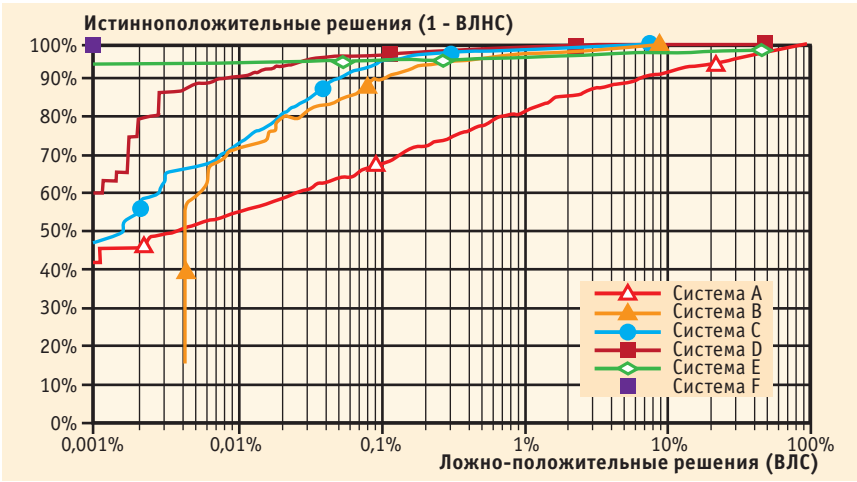


Рис. 5. Пример набора кривых РХ (ГОСТ Р ИСО/МЭК19795-1-2007)

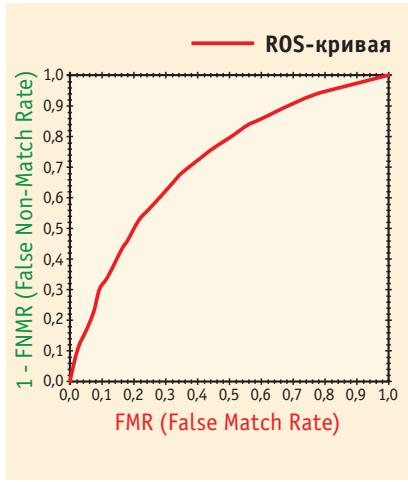


Рис. 6. Пример ROC-кривой

ROC – receiver operating characteristic curve) *(рис. 5-6)*, и, конечно, вы понимаете, что это далеко не последние кривые и зависимости, которые существуют в биометрии, но для ясности вопроса не будем на них останавливаться.

Кривые РХ (ROC) не зависят от порога, что позволяет проводить сравнение эксплуатационных характеристик различных биометрических систем, используемых в аналогичных условиях, или одной биометрической системы, используемой в различных условиях окружающей среды.

Кривые РХ (ROC) используют для изображения эксплуатационных характеристик алгоритма сравнения (1 – ВЛНС в зависимости от ВЛС), (1 – FNMR в зависимости от FMR), эксплуатационных характеристик биометрических систем верификации (1 – ВЛНД в зависимости от ВЛД), (1 – FRR в зависимости от FAR), а также эксплуатационных характеристик биометрических систем идентификации на открытом множестве (вероятность идентификации в зависимости от ВЛПИ).

Примечание: ВЛПИ – вероятность ложноположительной идентификации (англ. FPIR – false-positive identification-error rate), т.е. доля транзакций идентификации незарегистрированных в системе пользователей, в результате которых возвращается идентификатор (ГОСТ Р ИСО/МЭК19795-1-2007).

САМОЕ ГЛАВНОЕ ИЗ ПЕРЕЧИСЛЕННОГО

1) Параметры FAR (ВЛД), FRR (ВЛНД) и FMR (ВЛС) FNMR (ВЛРС) имеет смысл рассматривать только в совокупности.
2) Чем ниже коэффициент EER, тем выше точность биометрической системы.
3) Хорошим тоном для биометрической системы является наличие графиков DET (КОО) и ROC (РХ).

ГРАНИЦЫ ПАРАМЕТРОВ FAR И FRR БИОМЕТРИЧЕСКИХ СИСТЕМ

Теперь давайте прикинем, какие параметры FAR и FRR должны быть у биометрических систем. Обратимся за аналогией к требованиям для цифрового кодонаборника. Согласно ГОСТ число значимых десятичных цифр должно быть не менее 6, т.е. диапазон 0-999999, или 10^7 вариантов кода. Тогда вероятность FAR – 10^{-7}, а вероятность FRR определяется работоспособностью системы, т.е. стремится к нулю.

В банкоматах используется 4-разрядный десятичный код (что не соответствует ГОСТ), и тогда FAR будет составлять 10^{-5}. Возьмем FAR= 10^{-5} за определяющий параметр. Какое значение можно взять за приемлемое для FRR? Это зависит от задач биометрической системы, но нижняя граница должна находиться в диапазоне 10^{-2}, т.е. вас, как легального пользователя, система не допустит только один раз из ста попыток. Для систем с большой пропускной способностью, например, проходная завода, это значение должно быть 10^{-3}, иначе не понятно назначение биометрии, если мы не избавились от «человеческого» фактора.

Многие биометрические системы заявляют похожие и даже на порядок лучшие характеристики, но поскольку наши величины являются вероятностными, то необходимо указывать доверительный интервал этой величины. С этого момента производители биометрии предпочитают не вдаваться в подробности и не указывать данный параметр.

Если методика расчета, схема эксперимента и доверительный интервал не указаны, то по умолчанию подразумевается действие правила «тридцати», которое выдвинул J. F. Poter в работе «On the 30 error criterion» (1997).

Об этом же говорит и ГОСТ Р ИСО/МЭК19795-1-2007. В правиле «тридцати» утверждается, что для того, чтобы с доверительной вероятностью 90% истинная вероятность ошибки находилась в диапазоне ±30% от установленной вероятности ошибки, должно быть зарегистрировано не менее 30 ошибок. Например, если получены 30 ошибок ложного несоответствия в 3000 независимых испытаниях, можно с доверительной вероятностью 90% утверждать, что истинная вероятность ошибки находится в диапазоне от 0,7% до 1,3%. Правило следует непосредственно из биноминального распределения при независимых испытаниях и может применяться с учетом ожидаемых эксплуатационных характеристик для выполнения оценки.

После этого следует логичный вывод: чтобы получить величину ложного доступа в 10^{-5}, нужно провести $3x10^6$ опытов, что практически невозможно осуществить физически при реальном тестировании биометрической системы. Вот тут нас начинают мучить смутные сомнения.

Остается надеяться, что такое тестирование было проведено в лаборатории путем сравнения шаблонов вводимых биометрических признаков с шаблонами базы данных системы. Лабораторные испытания позволяют достаточно корректно оценить надежность заложенных алгоритмов обработки данных, но не реальную работу системы. Лабораторные испытания исключают такие воздействия на биометрическую систему, как электромагнитные наводки (актуально для всех систем биометрии), запыление или загрязнение контактных или дистанционных устройств считывания биометрического параметра, реальное поведение человека при взаимодействии с устройствами биометрии, недостаток или избыток освещения, периодическое изменение освещенности и т.д., да мало ли, что еще может повлиять на такую сложную систему, как система биометрии. Если бы человек мог заранее предугадать все негативно-действующие факторы, то можно было бы и не проводить натурные испытания.

Из опыта работы с другими охранными системами можем утверждать, что даже эксплуатация охранной системы в течение 45 суток не выявляет большинство скрытых проблем, и только опытная эксплуатация в течение 1-1,5 лет позволяет их устранить. У разработчиков существует даже термин – «детские болезни». Любая система должна ими переболеть.

Таким образом, кроме лабораторных испытаний необходимо проводить и натурные испытания, естественно, что оценки

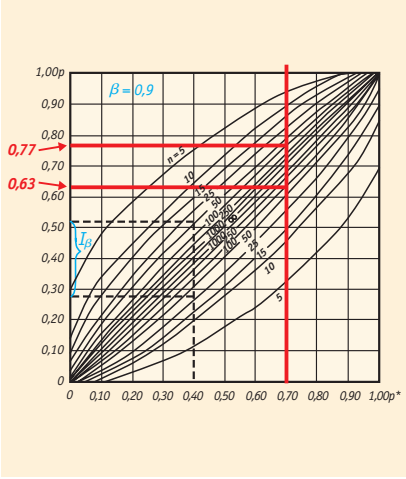


Рис. 7. Графическая зависимость доверительного интервала от количества проведенных опытов

доверительных интервалов при меньшем количестве опытов должны оцениваться по другим методикам.

Обратимся к учебнику Е.С. Вентцель «Теория вероятностей» (М.: «Наука», 1969. С. 334), который утверждает, если вероятность P очень велика или очень мала (что несомненно соответствует реальным результатам измерения вероятностей для биометрических систем), доверительный интервал строят, исходя не из приближенного, а из точного закона распределения частоты. Нетрудно убедиться, что это есть биномиальное распределение. Действительно, число появлений события A в *n*-опытах распределено по биномиальному закону: вероятность того, что событие A появится ровно *m* раз, равна

$$P_{m,n} = C_n^m \cdot p^m \cdot q^{n-m} \quad (14.5.11),$$

а частота p^* есть не что иное, как число появлений события, деленное на число опытов.

В данном труде приводится графическая зависимость доверительного интервала от количества проведенных опытов *(рис. 7)* для доверительной вероятности $\beta = 0{,}9$.

Рассмотрим пример. Мы провели 100 натурных опытов, из которых получили вероятность события равную *0,7*. Тогда по оси абсцисс откладываем значение частоты $p^* = 0{,}7$, проводим через эту точку прямую, параллельную оси ординат, и отмечаем точки пересечения прямой с парой кривых, соответствующих данному числу опытов n = 100; проекции этих точек на ось ординат и дадут границы $p^1 = 0{,}63$, $p^2 = 0{,}77$ доверительного интервала.

Для тех случаев, когда точность построения графического метода недостаточна, можно воспользоваться достаточно детальными табличными зависимостями *(рис. 8)* доверительного интервала, приведенными в труде И.В. Дунина-Барковского и Н.В. Смирнова «Теория вероятностей и математическая статистика в технике» (М.: Государственное издательство технико-теоретической литературы, 1955). В данной таблице *x*-числитель, *n*-знаменатель частости. Вероятности умножены на 1000.

Рассмотрим пример. Мы провели 204 натурных опытов, из которых событие произошло *4 раза*. Вероятность $P = 4/204 = 0{,}0196$, границы доверительного интервала $p^1 = 0{,}049$, $p^2 = 0{,}005$.

Теоретически подразумевается, что заявленные в документации параметры должны быть подтверждены сертификатами. Однако в России почти во всех областях жизни действует институт добровольной сертификации, поэтому сертифицируют на те требования, на которые хотят или могут получать сертификат.

Берем первый попавшийся сертификат на биометрическую систему, и видим 6 наименований ГОСТ, из которых ни один не содержит перечисленные выше параметры. Слава богу, что они хоть относятся к охранной технике и нормам безопасности. Это еще не самый худший вариант, приходилось встречать приемники и передатчики радиосистем передачи данных (РСПИ), сертифицированные как электрические машины.

САМОЕ ГЛАВНОЕ ИЗ ПЕРЕЧИСЛЕННОГО

1) Параметры FAR (ВЛД) должны быть не ниже 10^{-5}, а FRR (ВЛНД) должны находиться в диапазоне 10^{-2}-10^{-3}.
2) Не стоит безоговорочно доверять указанным в документации вероятностным параметрам, их можно воспринимать только как ориентир.
3) Кроме лабораторных испытаний необходимо проводить и натурные испытания биометрических систем.
4) Необходимо попытаться получить от разработчика, производителя, продавца как можно больше информации о реальных биометрических параметрах системы и методике их получения.
5) Не ленитесь расшифровывать, на какие ГОСТ(ы) и пункты ГОСТ(ов) сертифицирована биометрическая система.

В продолжение начатой темы о реальных системах биометрической идентификации предлагаем поговорить в статье «Основные биометрические системы».

Рис. 8. Фрагмент табличной зависимости доверительного интервала от количества проведенных опытов для доверительной вероятности $\beta = 0{,}95$

x \ n - x	30	35	40	45	50	60	80	100	200	500	∞
3	243	214	191	172	157	133	102	083	043	017	000
	019	**017**	**015**	**013**	**012**	**010**	**008**	**006**	**003**	**001**	**000**
4	275	242	217	196	179	152	118	096	049	020	000
	033	**029**	**025**	**023**	**021**	**017**	**013**	**011**	**005**	**002**	**000**
5	303	268	241	218	200	170	132	108	056	023	000
	048	**042**	**037**	**033**	**030**	**025**	**019**	**016**	**008**	**003**	**000**

ЛИТЕРАТУРА

1. *http://www.1zagran.ru*
2. *http://fingerprint.com.ua/*
3. *http://habrahabr.ru/post/174397/*
4. *http://sonda.ru/*
5. *http://eyelock.com/index.php/products/hbox*
6. *http://www.bmk.spb.ru/*
7. *http://www.avtelcom.ru/*
8. *http://www.nec.com/en/global/solutions/security/products/hybrid_finger.html*
9. *http://www.ria-stk.ru/mi «Мир измерений» 3/2014*
10. *http://www.biometria.sk/ru/principles-of-biometrics.html*
11. *http://www.biometrics.ru*
12. *http://www.guardinfo.ru/«Система физической защиты (СФЗ) ядерных материалов и ядерно-опасных объектов»*
13. *http://cbsrus.ru/*
14. *http: www.speechpro.ru*
15. *Poter J F. On the 30 error criterion. 1997.*
16. *ГОСТ Р ИСО/МЭК19795-1-2007. Автоматическая идентификация. Идентификация биометрическая. Эксплуатационные испытания и протоколы испытаний в биометрии. Часть 1. Принципы и структура.*
17. *Болл Р.М., Коннел Дж. Х., Ратха Н.К., Сеньор Э.У. Руководство по биометрии. М.: ЗАО «РИЦ Техносфера», 2006.*
18. *Симончик К.К., Белевитин Д.О., Матвеев Ю.Н., Дырмовский Д.В. Доступ к интернет-банкингу на основе бимодальной биометрии // Мир измерений. 2014. № 3.*
19. *Дунин-Барковский И.В., Смирнов Н.В. Теория вероятностей и математическая статистика в технике. М.: Государственное издательство технико-теоретической литературы, 1955.*