УДК 004.93”12

Сартасов Станислав Юрьевич

ОЦЕНКА НАДЁЖНОСТИ БИОМЕТРИЧЕСКИХ СИСТЕМ С НЕПРЕРЫВНОЙ КЛАССИФИКАЦИЕЙ И МНОЖЕСТВЕННОЙ РЕГИСТРАЦИЕЙ

Аннотация

Рассматривается теоретико-вероятностный подход к оценке качества биометрических систем, использующих биометрическое объединение уровня экземпляров и уровня принятия решения. Исследуется поведение систем, использующих непрерывный классификатор по стратегии «Лучшие K». Определяются качественные характеристики системы, и для каждой рассмотренной стратегии строятся теоретико-вероятностные модели качества. На основе численных экспериментов определяются границы применимости моделей.

Ключевые слова: биометрия, дактилоскопия, биометрические системы, биометрическое объединение, объединение уровня принятия решения, мультиэкземплярные системы.

ВВЕДЕНИЕ

Надёжность современных алгоритмов биометрической идентификации достигла очень высоких уровней. В частности, в литературе [5, 6] встречаются указания на коэффициент ложного доступа порядка 10^{-9} и коэффициент ложного отказа доступа порядка 10^{-6}. В то же время практика использования биометрических систем показала, что характеристики надёжности биометрической системы в целом значительно хуже, чем алгоритма идентификации. В этой связи предлагаются различные способы улучшения качества системы с помощью биометрической интеграции.

В литературе описаны различные способы биометрической интеграции [3]: путём совмещения изображений, путём объединения отличительных особенностей, путём комбинирования метрики или ранга различных классификаторов или на уровне принятия решения. Биометрическая интеграция может осуществляться как при решении задачи верификации (сопоставление 1 к 1), так и при решении задачи идентификации (сопоставление 1 ко многим). Большая часть представленных в литературе алгоритмов направлена на решение задачи верификации, и традиционным и хорошо себя зарекомендовавшим средством оценки качества алгоритма является кривая РХПУ (рабочей характеристики принимающего устройства) в координатах уровня ложного доступа и ложного отказа доступа (либо производных от них величинах). Такой подход оправдан в силу вероятностной природы биометрической идентификации [1]. Джейн и Прабакар в [4] показали, что одним из необходимых

условий для повышения качества верификации является низкий коэффициент корреляции итоговых оценок отдельных классификаторов. Между тем при решении задачи идентификации качество биометрической интеграции (например, цепочки или иерархии классификаторов при интеграции на уровне принятия решения) можно описать через надёжность отдельных составляющих, исходя из топологии системы. Такие модели оказываются полезными при разработке реальных идентификационных систем, поскольку позволяют оценить надёжность продукта до его введения в эксплуатацию, а также оценить принципиальную применимость той или иной схемы соединения уже имеющихся модулей.

В настоящей работе рассматриваются модели для одноуровневых идентификационных систем с одиночной и множественной регистрацией биометрического признака и двухуровневых идентификационных систем с непрерывным классификатором и одиночной и множественной регистрацией биометрического признака. Для всех рассмотренных моделей приведены формулы расчета основных коэффициентов, показаны результаты работы моделей для типовых значений коэффициентов, в качестве подведения итогов приводятся рекомендации по практическому применению различных способов биометрической интеграции.

В дальнейшем для наглядности мы будем использовать терминологию автоматизированных дактилоскопических идентификационных систем, однако полученные результаты применимы и для других источников биометрической информации, а также для мультимодальных систем.

ТЕРМИНЫ И ОПРЕДЕЛЕНИЯ

Алгоритм биометрической идентификации сопоставляет предоставленный отпечаток с отпечатком из базы данных, и результатом этого сопоставления является некоторая численная метрика. Когда она превышает некоторый порог, делается вывод, что два отпечатка принадлежат одному и тому же пальцу, в противном случае делается вывод об их принадлежности разным пальцам.

Идентификационная система может совершать ошибки двух типов:

1. Ситуация, когда злоумышленник, не зарегистрированный в базе, успешно сопоставлен с зарегистрированным пользователем – это ошибка ложного доступа, характеризующаяся коэффициентом ложного доступа (КЛД, False Acceptance Rate, *FAR*). Эта ошибка является критической для безопасности объекта, поэтому её стремятся минимизировать.

2. Ситуация, когда отпечаток зарегистрированного пользователя недостаточно хорошо сопоставлен с его зарегистрированными отпечатками – это ошибка ложного отказа доступа, характеризующаяся коэффициентом ложного отказа доступа (КЛОД, False Reject Rate, *FRR*). Эта ошибка описывает удобство пользования системой.

Этих двух коэффициентов достаточно для полного описания надёжности биометрического алгоритма. В то же время для оценки надёжности всей базы определение этих коэффициентов расширяется:

– $FAR(m)$ – вероятность ложного доступа злоумышленника в систему с m зарегистрированными отпечатками;

– $FRR(m)$ – вероятность ложного отказа доступа зарегистрированному пользователю в систему с m зарегистрированными отпечатками.

Мы также дополнительно определим коэффициент ошибочной идентификации (False Identification Rate, $FIR(m)$) как вероятность, что вместо одного зарегистрированного пользователя был распознан другой. Эта ошибка не нарушает безопасности объекта, так как некорректно распознанный человек обладает правом доступа к объекту, и не сказывается на удобстве пользования, однако может приводить к внутрисистемным ошибкам, в частности, в системах учёта времени или оповещения. Как мы покажем далее, это крайне маловероятная, но возможная ситуация.

Говоря о надёжности биометрических систем, используют также коэффициент правильного доступа (Genuine Acceptance Rate, GAR), который для алгоритма рассчитывается как $1 - FRR$. В биометрической системе $GAR(m)$, $FRR(m)$ и $FIR(m)$ связаны соотношением

$$GAR(m) + FRR(m) + FIR(m) = 1.$$

Коэффициент правильного отказа (Genuine Reject Rate, GRR) – это вероятность корректного отказа доступа злоумышленнику. Соотношения для случая системы и алгоритма схожи:

$$GRR + FAR = 1,$$
$$GRR(m) + FAR(m) = 1.$$

Далее по тексту мы будем обозначать общее число отпечатков в базе как m, число отпечатков на выходе непрерывного классификатора – k, а число зарегистрированных в базе представлений одного и того же отпечатка – n. В дальнейшем мы будем исходить из допущения, что n одинаково для всех зарегистрированных пальцев. Заметим, что количество зарегистрированных в базе пальцев, таким образом, составляет m/n.

Так как результаты каждого из сопоставлений не зависят от результата других, условная вероятность наступления того или иного события в системе определяется как произведение вероятностей наступления тех или иных событий для отдельных сопоставлений.

ОДНОУРОВНЕВАЯ СИСТЕМА С ЕДИНСТВЕННОЙ РЕГИСТРАЦИЕЙ

Случай биометрической системы, в которой каждому отпечатку соответствует один шаблон в базе, подробно рассмотрен в [1]. Кратко перечислим основные результаты.

Если злоумышленник пытается получить доступ к системе, то ему это удастся со следующей вероятностью:

$$FAR(m) = m(1 - FAR)^{m-1} FAR.$$

Так как отпечаток злоумышленника не зарегистрирован, любой из шаблонов в базе может привести к ложному доступу. Здесь и далее мы будем полагать, что система предоставляет доступ в том и только в том случае, когда все успешно сопоставленные шаблоны относятся к одному и тому же отпечатку, в данном случае успешно сопоставленный шаблон должен быть единственным.

Поскольку $(1 - FAR)^{m-1}$ стремятся поддерживать около 1, вероятность ложного доступа приближённо линейно возрастает с ростом числа записей в базе.

Теперь рассмотрим ситуацию, когда доступ к системе получает зарегистрированный пользователь. Правильный доступ будет осуществлён в случае успешного сопоставления с единственным отпечатком:

$$GAR(m) = (1 - FRR)(1 - FAR)^{m-1}.$$

А ошибочная идентификация произойдёт при удачном сопоставлении остальных $m - 1$ отпечатков и неудачном – единственного соответствующего:

$$FIR(m) = (m - 1)FRR \cdot FAR(1 - FAR)^{m-2}.$$

Отсюда вероятность ложного отказа – это вероятность всех остальных случаев:

$$FRR(m) = 1 - FIR(m) - GAR(m) = 1 - (1 - FAR)^{m-2}(1 - FAR - FRR + m \cdot FRR \cdot FAR).$$

ОДНОУРОВНЕВАЯ СИСТЕМА СО МНОЖЕСТВЕННОЙ РЕГИСТРАЦИЕЙ

Рассмотрим случай, когда в базе данных зарегистрировано n шаблонов для каждого зарегистрированного пальца, при этом общее число всех шаблонов m. Экспериментальное исследование биометрической интеграции на уровне комбинирования метрики для случая множественной регистрации при решении задачи верификации Джейн и Прабакар описали

в [4]. В то же время, множественная регистрация может быть рассмотрена для идентификационной системы на уровне принятия решения. В этом случае речь идёт о принятии решения на основе правила OR, то есть должен быть успешно сопоставлен хотя бы один шаблон из множества шаблонов, соответствующих данному отпечатку.

Чтобы успешно распознать зарегистрированного пользователя, необходимо идентифицировать как различные $m - n$ шаблонов в базе, не соответствующих зарегистрированному пользователю в базе, что соответствует вероятности $(1 - FAR)^{m-n}$. Кроме того, требуется, чтобы было успешно сопоставлено любое подмножество из n реально соответствующих шаблонов, что соответствует сумме $\sum_{i=1}^{n} c_i^n (1-FRR)^i FRR^{n-i}$ (рис. 1).

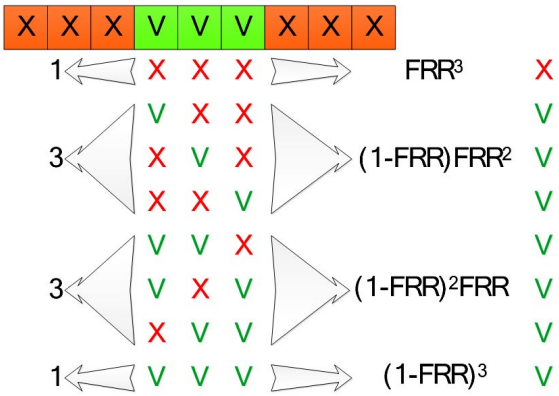

Рис. 1. Пример прямого подсчёта вероятностей для множественной регистрации из 3 шаблонов

Эту сумму можно упростить, заметив, что единственным случаем, когда при наших правилах система даст отрицательный ответ при сопоставлении соответствующих отпечатков, будет случай ложного отказа доступа во всех n случаях, вследствие чего нашу сумму можно переписать как $(1 - FRR^n)$. Таким образом, коэффициент правильного доступа будет равен

$$GAR(m, n) = (1 - FRR^n)(1 - FAR)^{m-n}.$$

Теперь найдём FAR для такой системы. Пусть для данного отпечатка злоумышленника успешно сопоставлено i отпечатков в базе. Если $i > n$, то система производит отказ доступа, так как хотя бы два шаблона будут относиться к разным отпечаткам. В случае, когда $i \le n$, из всех возможных C_i^m сочетаний шаблонов только $\frac{m}{n} C_i^n$ будет относиться к одному отпечатку. Отсюда

$$FAR(m,n) = \sum_{i=1}^{n} \frac{m}{n} C_i^n FAR^i (1 - FAR)^{m-i} = \frac{m}{n} \sum_{i=1}^{n} C_i^n FAR^i (1 - FAR)^{m-i}.$$

$FAR(m, n)$ линейно зависит от m, поэтому при разработке системы с множественной регистрацией приходится выбирать между безопасностью и числом пользователей. Также необходимо отметить, что при прочих равных условиях система с $n = 1$ будет априори реже пропускать злоумышленников, чем система с $n > 1$, хотя и незначительно. Это отчасти компенсируется более высоким уровнем GAR.

Поскольку ложная идентификация – это ложный отказ доступа для n корректных отпечатков при ложном доступе для хотя бы одного из $(m/n) - 1$ других зарегистрированных отпечатков, коэффициент ложной идентификации определяется как

$$FIR(m,n) = FRR^n \left(\frac{m}{n} - 1 \right) \sum_{i=1}^{n} C_i^n \, FAR^i \, (1 - FAR)^{m-n-i}.$$

БИОМЕТРИЧЕСКАЯ СИСТЕМА С НЕПРЕРЫВНЫМ КЛАССИФИКАТОРОМ И ЕДИНСТВЕННОЙ РЕГИСТРАЦИЕЙ

Методика непрерывной классификации была предложена Люминиа в [2]. Идея состоит в отказе от дискретной классификации, то есть разбиения всего множества отпечатков на непересекающиеся классы, в пользу вычисления некоторой метрики для всех отпечатков базы. Далее применяется одна из двух стратегий: либо выбирается k наиболее похожих

шаблонов, либо все шаблоны со значением метрики выше пороговой. Мы будем использовать первую стратегию.

Такой выбор позволяет рассматривать непрерывный классификатор как идентификационную систему, основанную на ранжировании. В [1] описан математический аппарат, который используется для описания таких систем. В частности, основной характеристикой является кривая суммарного сходства (Cumulative Match Curve, *CMC*) – накопленная сумма массы ранговой вероятности, то есть сумма вероятностей, что для входящего запроса правильная личность будет стоять на заданной позиции. Проще говоря, $CMC(k)$ – это вероятность, что в отсортированном по убыванию метрики массиве шаблонов в числе первых k шаблонов окажется соответствующий. В [1] показано, что для каждой базы в каждый момент времени $CMC(k)$ является постоянной. k выбирается таким образом, чтобы, с одной стороны, максимизировать $CMC(k)$, а с другой стороны, быть достаточно маленьким. Крайним случаем является $CMC(m) = 1$, но при этом модуль не фильтрует записи. Для $k < m$ всегда существует вероятность промаха классификатора, равная $1 - CMC(k)$.

Схема с цепочкой из непрерывного классификатора и мэтчера является частным случаем правила AND биометрической интеграции на уровне принятия решения для задачи идентификации. К априорным достоинствам такой схемы следует отнести потенциально лучшие временные характеристики в сравнении с одноуровневой системой за счёт оптимизации вычисления аргументов конъюнкции: не прошедшие классификатор шаблоны не участвуют в итоговом сопоставлении.

Когда к системе получает доступ злоумышленник, нам неважно, как отработает классификатор: записи злоумышленника в базе нет. Следовательно, в результате его работы база сокращается до k вариантов, что приводит к такой оценке вероятности ложного доступа:

$$FAR(k) = k(1 - FAR)^{k-1} FAR\,.$$

Найдём отношение $FAR(k)$ к $FAR(m)$:

$$\frac{FAR(k)}{FAR(m)} = \frac{k}{m(1 - FAR)^{m-k}} \approx \frac{k}{m}\,.$$

Таким образом, вероятность ложного доступа всегда снижается при $k << m$.

В системе с классификатором на вероятность корректного доступа будет также влиять и вероятность верной классификации:

$$GAR(k) = CMC(k)(1 - FRR)(1 - FAR)^{k-1}.$$

GAR и *FRR* влияют на удобство пользования системой: чем выше *GAR* и ниже *FRR* (при том же или меньшем *FIR*), тем удобнее система для пользователя. Аналогично с предыдущим случаем, с точки зрения *GAR*, система будет удобнее при условии:

$$\frac{GAR(k)}{GAR(m)} > 1,$$

$$\frac{CMC(k)(1 - FRR)(1 - FAR)^{k-1}}{(1 - FRR)(1 - FAR)^{m-1}} > 1,$$

$$CMC(k) > (1 - FAR)^{m-k}\,.$$

Ложная идентификация в системе с классификатором может произойти в двух случаях:

1) ошибка классификатора привела к тому, что в отфильтрованной базе нет нужной записи, и произошло ложное распознавание одного из отпечатков;

2) классификатор сработал корректно, но, вместо верного отпечатка, был сопоставлен другой.

Эти случаи описываются следующим соотношением:

$$FIR(k) = (1 - CMC(k))k(1 - FAR)^{k-1} FAR + CMC(k)(k - 1)FRR \cdot FAR(1 - FAR)^{k-2}\,.$$

Ложная идентификация в системе с классификатором будет лучше, если:

$$\frac{FIR(k)}{FIR(m)} < 1,$$

$$\frac{(1-CMC(k))k(1-FAR)^{k-1}FAR + CMC(k)(k-1)FRR \cdot FAR(1-FAR)^{k-2}}{(m-1)FRR \cdot FAR(1-FAR)^{m-2}} < 1,$$

$$\frac{(1-CMC(k))k(1-FAR)FAR + CMC(k)(k-1)FRR \cdot FAR}{(m-1)FRR \cdot FAR(1-FAR)^{m-k}} < 1,$$

$$CMC(k)(k-1)FRR \cdot FAR - CMC(k)k(1-FAR)FAR + k(1-FAR)FAR <$$

$$< (m-1)FRR \cdot FAR(1-FAR)^{m-k}.$$

Так как $(k-1)FRR - k(1-FAR) < 0$, при делении обеих частей неравенства знак меняется на противоположный:

$$CMC(k) > \frac{(m-1)FRR \cdot FAR(1-FAR)^{m-k} - k(1-FAR)FAR}{FAR((k-1)FRR - k(1-FAR))}.$$

Наконец, коэффициент ложного отказа доступа в системе с классификатором будет равен:

$$FRR(k) = 1 - CMC(k)(1-FRR)(1-FAR)^{k-1} -$$
$$-(1-CMC(k))k(1-FAR)^{k-1}FAR - CMC(k)(k-1)FRR \cdot FAR(1-FAR)^{k-2} =$$
$$= 1-(1-FAR)^{k-2}(CMC(k)(1-FRR)(1-FAR) +$$
$$+(1-CMC(k))k(1-FAR)FAR - CMC(k)(k-1)FRR \cdot FAR).$$

Система становится удобнее при $FRR(k) < FRR(m)$, отсюда:

$$1-(1-FAR)^{k-2}(CMC(k)(1-FRR)(1-FAR) + (1-CMC)(k))k(1-FAR)FAR +$$
$$+CMC(k)(k-1)FRR \cdot FAR) < 1-(1-FAR)^{m-2}(1-FAR-FRR+m \cdot FAR \cdot FRR),$$

$$CMC(k)(1-FRR)(1-FAR) + (1-CMC(k))k(1-FAR)FAR + CMC(k)(k-1)FRR \cdot FAR) >$$
$$> (1-FAR)^{m-k}(1-FAR-FRR+m \cdot FAR \cdot FRR),$$

$$CMC(k)((1-FRR)(1-FAR) - k(1-FAR)FAR + (k-1)FRR \cdot FAR) >$$
$$> (1-FAR)^{m-k}(1-FAR-FRR+m \cdot FAR \cdot FRR) - k(1-FAR)FAR,$$

$$CMC(k) > \frac{(1-FAR)^{m-k}(1-FAR-FRR+m \cdot FAR \cdot FRR) - k(1-FAR)FAR}{1-FRR-FAR+k \cdot FRR \cdot FAR - k(1-FAR)FAR}.$$

Так как неравенства совместны, для улучшения всех характеристик вероятность корректной классификации при заданном k должна быть выше максимального из полученных расчётных значений.

БИОМЕТРИЧЕСКАЯ СИСТЕМА С НЕПРЕРЫВНЫМ КЛАССИФИКАТОРОМ И МНОЖЕСТВЕННОЙ РЕГИСТРАЦИЕЙ

Необходимость во введении множественной регистрации может возникнуть, если непрерывный классификатор не обеспечивает достаточной точности классификации для приемлемого k. С точки зрения принятия решения, в рамках биометрической интеграции система идентификации сопоставляет предоставленный отпечаток с некоторым отпечатком в базе, согласно правилу, являющемуся дизъюнктивной нормальной формой из правил AND и OR, то есть идентификация будет успешной, если хотя бы один из шаблонов в базе, соответствующих некоторому отпечатку, будет сопоставлен и классификатором, и мэтчером.

Как и в предыдущем случае, на $FAR(k, n)$ не влияет работа классификатора, так как злоумышленника нет в базе. Между тем, так как в общем случае в k верхних результатах оказываются случайные шаблоны, это приводит к неравномерности представления m/n классов в итоговой выборке.

Как и в случае с одноуровневой системой, полная вероятность ложного доступа складывается из суммы вероятностей, что ложно распознаны $i = 1, 2, ..., n$ шаблонов, и все они принадлежат одному пальцу. Вероятность выборки из k шаблонов i, принадлежащих одному пальцу, такая же, как и для исходной базы в m отпечатков – $\frac{(m/n)C_i^n}{C_i^m}$. При этом i шаблонов может быть выбрано различными способами. Отсюда

$$FAR(k,n) = \frac{m}{n}\sum_{i=1}^{n}\frac{C_i^k C_i^n}{C_i^m} FAR^i (1-FAR)^{k-i}.$$

Вследствие ошибки классификатора в итоговую выборку может попасть любое число шаблонов, соответствующих зарегистрированному пользователю, от 0 до n. Чтобы найти GAR, необходимо сложить вероятности прохождения через классификатор i шаблонов, где $i \in [1, n]$, при этом хотя бы один из этих шаблонов должен быть распознан как правильный, а те, которые не принадлежат нужному отпечатку, – как не соответствующие. Кроме того, через классификатор i правильных шаблонов могут пройти любым из C_i^n способов.

Вероятность прохождения i данных шаблонов через классификатор можно оценить как $CMC(k)CMC(k-1)...CMC(k-i)(1-CMC(k-i-1))(1-CMC(k-i-2))...(1-CMC(k-n))$. В то же время n мало (в реальных системах крайне редко превышает 4), а функция $CMC(k)$ гладкая, и, как следствие, $CMC(k) \approx CMC(k-n)$. Поэтому грубо оценить вероятность прохождения i данных шаблонов можно как $CMC(k)^i(1-CMC(k))^{n-i}$. Это соотношение мы будем использовать далее. Отсюда по аналогии с предыдущими двумя случаями

$$GAR(k,n) = \sum_{i=1}^{n} C_i^n CMC(k)^i (1-CMC(k))^{n-i}(1-FRR^i)(1-FAR)^{k-i}.$$

По аналогии с предыдущим случаем, FIR состоит из суммы вероятностей двух ситуаций. В случае, когда ни один отпечаток не прошёл классификатор, из всех k отпечатков будет распознан один из $m/n - 1$ оставшихся пальцев. В случае, когда какие-то i из n правильных шаблонов прошли через классификатор, один из $m/n - 1$ отпечатков будет распознан из оставшихся $k - i$ шаблонов. Таким образом,

$$FIR(k,n) = (1-CMC(k)^n)\left(\frac{m}{n}-1\right)\sum_{i=1}^{n}\frac{C_i^k C_i^n}{C_i^{m-n}} FAR^j (1-FAR)^{k-i} + \\ + \sum_{i=1}^{n} c_i^n CMC(k)^i (1-CMC(k))^{n-i} FRR^i \left(\frac{m}{n}-1\right)\sum_{i=1}^{n}\frac{C_i^{k-i} C_i^n}{C_i^{m-n}} FAR^i (1-FAR)^{k-i-j}.$$

МОДЕЛИРОВАНИЕ ПОВЕДЕНИЯ СИСТЕМЫ

В таблице 1 приведены расчётные характеристики биометрических систем для всех четырёх схем организации при следующих параметрах: $m = 3000$, $k = 100$, $CMC(k) = 1 - 10^{-3}$. Рассматриваются два мэтчера: средней точности ($FAR = 10^{-6}$, $FRR = 10^{-3}$) и высокоточный ($FAR = 10^{-9}$, $FRR = 10^{-6}$). Для множественной регистрации использованы параметры $n = 2$ и $n = 4$.

Полученные результаты позволяют сделать следующие выводы:

1. Непрерывный классификатор всегда заметно снижает вероятность ложного доступа. Увеличение коэффициента ложного доступа при внедрении множественной регистрации пренебрежимо мало.

Табл.1. Характеристики различных схем биометрических систем при различных исходных характеристиках мэтчера

	FAR, %	*GAR*, %	*FIR*, %	*FRR*, %
Мэтчер средней точности (МСТ)	0.2991	99.60085	0.0003	0.39885
МСТ + Классификатор (К)	0.01	99.79022	0.00002	0.20976
МСТ, $n = 2$	0.2991	99.70055	0.0000003	0.29945
МСТ, $n = 4$	0.2991	99.70085	0.0000000000003	0.29915
МСТ + К, $n = 2$	0.009999	99.9898	0.00000001	0.0102
МСТ + К, $n = 4$	0.009999	99.9904	0.00000000000001	0.0096
Высокоточный мэтчер (ВМ)	0.0003	99.9996	0.0000000003	0.0004
ВМ + К	0.00001	99.89989	0.00000001	0.10011
ВМ, $n = 2$	0.0003	99.9997	$< 10^{-15}$	0.0003
ВМ, $n = 4$	0.0003	99.9997	$< 10^{-15}$	0.0003
ВМ + К, $n = 2$	0.00001	99.99989	0.00000000001	0.00011
ВМ + К, $n = 4$	0.00001	99.99999	$< 10^{-15}$	0.00001

2. Оба предлагаемых метода заметно снижают вероятность ошибочной идентификации. При их совместном использовании вероятность ошибочной идентификации становится пренебрежимо малой.

3. Если мэтчер значительно качественнее классификатора, двухуровневая схема может оказаться менее удобной (в терминах *FRR*), чем одноуровневая. При этом внедрение множественной регистрации сглаживает этот недостаток.

4. При введении множественной регистрации значительный прирост удобства наблюдается при $n = 2$. Увеличение n ведёт к дальнейшему увеличению удобства, но к существенно меньшему.

При разработке двухуровневых систем с непрерывным классификатором необходимо выбрать параметр k и соответствующий ему $CMC(k)$. Рассмотрим влияние этого параметра на характеристики качества в системе с высокоточным мэтчером с одиночной регистрацией. Для удобства мы будем рассматривать не конкретное значение k, а k', где $k' = k/m$. На рис. 2 представлена зависимость FAR системы от k'. Так как порядок FAR наиболее показателен, FAR отложен по логарифмической шкале. Порядок FAR системы уменьшается при стремлении k' к нулю.

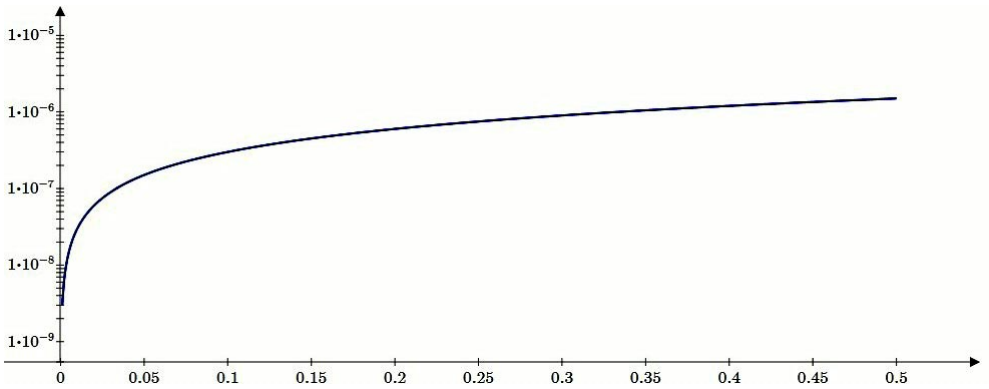

Рис. 2. Зависимость FAR от k'

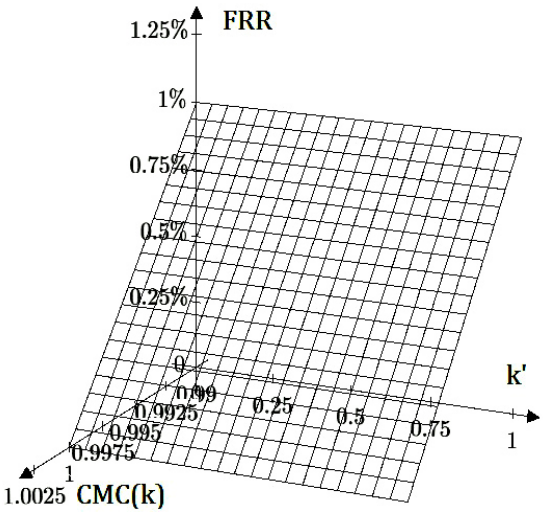

Рис. 3. Поведение *FRR* при различных значениях k' и $CMC(k)$

В случае с *FRR* (рис. 3) при уменьшении k' его снижение незначительно, однако коэффициент резко снижается при увеличении $CMC(k)$. Таким образом, надёжность двухуровневой идентификационной системы так же, как и надёжность одноуровневой, является выбором между удобством (низкий *FRR*) и безопасностью (низкий *FAR*).

ЗАКЛЮЧЕНИЕ

Полученные соотношения и их анализ позволяют описать широкий класс биометрических систем. Они не привязаны к конкретному биометрическому признаку, а потому могут быть использованы для оценки систем, построенных как на одном типе признака, так и на нескольких. Мы считаем, что теоретико-вероятностный подход к оценке надёжности биометрических идентификационных систем, особенно полученных с использованием биометрической интеграции, является не менее важным методом описания системы, наряду с её РХПУ, поскольку позволяет предсказывать максимально достижимые параметры *FAR* и *FRR* ещё на этапе проектирования системы.

Литература

1. *Bolle Ruud и др.* Guide to Biometrics. London: Springer, 2003.
2. *Luminia A., Maio D., Maltoni D.* Continuous versus exclusive classification for fingerprint retrieval // Pattern Recognition Letters, 1997. № 10.
3. *Maltoni D., Maio D., Jain A., Prabhakar S.* Handbook of Fingerprint Recognition. London: Springer, 2009.
4. *Prabhakar S, Jain A.K.* Decision-level fusion in fingerprint verification // Pattern Recognition, 2002. № 35.
5. *Сорокин К.* Биометрические системы: взгляд на рынок // Технологии защиты, 2010. № 5.
6. http://www.cscomm.com/access_control.html Access Control. Clearstream Communications, 2009 (дата обращения 25.08.2013).

QUALITY EVALUATION OF THE BIOMETRIC SYSTEMS IMPLEMENTING CONTINUOUS CLASSIFICATION AND MULTIPLE IMPRESSION STORAGE

Abstract

A probabilistic approach to the quality of multi-finger and decision level fusion-based biometric systems is described. The behavior of the systems implementing a «Top K» continuous classifier is investigated. System quality characteristics are defined and probabilistic quality models for each of the mentioned strategies are devised. The boundaries for the strategies are then determined with the numerical experiments.

Keywords: biometrics, fingerprinting, biometric systems, biometric fusion, decision-level fusion, multi-finger fusion.

Сартасов Станислав Юрьевич,
аспирант математико-механического факультета СПбГУ,
разработчик «Ланит-Терком»,
Stanislav.Sartasov@lanit-tercom.com.