

**Ахметов Б.С., Надеев Д.Н., Фунтиков В.А.,
Иванов А.И., Малыгин А.Ю.**

ОЦЕНКА РИСКОВ ВЫСОКОНАДЕЖНОЙ БИОМЕТРИИ

Монография

Алматы 2014

УДК 57.087.1

ББК 65.011

О93

Рецензенты:

д.т.н., профессор Тукаев У.А., д.т.н., профессор Джурунтаев Д.З., д.т.н., профессор Утепбергенов И.Т.

Ахметов Б.С., Надеев Д.Н., Иванов А.И., Фунтиков В.А., Малыгин А.Ю.
О93 Оценка рисков высоконадежной биометрии: монография. – Алматы: ТОО «Издательство LEM», 2014. – 108 с.

ISBN 978-601-239-333-0

В коллективной монографии сотрудников Казахского национального технического университета им. К.И. Сатпаева, Пензенского научно-исследовательского электротехнического института и Пензенского государственного университета излагаются вопросы статистических оценок рисков информационной безопасности средств высоконадежной биометрико-нейросетевой аутентификации.

Показано, что используемые зарубежными авторами классический биномиальный закон распределения независимых данных при оценке рисков информационной безопасности не позволяет учитывать естественные корреляционные связи битов выходного кода преобразователя биометрия-код. Для учета коррелированности битов выходного кода преобразователя в работе предложен неклассический биномиальный закон распределения зависимых данных. На основе результатов численного моделирования синтезированы таблицы плотности распределения значений меры Хэмминга кодов-откликов «Чужой» обученного нейросетевого преобразователя биометрия-код. Установлено, что для качественной оценки рисков необходимо переходить от расчета основных параметров статистической модели к наблюдению формы закона распределения, при условии наличия соответствующих эталонов формы. В качестве таких эталонов формы закона распределения служат синтезированные в работе таблицы биномиального зависимого распределения.

Монография подготовлена в рамках выполнения комплексного проекта по теме: «Разработка и подготовка производства телекоммуникационного оборудования, разработка программного сетевого, прикладного и специального обеспечения для создания цифровых сетей связи с персонализированным доступом» в соответствии с Постановлением Правительства РФ №218 от 09.04.2010 г. (Договор № 13G25.31.0039 от 07.09.2010 г.) и научного проекта по теме: №753. МОН.ГФ.13.13 «Исследование вариантов реализации и разработка действующего лабораторного образца ON-LINE системы биометрического обезличивания электронных историй болезней для медицинского учреждения».

Для специалистов по искусственным нейронным сетям, защите информации в системах медицинского электронного документооборота, созданию электронных карт и паспортов с автоматической биометрической аутентификации личности.

УДК 57.087.1

ББК 65.011

***Печатается по решению Ученого Совета КазНТУ
имени К.И.Сатпаева***

ISBN 978-601-239-333-0

© Ахметов Б.С., Надеев Д.Н., Иванов А.И.,

© Фунтиков В.А., Малыгин А.Ю.

© КазНТУ имени К.И.Сатпаева, 2014

СОДЕРЖАНИЕ

Введение	4
1 Модификация схемы Бернулли при описании схемы испытаний с зависимыми исходами	11
2 Тестирование нейросетевого преобразователя биометрия-код с корреляционными связями между битами выходного кода.....	20
3 Моделирование биномиального зависимого закона распределения меры Хэмминга кодов-откликов «Чужой» преобразователя биометрия-код.....	28
4 Синтез таблиц биномиального зависимого закона распределения меры Хэмминга кодов-откликов «Чужой» преобразователя биометрия-код.....	47
5 Оценка рисков ложной аутентификации «Чужого» через среднеквадратическое отклонение нормированной меры Хэмминга и математическое ожидание модуля коэффициентов парной корреляции	57
6 Использование таблиц биномиального зависимого закона распределения при измерении энтропии высоких порядков	65
Заключение.....	87
Список использованных источников	89
Приложение А.....	98

ВВЕДЕНИЕ

В настоящее время во всем мире идет активный переход к информационному обществу. С каждым годом происходит рост потоков обмена информацией и процедур идентификации и аутентификации личности как локально, так и в открытом информационном пространстве, например, сети Internet. Наблюдается тенденция увеличения длины ключей аутентификации с целью обеспечения высокого уровня безопасности личной информации пользователя. В то же время используемые в настоящее время протоколы аутентификации по паролю и личному ключу провоцируют атаки, как на самого пользователя, так и его персональный носитель ключа [1-7].

Для того, чтобы обеспечить высоконадежную дистанционную идентификацию человека, необходимо привлекать биометрические технологии, способные безопасно взаимодействовать с криптографическими механизмами [1, 8-12].

В настоящее время практически все страны, имеющие значимый национальный научно-технический потенциал, пытаются решать эти задачи.

Ученые России, Казахстана и Белоруссии являются лидерами по разработке и использованию технологии биометрико-нейросетевой защиты информации [1, 8-12].

США, Канада и страны Евросоюза идут по пути использования нечеткой математики, ученые этих стран предлагают мировому сообществу специализированные «fuzzy» обогатители (экстракторы) превращающие бедную неоднозначную размытую биометрическую информацию в сильный личный ключ пользователя [13-21]. Особую важность биометрические технологии идентификации получают в связи с введением в ближайшем будущем нового поколения загранпаспортов и визовых документов. С этой целью специальный подкомитет ISO/IEC JTC1 SC37 (Biometrics) [22] принял более 90 международных стандартов, которые касаются различных особенностей биометрии. Они описывают, как нужно снимать, хранить, обрабатывать: отпечатки пальцев, изо-

бражения лица, рукописные образы. Двадцать международных стандартов посвящено вопросам тестирования биометрических устройств и технологий [22], в настоящее время более пятидесяти стандартов находятся в стадии разработки.

Россией в 2004 г. впервые в мире предложено использование полностью автоматической локальной и дистанционной биометрико-нейросетевой аутентификации пользователя в открытом информационном пространстве. Ее принцип основан на использовании нейросетевых преобразователей биометрия-код, построенных на основе больших и сверхбольших искусственных нейронных сетей. Преобразователь биометрия-код является частью средства высоконадежной биометрико-нейросетевой аутентификации пользователя. Правильно построенный преобразователь биометрия-код ведет себя как классическая необратимая хэш-функция. Случайные входные биометрические образы множества «Все чужие» нейросетевой преобразователь биометрия-код перемешивает (хэширует), а заранее известное множество нечетких образов «Свой» преобразователь свертывает в единственное значение личного криптографического ключа. При этом достаточно сложная нейронная сеть с 256 выходами позволяет обеспечивать стойкость к атакам случайного подбора на уровне 10^{22} (22 степень) попыток. Такие преобразователи позволяют отказаться от хранения паролей и личных ключей. Формирование личного кода аутентификации происходит на основе нечетких биометрических образов пользователя в соответствие с российским базовым стандартом [11].

Сдерживающим фактором применения нейросетевых преобразователей биометрия-код, использующих искусственный интеллект высокой размерности, было отсутствие быстрых алгоритмов автоматического обучения и тестирования преобразователей. Данная научно-техническая проблема была решена в научной школе Пензенского государственного университета под руководством д.т.н., профессора Волчихина В.И. д.т.н. Ивановым А.И., д.т.н. Малыгиным А.Ю., к.т.н. Фунтиковым В.А. в период с 1998 по 2012 годы [1, 8].

Дальнейшим положительным импульсом для ускорения исследований в данном направлении стал подписанный 22 февраля

2012 г. Меморандум о взаимопонимании и взаимовыгодном сотрудничестве между одним из ведущих технических вузов Казахстана – Казахским национальным техническим университетом им. К.И. Сатпаева (КазНТУ им. К.И. Сатпаева) и Пензенским государственным университетом (ПГУ). Для исследований в области биометрии в КазНТУ им. К.И. Сатпаева на базе Института информационных и телекоммуникационных технологий была создана «Межотраслевая научно-исследовательская лаборатория по созданию и тестирования средств для биометрии». Научное направление возглавляет директор института, д.т.н., профессор Ахметов Б.С.

Прошедшие годы совместных исследований дали определенный результат: опубликован ряд научных статей в высокорейтинговых рецензируемых журналах и аннотированных в базе цитирования SCOPUS, опубликованы две коллективные монографии. Ученые КазНТУ имени К.И. Сатпаева и ПГУ участвовали в проведении НИР и НИОКТР по данной тематике.

Как показали проведенные совместные исследования, использование нейросетевых преобразователей биометрия-код в ответственных приложениях безопасности требует проводить оценку рисков нарушения информационной безопасности при обработке информации в средствах высоконадежной биометрико-нейросетевой аутентификации пользователя.

Вероятность ложной аутентификации «Чужого» является основным показателем защищенности личной информации при выборе парольного слова для конкретного пользователя администратором безопасности защищенной автоматизированной системы. Основной проблемой при оценке рисков нарушения информационной безопасности процессов обработки информации преобразователем биометрия-код и определении вероятности ложной аутентификации «Чужого» является отсутствие модели статистического описания выходных состояний обученного нейросетевого преобразователя биометрия-код.

Зарубежные исследователи в этом случае в качестве модели статистического описания выходных состояний преобразователя биометрия-код используют гипотезу классического биномиаль-

ного закона распределения независимых данных. Однако, как показали проведенные исследования, при этом не учитывается то, что биометрические данные обладают корреляционными связями. Причина заключается в том, что человек не может воспроизводить личные биометрические образы совершенно случайно. Общими для большинства биометрических образов является, например, направление письма при воспроизведении рукописного парольного слова, спектральные полосы букв при произнесении голосовой парольной фразы и т.д. Это обстоятельство препятствует использованию принятого в работах зарубежных исследователей классического биномиального закона распределения независимых данных в качестве модели статистического описания выходных состояний обученного нейросетевого преобразователя биометрия-код. Корреляционные связи входных биометрических данных преобразователя биометрия-код являются причиной коррелированности разрядов выходного кода преобразователя. Игнорирование свойства коррелированности разрядов выходного кода преобразователя приводит к тому, что результаты оценки рисков нарушения информационной безопасности оказываются неоправданно оптимистичными [23-26]. По этой причине разрабатываемые на основе данных результатов методики противодействия угрозам нарушения информационной безопасности не приводят к желаемому эффекту. Прогнозы вероятности ложной аутентификации «Чужого» оказываются сильно заниженными по сравнению с результатами, полученными при прямом тестировании на статических базах биометрических образов «Все чужие» больших и сверхбольшого объема [27-32]. В этой ситуации возникает необходимость в разработке такой модели статистического описания преобразователя биометрия-код, которая бы учитывала свойство коррелированности разрядов выходного кода преобразователя.

В данной работе авторами предложена модель статистического описания выходных состояний обученного нейросетевого преобразователя биометрия-код при прогнозировании вероятности ложной аутентификации «Чужого», учитывающая свойства коррелированности разрядов выходного кода преобразователя [32-36]. С этой целью разработана численная модель преобразо-

вателя биометрия-код. На основе этой модели получен закон распределения коррелированных выходных данных преобразователя биометрия-код – биномиальный закон распределения зависимых данных (биномиальный зависимый закон распределения). При этом замечено, что при увеличении числа входов и выходов преобразователя и использовании гипотезы независимости биометрических данных происходит накопление ошибки оценки вероятности ложной аутентификации «Чужого» с ростом корреляции битов выходного кода по сравнению с гипотезой независимости биометрических данных. При корреляции равной 0,5 топологическое независимое от размерности преобразователя превращение биномиального зависимого распределения в равномерное распределение значений.

По мере увеличения корреляции при использовании биномиального зависимого закона распределения риски нарушения информационной безопасности самых вероятных источников опасности становятся минимальными по сравнению со штатной ситуацией оценки рисков. Наоборот, риски нарушения информационной безопасности самых невероятных источников опасности, становятся максимальными по сравнению со штатной ситуацией, когда они являются наименее важными. Становится обманчивой гипотеза о независимости данных при оценке рисков, она заменяется доказательством существенной коррелированности и сильной зависимости от корреляции результатов оценки рисков.

По результатам численного моделирования синтезированы таблицы биномиального закона распределения зависимых данных.

Они задают значения функции $\alpha_n \left(m(r) \cdot m\left(\frac{H}{n}\right) \right)$, позволяющей

перейти от выражения классического биномиального распределения к неклассическому биномиальному зависимому распределению. Эта функция определяет топологические преобразования классического биномиального распределения при введении корреляционных связей в схему эксперимента. Значения функции

$\alpha_n \left(m(r) \cdot m\left(\frac{H}{n}\right) \right)$ задаются численно в виде таблиц биномиаль-

ного зависимого распределения в интересующих нас точках по корреляции.

Получено аналитическое описание оценки рисков ложной аутентификации «Чужого» как функции от качества принимаемых преобразователем решений. Предложенное аналитическое описание позволило снизить ошибки прогнозирования вероятности ложной аутентификации «Чужого» с 8% до 0,05% от 256 бит выходного кодов преобразователя.

Предложенная модель статистического описания преобразователя биометрия-код используется при оценке вероятности ложной аутентификации «Чужого» преобразователем биометрия-код. Оценка вероятности ложной аутентификации «Чужого» проводится через среднеквадратическое отклонение нормированной меры Хэмминга и математическое ожидание коэффициентов парной корреляции при условии знания кода «Свой».

Если код «Свой» неизвестен, то вероятность ложной аутентификации «Чужого» определяется с помощью высокоразмерного наблюдателя энтропии преобразователя биометрия-код с 256 выходами на базе 10 000 образов [35].

Высокоразмерный наблюдатель энтропии настраивается путем усреднения энтропий, полученных по нескольким образам «Чужой» и множеству «Все Чужие» по параметрам среднеквадратического отклонения нормированной меры Хэмминга и математическому ожиданию коэффициентов парной корреляции. Проверить эффективность работы наблюдателя энтропии можно на примере вычисления энтропии кодов букв русского и английского языка на базе 100 000 примеров кодов.

Пути понижения размерности требуемой базы образов «Все Чужие» при энтропийной оценке рисков являются выбор кодов со статистической симметрией (псевдо ортогональностью) и переход от двух статистических моментов к эталонам формы закона распределения. Пытаясь понизить размеры базы «Все Чужие» при энтропийной оценке рисков, получаем ситуацию, когда требуется учитывать не два статистических момента, а форму закона распределения [37, 38]. Чтобы проводить качественную оценку рисков с помощью высокоразмерного наблю-

дателя энтропии, необходимы эталоны формы закона распределения.

Оценивать риски информационной безопасности предлагается путем создания высокоразмерных наблюдателей энтропии и настройке их по эталонам формы закона распределения. В качестве таких эталонов формы закона распределения служат синтезированные в работе таблицы биномиального зависимого

распределения, дающие значения функции $\alpha_n \left(m(|r|) \cdot m\left(\frac{H}{n}\right) \right)$

в интересующих нас точках. В общем случае эта функция может быть восстановлена при любых значениях длины выходного кода преобразователя. Для исследований интересна функция

$\alpha_{256} \left(m(|r|) \cdot m\left(\frac{H}{n}\right) \right)$ для 256-выходного преобразователя исходя из

требований стандартов [11, 39] и необходимой длине кода аутентификации при использовании в приложения безопасности.

При синтезе высокоразмерных наблюдателей энтропии функ-

цию $\alpha_n \left(m(|r|) \cdot m\left(\frac{H}{n}\right) \right)$ необходимо разбивать на диапазоны по

корреляции. В середине каждого диапазона значение энтропии, получаемое с помощью высокоразмерного наблюдателя энтро-

пии должно совпадать со значением функции $\alpha_n \left(m(|r|) \cdot m\left(\frac{H}{n}\right) \right)$.

В окрестности этой точки наблюдатель дает оценку энтропии с погрешностью определяемой выбором диапазонов разбиения [40].

1 МОДИФИКАЦИЯ СХЕМЫ БЕРНУЛЛИ ПРИ ОПИСАНИИ СХЕМЫ ИСПЫТАНИЙ С ЗАВИСИМЫМИ ИСХОДАМИ

При оценке рисков нейросетевого преобразователя биометрия-код приходится решать задачу идентификации закона распределения его выходных состояний. Требуется разработать математическую модель процесса тестирования заранее обученного преобразователя, когда на его входы подаются случайные образы и контролируются состояния его выходов. В данном разделе дано описание схемы испытаний используемой при описании схемы испытаний с зависимыми исходами. Процесс тестирования обученного преобразователя можно представить как некоторую схему испытаний, при которой пошаговое повторение опытов заменяется контролированием состояний выходов преобразователя. Один и тот же опыт или аналогичные опыты (считывание состояния выходов обученного преобразователя) повторяются неоднократно. Если преобразователь имеет 256 выходов, то происходит повторение или серия из 256 опытов или одно считывание состояния 256 выходов преобразователя.

В результате каждого опыта (на каждом выходе преобразователя) может появиться или не появиться некоторое событие (например, состояние «0» на выходе преобразователя), причем интересен не результат каждого отдельного опыта, а общее число появлений события в результате серии опытов (вероятность появления состояния «0» на 256 выходах преобразователя). Например, если на входы преобразователя подаются случайные биометрические образы и контролируются состояния «0» на его выходах, то нас в данной ситуации интересует не состояние каждого конкретного выхода, а вероятность появления состояния «0» на 256 выходах преобразователя. В этом случае требуется определять вероятность любого заданного числа появлений события (состояния «0») в результате серии опытов (на 256 выходах преобразователя). При независимых испытаниях схемой испытаний служит классическая схема Бернулли. Эта

схема описывает проведение испытаний с независимыми исходами случайной величины. При этом несколько опытов называются независимыми, если вероятность того или иного исхода каждого опыта не зависит от того, какие исходы имели другие опыты. Например, несколько последовательных подбрасываний монеты представляют собой независимые опыты. Несколько последовательных извлечений карты из колоды представляют собой независимые опыты при условии, что вынутая карта каждый раз возвращается в колоду и карты перемешиваются; в противном случае это зависимые опыты. Несколько выстрелов представляют собой независимые опыты только в случае, если прицеливание производится заново перед каждым выстрелом; в случае, когда прицеливание производится один раз перед всей стрельбой или непрерывно осуществляется в процессе стрельбы (стрельба очередью, бомбометание серией), выстрелы представляют собой зависимые опыты. В нашем случае 256 битов выходного кода преобразователя биометрия-код являются независимыми, если корреляция между ними равна нулю.

Примером проведения испытаний по схеме Бернулли является подбрасывание одной монеты с заданной вероятностью появления решки или орла. При тестировании преобразователя биометрия-код происходит контроль его выходных кодов-откликов с заданной вероятностью появления состояния «0» в одном бите кода. Иллюстрация проведения испытаний по классической схеме Бернулли приведена на рисунке 1.



Рисунок 1 – Классическая схема Бернулли

Испытания с независимыми исходами случайной величины могут производиться в одинаковых или различных условиях. В первом случае вероятность события во всех опытах одна и та же. В этом случае вероятность появления k событий из n возможных

событий при заданной вероятности p этого события вычисляется следующим образом:

$$D_k(n, \delta) = \frac{n!}{k!(n-k)!} \delta^k (1-\delta)^{n-k}. \quad (1)$$

Во втором случае вероятность события от опыта к опыту меняется. В этом случае используется общее выражение схемы испытаний, из которого можно определить вероятность появления k событий из n возможных событий при заданной вероятности p этого события:

$$\sum_{k=0}^n P_k(n, p) z^k = \prod_{i=1}^n (q_i + p_i z), \quad (2)$$

где z используется для перебора всех вариантов сочетаний k из n возможных событий с вероятностями появления события p_i , i – номер опыта; $q_i = 1 - p_i$.

Выражения (1) и (2) могут использоваться при вычислении вероятности ложной аутентификации «Чужого» P_2 преобразователем биометрия-код. Согласно выражению (1) вероятность ложной аутентификации «Чужого» P_2 преобразователем биометрия-код при заданной вероятности p появления «0» в бите выходного кода вычисляется следующим образом:

$$D_2 = \delta^n. \quad (3)$$

Согласно выражению (3) если вероятность $p=0,5$, то оценка вероятности ложной аутентификации «Чужого» преобразователем с 256 выходами составляет $D_2 = 2^{-256} \approx 10^{-77}$. Однако согласно [36] вероятность p появления «0» в бите выходного кода равна 0,5 только для идеального преобразователя, биты выходного кода которого являются некоррелированными. При тестировании реального преобразователя биометрия-код случайными биометрическими образами «Все Чужие», биты выходного кода являются

коррелированными и расчеты по выражению (3) дают неоправданно оптимистические прогнозы вероятности ложной аутентификации «Чужого».

В связи с этим в качестве модели схемы испытаний при тестировании преобразователя биометрия-код должна служить модификация классической схемы Бернулли – схема испытаний с зависимыми исходами случайной величины, которая учитывает изменение в распределении вероятностей при введении зависимостей в серию опытов. По аналогии с классической схемой Бернулли в результате каждого опыта (на каждом выходе преобразователя) может появиться или не появиться некоторое событие (например, состояние «0» на выходе преобразователя), причем нас интересует не результат каждого отдельного опыта, а общее число появлений события в результате серии опытов (вероятность появления состояния «0» на 256 выходах преобразователя).

Например, на входы преобразователя подаются случайные биометрические образы и контролируются состояния «0» на его выходах. В этом случае требуется определять вероятность любого заданного числа появлений события (состояния «0») в результате серии опытов (на 256 выходах преобразователя). При использовании модифицированной схемы Бернулли вероятность того или иного исхода каждого опыта зависит от того, какие исходы имели другие опыты. Например, подбрасывания сразу нескольких монет одновременно одной рукой дают зависимые опыты для каждой монеты. В нашем случае 256 битов выходного кода преобразователя биометрия-код являются зависимыми, если корреляция между ними отлична от нуля. Иллюстрация проведения испытаний по модифицированной схеме Бернулли, учитывающей появление корреляционных связей, приведена на рисунке 2.

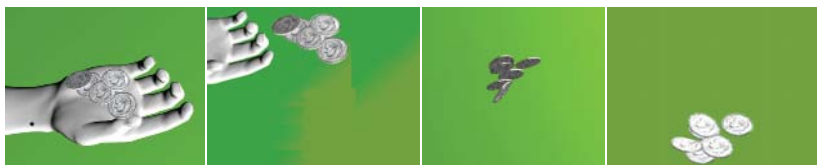


Рисунок 2 – Модификация классической схемы Бернулли, учитывающая появление корреляционных связей

Испытания с зависимыми исходами случайной величины (зависимые опыты) могут производиться в одинаковых или различных условиях.

В первом случае вероятность события во всех опытах одна и та же. Этот случай является наиболее интересным для ситуации тестирования преобразователя биометрия-код.

Во втором случае вероятность события от опыта к опыту меняется. Этот вариант является обобщением модифицированной схемы Бернулли и может быть получен из первого изменением параметров статистической модели.

Модификация схемы Бернулли дает вероятность появления k событий из n возможных событий при заданной вероятности p этого события вычисляется следующим образом:

$$P_k(n, p) = \left[\frac{n!}{k!(n-k)!} \right]^{\alpha_n \left(m(r), m\left(\frac{H}{n}\right) \right)} [p^k]^{\alpha_n \left(m(r), m\left(\frac{H}{n}\right) \right)} [(1-p)^{n-k}]^{\alpha_n \left(m(r), m\left(\frac{H}{n}\right) \right)}, \quad (4)$$

где $\alpha_n \left(m(r), m\left(\frac{H}{n}\right) \right)$ – функция перехода от классического биномиального распределения к неклассическому биномиальному зависимому распределению.

Эта функция определяет топологические преобразования классического биномиального распределения при введении корреляционных связей в схему эксперимента. Значения функции

$\alpha_n \left(m(r), m\left(\frac{H}{n}\right) \right)$ в выражении (4) задаются численно в виде

таблиц биномиального зависимого распределения в интересующих нас точках по корреляции. Основной задачей является вос-

становление функций $\alpha_n \left(m(r), m\left(\frac{H}{n}\right) \right)$ в интересующих нас

точках по корреляции. В общем случае эта функция может быть восстановлена при любых значениях длины выходного кода пре-

образователя. Нас интересует функция $\alpha_{256} \left(m(r), m\left(\frac{H}{n}\right) \right)$ для

256-выходного преобразователя, исходя из требований стандартов [11, 39] и необходимой длине кода аутентификации при использовании в приложения безопасности.

В связи с тем, что одним из основных признаков не идеальности нейросетевого преобразователя являются корреляционные связи между битами выходного кода преобразователя, необходимо вычислять величину корреляции. Для этой цели необходимо контролировать парную корреляцию выходов преобразователя при подаче на его входы случайных биометрических образов «Все Чужие». При измерениях необходимо использовать не менее [36] 300 случайных образов и проводить измерение не менее чем на 100 случайных парах i, j выходов. Расчет корреляционной связи по каждой из пар осуществляется по традиционной формуле:

$$r_{i,j} = \frac{1}{N} \sum_{k=1}^N \frac{(m(y_i) - y_{i,k}) \cdot (m(y_j) - y_{j,k})}{\sigma_i \cdot \sigma_j}, \quad (5)$$

где N – число независимых экспериментов; i, j – номера контролируемых выходных разрядов; y_i, y_j – значения контролируемых разрядов, полученные при тестировании; $m(\cdot)$ – символ операции вычисления математического ожидания; σ_i, σ_j – среднеквадратические отклонения состояний контролируемых разрядов.

Проведенные статистические исследования реальных биометрико-нейросетевых преобразователей с большим числом выходов показали, что наиболее вероятным значением коэффициентов корреляции являются нулевые значения [41]. Тем не менее, достаточно часто встречаются сравнительно сильно коррелированные выходные данные. На сегодняшний день считается приемлемым иметь среднее значение модулей коэффициентов корреляции менее 0.15 [11]. Пример распределения коэффициентов парной корреляции преобразователя биометрия-код, вычисленных по выражению (5), приведен на рисунке 3.

Проблема не идеальности нейросетевых преобразователей пока не имеет однозначного решения. Пока не ясно, какова причина появления сравнительно сильных корреляционных связей. Вполне возможно, что это следствие присутствия в рукописном

почерке некоторых корреляционных связей. Мы все пишем слева направо (кириллическая рукописная традиция), кроме того, некоторые знаки чаще встречаются при воспроизведении слов русского и любого иного языка. Это означает, что примеры кривых колебаний пера по оси Y при воспроизведении случайно выбранных слов, отображенные на рисунке 4, нельзя считать белым шумом.

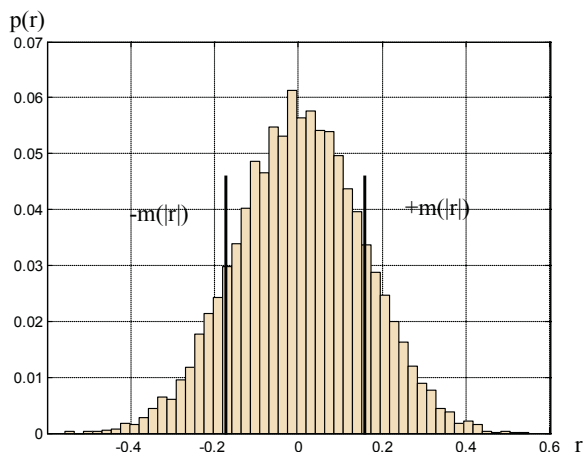


Рисунок 3 – Пример распределения коэффициентов парной корреляции преобразователя биометрия-код

Человек не может воспроизводить личные биометрические образы совершенно случайно. Общими для большинства биометрических образов является, например, направление письма при воспроизведении рукописного парольного слова, спектральные полосы букв при произнесении голосовой парольной фразы и т.д. Это обстоятельство препятствует использованию принятого в работах зарубежных исследователей классического биномиального закона распределение независимых данных в качестве модели статистического описания выходных состояний обученного нейросетевого преобразователя биометрия-код.

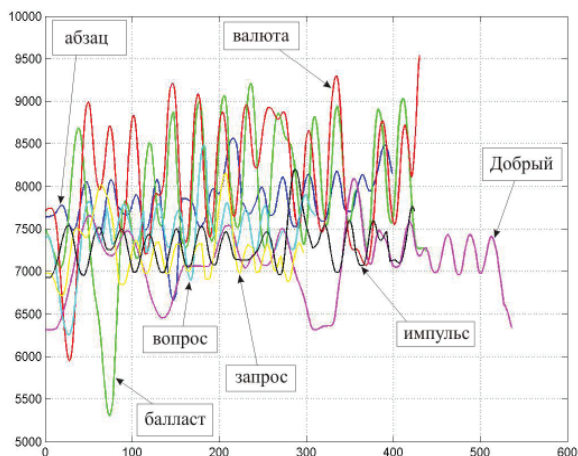


Рисунок 4 – Примеры кривых колебаний пера по оси Y при воспроизведении случайно выбранных слов

Корреляционные связи входных биометрических данных преобразователя биометрия-код являются причиной коррелированности разрядов выходного кода преобразователя. Игнорирование свойства коррелированности разрядов выходного кода преобразователя приводит к тому, что результаты оценки рисков нарушения информационной безопасности оказываются неоправданно оптимистичными. По этой причине разрабатываемые на основе данных результатов методики противодействия угрозам нарушения информационной безопасности не приводят к желаемому эффекту.

Второй причиной могут являться случайно появляющиеся корреляционные связи при обучении нейронов одного слоя. Даже если используются декорреляционные алгоритмы обучения [11, 42] они работают только при обучении одного нейрона. Если алгоритм устранения корреляции одинаков и применяется к разным нейронам, то существует вероятность усиления им корреляционных связей нейронов в одном слое. Вероятность такого события растет в случае, если нейроны одного слоя обрабатывают частично перекрывающиеся (частично общие) биометрические параметры. Избежать частичного перекрытия контролируемых пара-

метров у нейронов технически невозможно. Так при случайном выборе 7 входных связей у каждого из 416 нейронов сети достаточно часто будут встречаться нейроны, имеющие по 1, 2, 3 общим связям. Естественно, что выходы этих нейронов будут иметь повышенный уровень корреляционной связи. При создании биометрико-нейросетевых преобразователей необходимо следить за тем, чтобы ощутимо коррелированных разрядов выходного кода было мало.

Основной задачей при оценке вероятности ложной аутентификации «Чужого» является восстановление функций

$$\alpha_n \left(m(r) \cdot m\left(\frac{H}{n}\right) \right) \text{ в интересующих нас точках по корреляции.}$$

В связи с тем, что корреляция является определяющим параметром, влияющим на качество преобразователя биометрия-код необходимо установить, как меняется распределение кодов-откликов «Все чужие» преобразователя при изменении корреляции. Чтобы выяснить каким образом изменяется закон распределения при разных значениях корреляции необходимо проводить тестирование нейросетевого преобразователя биометрия-код с коррелированными связями между битами выходного кода.

2 ТЕСТИРОВАНИЕ НЕЙРОСЕТЕВОГО ПРЕОБРАЗОВАТЕЛЯ БИОМЕТРИЯ-КОД С КОРРЕЛЯЦИОННЫМИ СВЯЗЯМИ МЕЖДУ БИТАМИ ВЫХОДНОГО КОДА

Ранее было установлено, что основным свойством преобразователя биометрия-код является его свойство высокой размерности со стороны входов и выходов. Свойство высокой размерности преобразователя со стороны входов и выходов является основной причиной, препятствующей использованию классического биномиального закона распределения при прогнозировании вероятности ложной аутентификации «Чужого».

В работах Иванова А.И. [41, 43-46] обоснована и проведена оценка размерности преобразователя рукописного образа парольного слова в код со стороны входов и выходов. Путем статистических вычислений получено, что для выходных кодов преобразователя длиной 256 бит достаточно использования 416 входов, 3 слоя нейронов по 256 нейронов в каждом слое, 32 входа у каждого нейрона. Такой преобразователь позволяет обеспечивать вероятности ложной аутентификации «Чужого» порядка 10^{-12} попыток подбора.

Обучение преобразователя происходит в соответствии с российским национальным стандартом [39]. Схема обучения преобразователя приведена на рисунке 5.

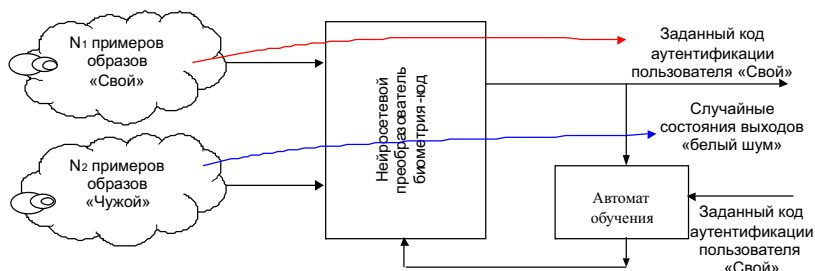


Рисунок 5 – Структурная схема обучения нейросетевого преобразователя биометрия-код в соответствии с российским национальным базовым стандартом ГОСТ Р 52633.0-2006

При обучении используются биометрические образы пользователя «Свой», биометрические образы пользователя «Чужой» и заданный код аутентификации. При подаче на входы правильно обученного преобразователя биометрических образов пользователя «Свой» на выходе формируется заданный при обучении код аутентификации. Расстояние между кодом-откликом «Свой» и кодами-откликами «Все чужие» оценивается с помощью меры Хэмминга. Мера Хэмминга описывает расстояние между сравниваемыми кодами, являясь дискретной величиной. Нам необходима мера Хэмминга при сравнении кода «Свой» и кодов «Чужие». Для вычисления меры Хэмминга необходимо просуммировать сравниваемые коды по модулю два:

$$\text{код}_0 \oplus \text{код}_Y = \text{код}_n,$$

что дает код Хэмминга – «код_n», в котором единицы соответствуют расхождению разрядов складываемых кодов:

$$\begin{array}{rcl} \text{код}_0 & & 0000001111111 \\ \text{код}_Y & & 0001001100111 \\ \hline \text{код}_n & = & 0001000011000 \end{array}$$

Суммирование по модулю два – $\oplus$ дает «1» при одинаковых значениях суммируемых разрядов и «0» если они отличаются:

$$\begin{array}{ll} \langle 0 \rangle \oplus \langle 0 \rangle = \langle 1 \rangle; & \langle 0 \rangle \oplus \langle 1 \rangle = \langle 0 \rangle; \\ \langle 1 \rangle \oplus \langle 1 \rangle = \langle 1 \rangle; & \langle 1 \rangle \oplus \langle 0 \rangle = \langle 0 \rangle. \end{array}$$

При этом мера Хэмминга – это скаляр, суммирующий все единицы в коде Хэмминга, или подсчитывающая общее число несовпавших бит двух сравниваемых кодов. Для приведенного выше примера мера Хэмминга составляет значение 3 ($H = 3$). Очевидно, что минимальное значение меры Хэмминга $H = 0$ соответствует полному тождеству сравниваемых кодов. Максимальное значение меры Хэмминга может принимать равное длине сравниваемых кодов $H = n$. Это соответствует ситуации, когда сравниваемые коды не совпадают ни в одном разряде.

Мера Хэмминга (число несовпавших бит) кодов-откликов «Свой» и заданного при обучении кода аутентификации равно

нулю. При подаче на входы правильно обученного преобразователя случайных биометрических образов множества «Чужие» на выходе формируются случайные коды-отклики.

Сравнение кодов-откликов преобразователя биометрия-код по мере Хэмминга используется при тестировании преобразователя. Структурная схема проведения численного эксперимента по тестированию нейросетевого преобразователя биометрия-код приведена на рисунке 6.

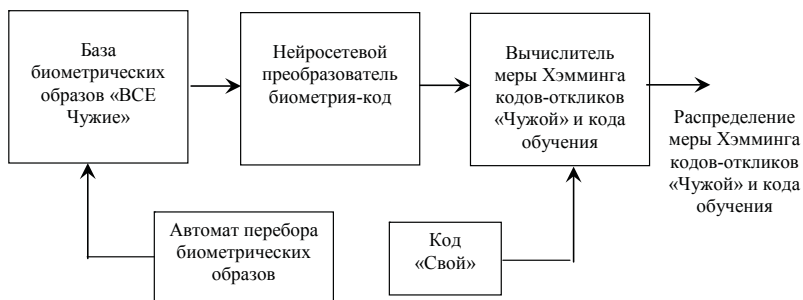


Рисунок 6 – Структурная схема проведения численного эксперимента по тестированию нейросетевого преобразователя биометрия-код

На схеме рисунка 6 используется база биометрических образов «Все Чужие». Объем такой базы обратно пропорционально значению вероятности ложной аутентификации «Чужого», обеспечиваемому тестируемым преобразователем. Например, если тестируемый преобразователь обеспечивает значение вероятности ложной аутентификации «Чужого» порядка $P_2=10^{-12}$, то для тестирования требуется база образов «Все Чужие» объемом как минимум $N=10^{13}$ с целью получения статистической достоверности результатов тестирования. В работах [29, 30] были разработаны методики получения таких баз большого и сверхбольшого объема, например, порядка $N=10^{12}$. Они заключаются в сборе реальных образов, воспроизводимых множеством пользователей «Все Чужие» и размножении реальных образов синтетическими образами по методам, предложенным в [47]. Возможно сжатие баз до объема $N=10^3$ с целью проведения ускоренного тестирования преобразователя [48].

Перебор биометрических образов базы «Все Чужие» производит автомат перебора, который строится в соответствии с требованиями стандартов [11, 39]. Данный автомат может реализовывать схему прямого или направленного перебора.

При прямом переборе происходит последовательный обход образов содержащихся в базе «Все Чужие» и подача образа на входы преобразователя. При направленном переборе используются методы разделения входного пространства образов на группы по наибольшему подобию, методы ортогонализации поиска и т.д. Эти методы служат ускорению поиска и получению результата за то же самое время при меньшем объеме базы или при фиксированном объеме базы за меньшее время. Выбранные образы базы «Все Чужие» подаются на входы преобразователя.

Перед тестированием преобразователя его необходимо обучить распознавать множество биометрических образов «Свой» в соответствии с требованиями стандартов [11, 39]. При обучении используются биометрические образы пользователя «Свой», биометрические образы пользователя «Чужой» и заданный код аутентификации. При подаче на входы правильно обученного преобразователя биометрических образов пользователя «Свой» на выходе формируется заданный при обучении код аутентификации. Мера Хэмминга (число несовпавших бит) кодов-откликов «Свой» и заданного при обучении кода аутентификации равно нулю.

При подаче на входы правильно обученного преобразователя случайных биометрических образов множества «Чужие» на выходе формируются случайные коды-отклики (рисунок 7). При этом, регулируя параметры нелинейных элементов, можно смещать вероятности состояний на выходе преобразователя биометрия-код. Идеальный преобразователь биометрия-код построен таким образом, чтобы входные данные каждого нейрона не повторялись во входных данных других нейронов.

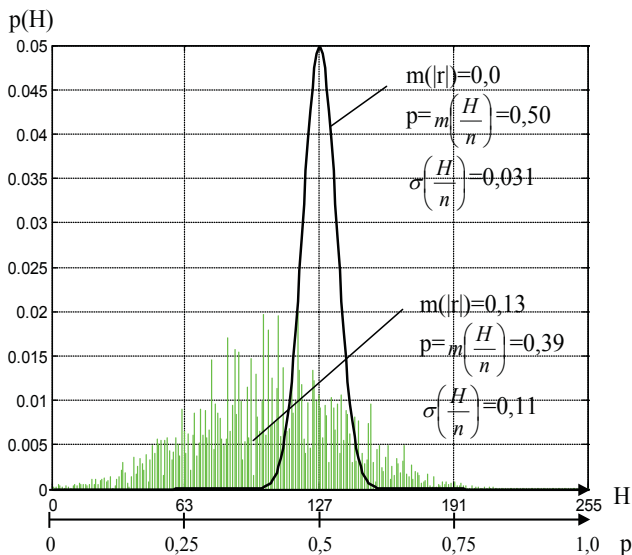


Рисунок 7 – Результаты тестирования 256-выходного преобразователя

биометрия-код при $m(|r|)=0,13$; $p=m\left(\frac{H}{n}\right)=0,39$; $\sigma\left(\frac{H}{n}\right)=0,11$

На выходе идеального преобразователя формируются N случайных кодов-откликов, вычисляется мера Хэмминга кодов-откликов «Чужой» и заданного при обучении преобразователя кода аутентификации пользователя «Свой». Формируется распределение меры Хэмминга кодов-откликов «Чужой» и заданного при обучении преобразователя кода аутентификации пользователя «Свой».

Результаты тестирования 256-выходного преобразователя биометрия-код при $m(|r|)=0,13$; $p=m\left(\frac{H}{n}\right)=0,39$; $\sigma\left(\frac{H}{n}\right)=0,11$ приведены на рисунке 7. На рисунке 7 математическое ожидание меры Хэмминга равно половине длины выходного кода $m(H)=128$, среднеквадратическое отклонение равно $\sigma(H)=16$. При вычислении нормированной меры Хэмминга данные параметры равны соответственно $m\left(\frac{H}{n}\right)=0,5$ и $\sigma\left(\frac{H}{n}\right)=0,0625$.

Из данных, приведенных на рисунке 7, видно, что распределение меры Хэмминга кодов-откликов «Чужой» реального преобразователя существенным образом отличается от распределения для идеального преобразователя. Распределение смещено влево и имеет больший разброс значений, чем у идеального преобразователя.

В параметрах классической схемы Бернулли смещение влево говорит о том, что вероятность каждого опыта не равна 0,5. Увеличение разброса значений не может быть описано с помощью классической схемы Бернулли. Основной причиной является появление корреляционных связей между битами выходного кода реального преобразователя. При использовании модификации схемы Бернулли увеличение разброса значений описывается увеличением параметра корреляции. Корреляция битов выходного кода и вероятность появления «0» в бите кода как параметры модифицированной схемы Бернулли позволяют описать изменения закона распределения меры Хэмминга кодов-откликов «Чужой».

В работах [49-54] показано, что основным свойством преобразователя биометрия-код наряду с высокой размерностью преобразователя со стороны входов и выходов является коррелированность разрядов выходного кода. Установлено что доля стабильной части индивидуальных признаков в значении меры Хэмминга, которая составляет 85% от кода аутентификации. Свойство коррелированности разрядов выходного кода также не позволяет использовать классический биномиальный закон распределения при прогнозировании вероятности ложной аутентификации «Чужого». Поэтому при оценке вероятности ложной аутентификации «Чужого» необходимо вычислять коэффициенты корреляции разрядов выходного кода преобразователя. При этом требуется провести оценку вычислительных ресурсов при определении коэффициентов корреляции. Также установлено, что число внутренних состояний преобразователя составляет порядка 10^{1000} состояний. При этом возникло противоречие, заключающееся в том, что прямое решение задачи вычисления коэффициентов корреляции технически нереализуемо, так как время перебора такого

огромного числа внутренних состояний преобразователя может занять порядка 10^{13} лет.

Поэтому предложено вычислять коэффициенты корреляции через закон распределения меры Хэмминга кодов-откликов «Чужой». Использование закона распределения меры Хэмминга позволяет сократить время оценки вероятности ложной аутентификации «Чужого» с 10^{13} лет до 3,5 секунд. Использование меры Хэмминга вместо прямого вычисления коэффициентов парной корреляции и определение доли стабильной части индивидуальных признаков позволило сократить на 15 порядков время вычисления коэффициентов корреляции [55].

Таким образом, основными причинами, препятствующими использованию принятого в работах зарубежных исследователей классического биномиального закона при описании распределения меры Хэмминга кодов-откликов «Чужой», являются свойство высокой размерности преобразователя биометрия-код со стороны входов и выходов и свойство коррелированности разрядов выходного кода преобразователя биометрия-код.

В качестве статистической модели преобразователя биометрия-код целесообразно использовать модифицированную схему Бернулли проведения испытаний с зависимыми исходами. Параметрами модифицированной схемы Бернулли являются в данном случае математическое ожидание модуля коэффициентов парной корреляция битов выходного кода $m(|r|)$, математическое ожида-

ние нормированной меры Хэмминга $m\left(\frac{H}{n}\right)$ (топологический эквивалент вероятности появления «0» в бите кода p) и длина выходного кода n . Длина выходного кода преобразователя, исходя из требований стандарта [11], выбирается $n=256$ при использовании кода-отклика «Свой» как кода аутентификации. Модификацию схемы Бернулли целесообразно использовать как основу для синтеза закона набора законов распределения при заданных значениях параметров схемы испытаний. По синтезированным законам определяется значение вероятности ложной аутентификации «Чужого».

Восстановить функцию $\alpha_n \left(m(|r|) \cdot m \left(\frac{H}{n} \right) \right)$, определяющую

топологические преобразования формы закона распределения, целесообразно численно в виде таблиц биномиального зависимого распределения в интересующих нас точках по корреляции.

В общем случае эта функция может быть восстановлена при любых значениях длины выходного кода преобразователя. Так

как рассматривается функция $\alpha_{256} \left(m(|r|) \cdot m \left(\frac{H}{n} \right) \right)$ для 256-выход-

ного преобразователя, исходя из требований стандартов [11, 39] и необходимой длине кода аутентификации при использовании в приложения безопасности, то численное моделирование, проводится при длине выходного кода преобразователя равной 256 бит.

3 МОДЕЛИРОВАНИЕ БИНОМИАЛЬНОГО ЗАВИСИМОГО ЗАКОНА РАСПРЕДЕЛЕНИЯ МЕРЫ ХЭММИНГА КОДОВ-ОТКЛИКОВ «ЧУЖОЙ» ПРЕОБРАЗОВАТЕЛЯ БИОМЕТРИЯ-КОД

Вначале рассмотрим идеальный преобразователь, у которого отсутствуют корреляционные связи между разрядами выходного кода [50]. Идеальный преобразователь является основой для разработки методики повышения достоверности автоматизированного прогнозирования вероятности ложной аутентификации «Чужого».

Схема моделирования закона распределения меры Хемминга кодов-откликов «Чужой» идеального преобразователя биометрия-код приведена на рисунке 8.

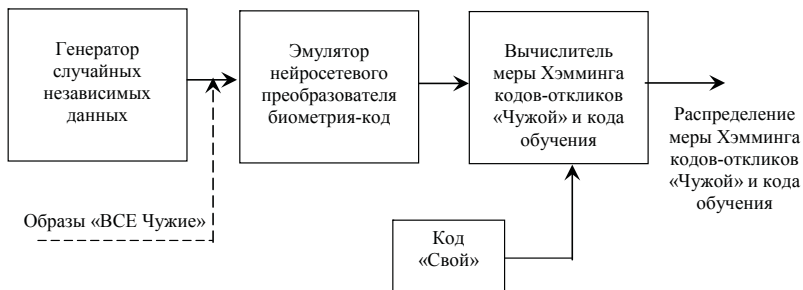


Рисунок 8 – Структурная схема проведения численного эксперимента по моделированию распределения меры Хэмминга кодов-откликов «Чужой» идеального нейросетевого преобразователя биометрия-код

На схеме рисунка 8 блок 1 генерирует N векторов независимых нормально распределенных данных, эмулируя множество биометрических образов «Все Чужие». Данные с идеального генератора поступают на вход эмулятора преобразователя биометрия-код, заранее обученного распознавать множество биометрических образов «Свой». При обучении используются биометрические об-

разы пользователя «Свой», биометрические образы пользователя «Чужой» и заданный код аутентификации. При подаче на входы правильно обученного преобразователя биометрических образов пользователя «Свой» на выходе формируется заданный при обучении код аутентификации. Мера Хэмминга (число несовпавших бит) кодов-откликов «Свой» и заданного при обучении кода аутентификации равно нулю.

При подаче на входы правильно обученного преобразователя случайных биометрических образов множества «Чужие» на выходе формируются случайные коды-отклики. При этом, регулируя параметры нелинейных элементов, можно смещать состояния кодов-откликов на выходе преобразователя биометрия-код. Идеальный преобразователь биометрия-код построен таким образом, чтобы входные данные каждого нейрона не повторялись во входных данных других нейронов. На выходе идеального преобразователя формируются N случайных кодов-откликов, вычисляется мера Хэмминга кодов-откликов «Чужой» и заданного при обучении преобразователя кода аутентификации пользователя «Свой». Формируется распределение меры Хэмминга кодов-откликов «Чужой» и заданного при обучении преобразователя кода аутентификации пользователя «Свой».

Проведенные численные эксперименты по схеме рисунка 1 показали, что выходные случайные коды идеального преобразователя биометрия-код достоверно описываются классическим биномиальным законом распределения независимых данных значений с параметрами числа степеней свободы – n и математиче-

ского ожидания нормированной меры Хэмминга – $m\left(\frac{H}{n}\right)$.

В этом случае вероятность появления k событий из n возможных событий при заданной вероятности p этого события вычисляется следующим образом:

$$P\left(n, m\left(\frac{H}{n}\right)\right) = \frac{n!}{k!(n-k)!} \left(m\left(\frac{H}{n}\right)\right)^k \left(1 - m\left(\frac{H}{n}\right)\right)^{n-k} \quad (6)$$

Выражение (6) аналогично выражению (1), где вместо вероятности появления «0» в бите кода p используется математическое ожидание нормированной меры Хэмминга $m\left(\frac{H}{n}\right)$ как ее топологический эквивалент.

Гистограммы распределения значений меры Хэмминга кодов-откликов «Чужой» идеального преобразователя при различных значениях параметра $m\left(\frac{H}{n}\right)$ для $n=256$ приведены на рисунке 9.

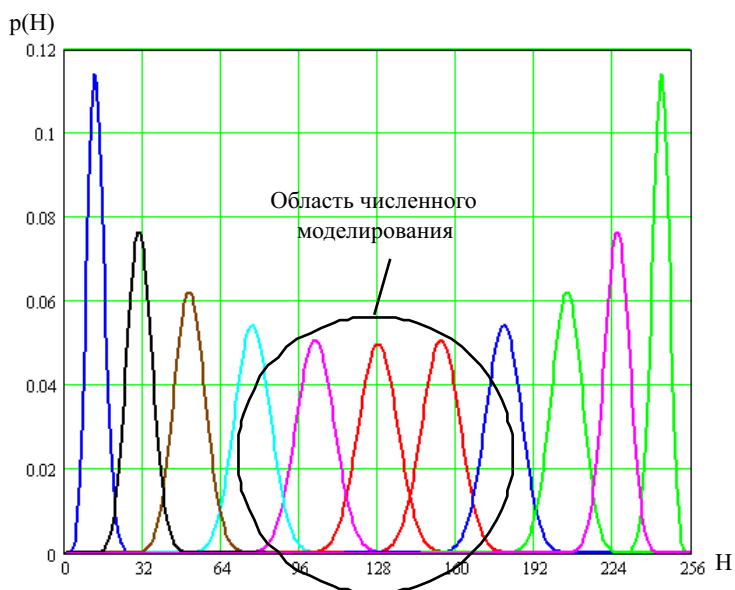


Рисунок 9 – Гистограммы распределения значений меры Хэмминга кодов-откликов «Чужой» идеального преобразователя при различных значениях параметра $m\left(\frac{H}{n}\right)$ для $n=256$

Выбор обозначенной на рисунке 9 области численного моделирования обусловлен предположением о правильном обучении идеального преобразователя биометрия-код, когда $m\left(\frac{H}{n}\right)=0,5$.

При ошибках обучения параметр $m\left(\frac{H}{n}\right)$ принимает другие значения. Установлено, что результаты численного эксперимента согласуются с теоретическими расчетами. Среднеквадратическое отклонение численных данных падает пропорционально квадратному корню числа экспериментов.

Проведение вычислений по выражению (3) позволяет получать оценку вероятности ложной аутентификации «Чужого» обученным преобразователем с некоррелированными выходными данными. Вероятность ложной аутентификации «Чужого» вычисляется следующим образом:

$$D_2 = \left(m\left(\frac{H}{n}\right) \right)^n \quad (7)$$

Однако использование выражения (7) дает неоправданно оптимистические прогнозы вероятности ложной аутентификации «Чужого» если мы начинаем тестировать реальный преобразователь с 256 выходами. При нулевом значении корреляции $m(|r|)$ оценка (4) для идеального и реального преобразователя дает одинаковые результаты и согласуется с результатами реального тестирования. По мере увеличения корреляции результаты прогнозирования по выражению (7) для реального преобразователя начинают сильно расходиться с результатами реального тестирования.

При корреляции равной $m(|r|)=0,15$ ошибка составляет 70% от прогнозируемого значения вероятности ложной аутентификации «Чужого». При корреляции равной $m(|r|)=0,5$ преобразователь теряет способность к защите. В этой ситуации на его выходах могут формироваться прямо противоположные по расстоянию Хэмминга коды-отклики при подаче на его входы образов множества «Свой» или «Все Чужие».

И, наоборот, на выходах преобразователя могут формироваться одинаковые по расстоянию Хэмминга коды-отклики при подаче на его входы разных образов из множеств «Свой» и об-

разов из множества «Все Чужие». Такой преобразователь в этих условиях теряет способность к защите. При корреляции близкой к $m(|r|)=1,0$ ошибка достигает 100% от прогнозируемого значения вероятности [56-60].

С целью получения достоверного прогноза вероятности ложной аутентификации «Чужого» реальным преобразователем биометрия-код далее разработана методика численного моделирования закона распределения меры Хэмминга кодов-откликов «Чужой» реального преобразователя, выходные данные которого обладают корреляционными связями [61]. Методика, в свою очередь, является основой для разработки методики повышения достоверности автоматизированного прогнозирования вероятности ложной аутентификации «Чужого». Схема реализации методики численного моделирования закона распределения меры Хэмминга кодов-откликов «Чужой» реального преобразователя биометрия-код приведена на рисунке 10.

На схеме рисунка 10 блок 1, как при моделировании идеального преобразователя, генерирует N векторов независимых нормально распределенных данных множества биометрических образов «Все Чужие». Коррелятор (блок 2) вводит корреляционные связи между входными данными, регулируемым параметром является модуль коэффициентов парной корреляции $|r|$.

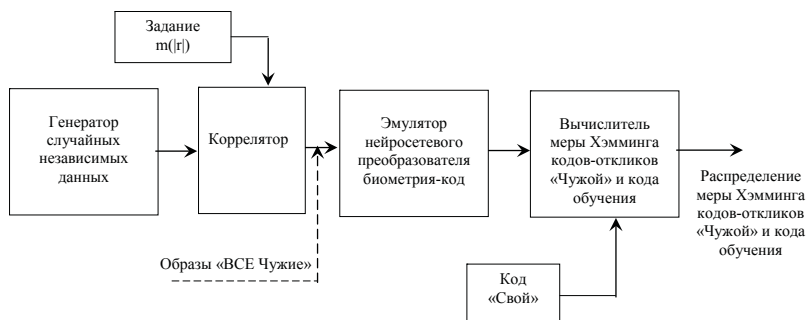


Рисунок 10 – Структурная схема проведения численного эксперимента по моделированию распределения меры Хэмминга кодов-откликов Чужой реального нейросетевого преобразователя биометрия-код

Данные с коррелятора поступают на вход эмулятора преобразователя биометрия-код, заранее обученного распознавать множество биометрических образов «Свой». При обучении используются биометрические образы пользователя «Свой», биометрические образы пользователя «Чужой» и заданный код аутентификации. При подаче на входы правильно обученного преобразователя биометрических образов пользователя «Свой» на выходе формируется заданный при обучении код аутентификации. Мера Хэмминга (число несовпавших бит) кодов-откликов «Свой» и заданного при обучении кода аутентификации равно нулю. При подаче на входы правильно обученного преобразователя случайных биометрических образов множества «Чужие» на выходе формируются случайные коды-отклики. При этом, регулируя параметры нелинейных элементов, мы можем смещать вероятности состояний на выходе преобразователя биометрия-код. На выходе преобразователя формируются N случайных кодов-откликов, вычисляется мера Хэмминга кодов-откликов «Чужой» и заданного при обучении преобразователя кода аутентификации пользователя «Свой». Формируется распределение меры Хэмминга кодов-откликов «Чужой» и заданного при обучении преобразователя кода аутентификации пользователя «Свой».

Переход от классического биномиального распределения к неклассическому биномиальному зависимому распределению

происходит с помощью функции $\alpha_n \left(m(r), m\left(\frac{H}{n}\right) \right)$. При длине вы-

ходного кода преобразователя 256 бит эта функция записывается

как $\alpha_{256} \left(m(r), m\left(\frac{H}{n}\right) \right)$. Модификация схемы Бернулли дает веро-

ятность появления k событий из n возможных событий при заданной вероятности p этого события вычисляется следующим образом:

$$D_k \left(n, m\left(\frac{H}{n}\right) \right) = \left[\frac{n!}{k!(n-k)!} \right]^{\alpha_{256} \left(m(r), m\left(\frac{H}{n}\right) \right)} \left[m\left(\frac{H}{n}\right)^k \right]^{\alpha_{256} \left(m(r), m\left(\frac{H}{n}\right) \right)} \left[\left(1 - m\left(\frac{H}{n}\right) \right)^{n-k} \right]^{\alpha_{256} \left(m(r), m\left(\frac{H}{n}\right) \right)} \quad (8)$$

Функция $\alpha_{256}\left(m(r), m\left(\frac{H}{n}\right)\right)$ задается в виде таблиц биномиального зависимого распределения синтезированных по результатам численного моделирования по схеме рисунка 10.

Установлено, что биномиальный зависимый закон, получаемый по выражению (8), существенным образом отличается от классического биномиального закона распределения независимых данных. Установлено, что при малых значениях математического ожидания модуля коэффициентов парной корреляции биномиальный зависимый закон распределения похож на нормальный закон. При этом оценки вероятности ложной аутентификации «Чужого» зарубежных авторов практически совпадают с оценками, полученными на основе численного моделирования. Результаты имитационного моделирования нейросетевого преобразователя

биометрия-код при $m(r)=0,15$; $p=m\left(\frac{H}{n}\right)=0,50$; $\sigma\left(\frac{H}{n}\right)=0,10$ приведены на рисунке 11.

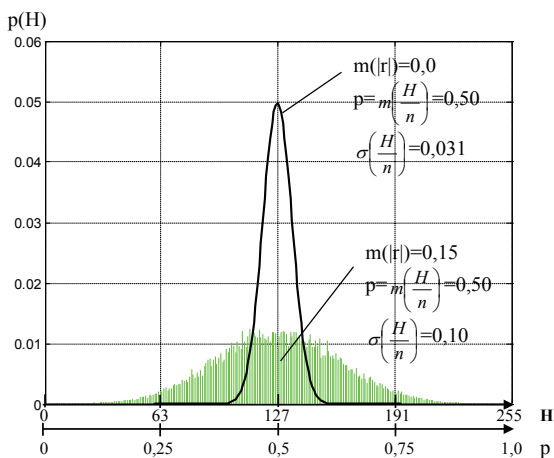


Рисунок 11 – Результаты имитационного моделирования нейросетевого преобразователя биометрия-код при $m(r)=0,15$;

$$p=m\left(\frac{H}{n}\right)=0,50; \sigma\left(\frac{H}{n}\right)=0,10$$

Установлено, что по мере увеличения корреляции $m(|r|)$ растет среднеквадратическое отклонение меры Хемминга $\sigma\left(\frac{H}{n}\right)$. При этом увеличиваются ошибки оценки вероятности ложной аутентификации «Чужого».

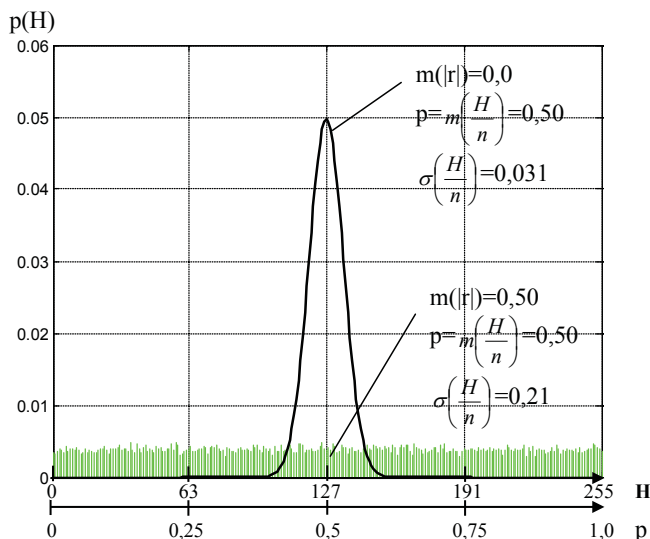


Рисунок 12 – Результаты имитационного моделирования нейросетевого преобразователя биометрия-код при $m(|r|)=0,50$;

$$p=m\left(\frac{H}{n}\right)=0,50; \sigma\left(\frac{H}{n}\right)=0,21$$

При корреляции $m(|r|)=0,15$, допустимой российским национальным стандартом ГОСТ Р 52633.0 – 2006 [11], оценки вероятности ложной аутентификации «Чужого» зарубежных авторов получаются сильно завышенными по сравнению с реальными оценками, полученными на основе численного моделирования, и дают ошибки порядка 70% от 256 бит выходного кода преобразователя.

Установлено, что при корреляции $m(|r|) = 0.5$ биномиальное зависимое распределение вырождается в равномерное. При этом способность преобразователя биометрия-код к защите вырожда-

ется. Результаты имитационного моделирования нейросетевого преобразователя биометрия-код при $m(|r|)=0,50$; $p=m\left(\frac{H}{n}\right)=0,50$; $\sigma\left(\frac{H}{n}\right)=0,21$ приведены на рисунке 12.

На основе численного моделирования установлено, что при $m(|r|) > 0,5$ происходит перерождение биномиального зависимого распределения в двумодальный закон с минимальной плотностью распределения значений в центре. Установлено, что при корреляции, близкой к единице, ошибки оценки вероятности ложной аутентификации «Чужого» зарубежными авторами по сравнению с оценками, полученными на основе численного моделирования, приближаются к 100% от 256 бит выходного кода преобразователя. Результаты имитационного моделирования нейросетевого преобразователя биометрия-код при $m(|r|)=0,60$; $p=m\left(\frac{H}{n}\right)=0,50$; $\sigma\left(\frac{H}{n}\right)=0,29$ приведены на рисунке 13.

Данные топологические преобразования биномиального зависимого закона распределения позволяют вычислять среднеквадратическое отклонение нормированной меры Хэмминга $\sigma\left(\frac{H}{n}\right)$ при малых значениях корреляции.

В противном случае вычисляется математическое ожидание модуля коэффициентов парной корреляции $m(|r|)$. Таким образом, на основе численного моделирования, установлено, что распределение значений меры Хэмминга кодов-откликов «Чужой» идеального нейросетевого преобразователя биометрия-код описывается классическим биномиальным законом распределения. Установлено, что распределение значений меры Хэмминга кодов-откликов «Чужой» реального нейросетевого преобразователя биометрия-код описывается неклассическим биномиальным зависимым законом распределения.

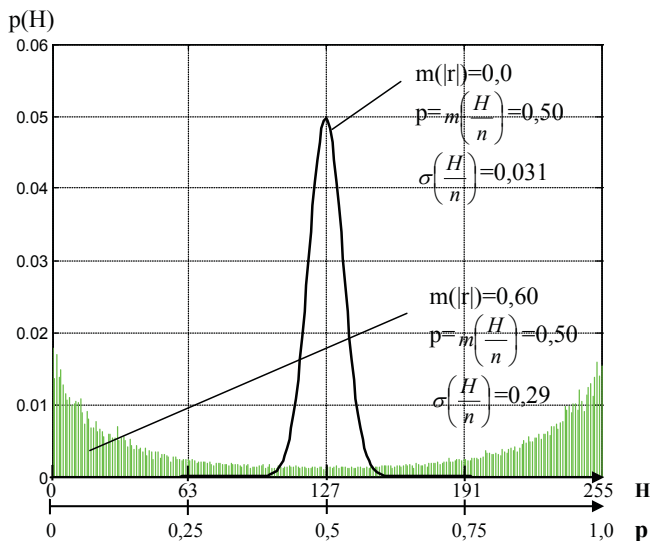


Рисунок 13 – Результаты имитационного моделирования нейросетевого преобразователя биометрия-код при $m(r)=0,60$;

$$p=m\left(\frac{H}{n}\right)=0,50; \sigma\left(\frac{H}{n}\right)=0,29$$

Получить вектор независимых случайных параметров с нормальным законом распределения несложно. Для этого можно воспользоваться любым из известных вариантов программных или аппаратных генераторов случайных чисел. Для того, что бы сделать данные зависимыми, необходимо их вектор умножить на некоторую матрицу связанности. Например, может использоваться матрица, состоящая из единиц с одинаковыми числами на диагонали:

$$\begin{bmatrix} a & 1 & :: & 1 \\ 1 & a & :: & 1 \\ :: & :: & :: & :: \\ 1 & 1 & :: & a \end{bmatrix} \times \begin{bmatrix} x_1 \\ x_2 \\ :: \\ x_n \end{bmatrix} = \begin{bmatrix} v_1 \\ v_2 \\ :: \\ v_n \end{bmatrix} \Rightarrow [r_{v_i v_j}] = \begin{bmatrix} 1 & r & :: & r \\ r & 1 & :: & r \\ :: & :: & :: & :: \\ r & r & :: & 1 \end{bmatrix}, \quad (9)$$

где x_i – случайные числа, полученные от генератора с нормальным законом распределения значений, нулевым математическим ожиданием и единичной дисперсией; a – регулируемый параметр, влияющий на значение коэффициентов корреляции – r .

Зависимость коэффициента корреляции от параметра « a » при разной размерности генерируемого вектора приведена на рисунке 14. Можно показать, что описанная выше процедура генерации зависимых данных дает вектор с одинаковыми среднеквадратическими отклонениями:

$$\sigma(v_i) = \sqrt{n - 1 + a^2} \quad (10)$$

и нулевыми математическими ожиданиями для каждой компоненты v_i .

На рисунке 14 показана номограмма связи коэффициента – r с параметром « a » при длине вектора коррелированных данных $n=4;8;12;16;32;48;64$ при использовании модели равнокоррелированных входных данных описываемой выражениями (9) и (10). Из рисунка 14 видно, что зависимость коэффициента – r от параметра « a » сохраняет масштаб при изменении длины ключа n длины вектора коррелированных данных n . Одновременно происходит некоторый спад кривых в средней при длине вектора коррелированных данных что может быть связано с высокой чувствительностью преобразования (9) к параметру a при малых размерах трансформирующей матрицы. По мере увеличения размера матрицы трансформации зависимость a от r становится практически линейной.

Корреляционная матрица с одинаковыми коэффициентами корреляции является простой, однако она оказывается очень удобной при моделировании. Более сложные корреляционные матрицы получаются незначительными вариациями данной процедуры.

Необходимо отметить, что биометрические данные с одинаковыми по знаку корреляционными функциями не встречаются на практике. В связи с этим, необходимо усовершенствовать механизм синтеза зависимых данных с тем, чтобы он позволял по-

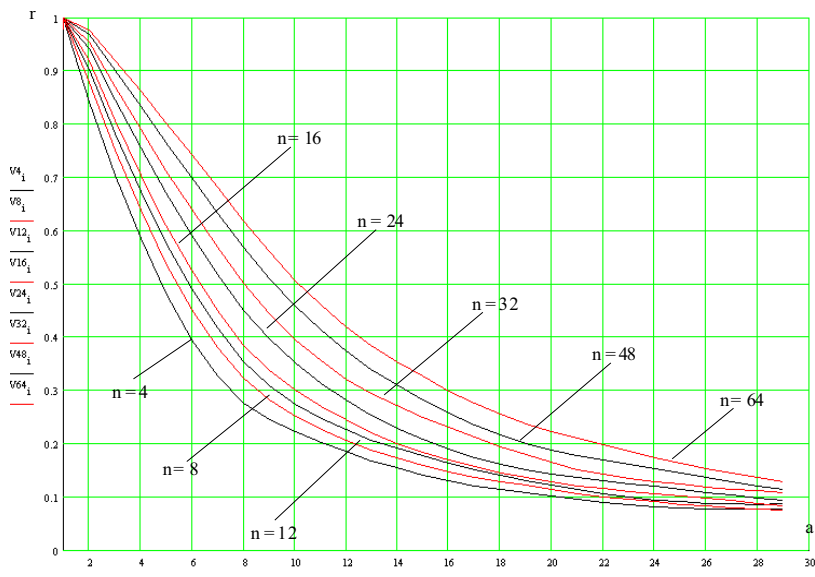


Рисунок 14 – Номограмма связи коэффициента r с параметром «a» при длине вектора коррелированных данных $n=4;8;12;16;32;48;64$

лучать более сложные данные с корреляционными матрицами, в которых знаки коэффициентов корреляции меняются.

Для этой цели необходимо использовать матрицу преобразования исходных независимых случайных данных со строками, имеющими разный знак:

$$\begin{bmatrix} a & 1 & :: & 1 \\ -1 & -a & :: & -1 \\ :: & :: & :: & :: \\ 1 & 1 & :: & a \end{bmatrix} \times \begin{bmatrix} x_1 \\ x_2 \\ :: \\ x_n \end{bmatrix} = \begin{bmatrix} v_1 \\ v_2 \\ :: \\ v_n \end{bmatrix} \Rightarrow [r_{v_i v_j}] = \begin{bmatrix} 1 & -r & :: & r \\ -r & 1 & :: & r \\ :: & :: & :: & :: \\ r & r & :: & 1 \end{bmatrix} \quad (11)$$

При реализации процедуры со случайным распределением знаков коэффициентов корреляции необходимо правую часть матричного уравнения (11) до множить на $\varepsilon_i = \pm 1$ или

$$\begin{bmatrix} a & 1 & \ddots & 1 \\ 1 & a & \ddots & 1 \\ \ddots & \ddots & \ddots & \ddots \\ 1 & 1 & \ddots & a \end{bmatrix} \times \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} \varepsilon_1 \cdot v_1 \\ \varepsilon_2 \cdot v_2 \\ \vdots \\ \varepsilon_n \cdot v_n \end{bmatrix} \quad (12)$$

где ε_i – числа ± 1 со случайным знаком.

На рисунке 15 показаны номограммы связи коэффициента – γ с параметром «а» при длине вектора коррелированных данных $n=4;8;12;16;32;48;64$ при использовании генератора с равнокоррелированными по модулю входными данными который описывается выражениями (11),(12).

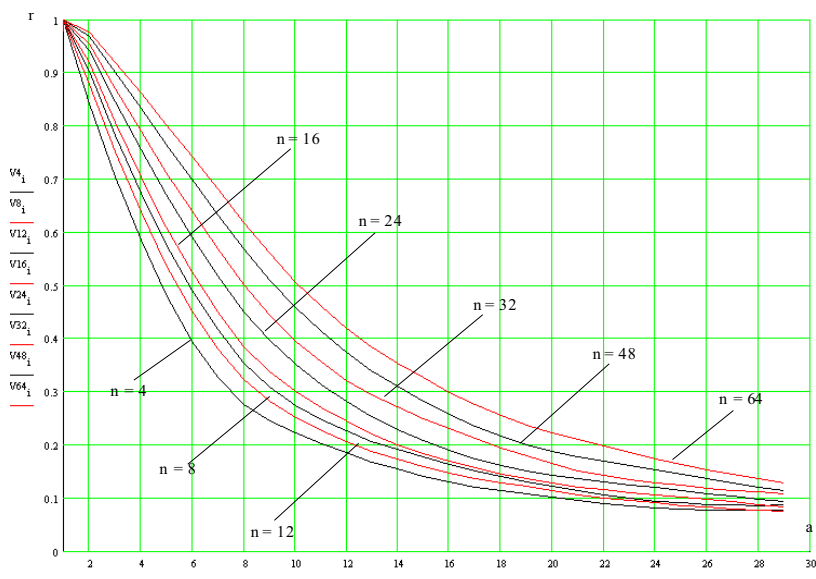


Рисунок 15 – Номограмма связи коэффициента – γ с параметром «а» при длине вектора коррелированных данных $n=4;8;12;16;32;48;64$

Из рисунка 15 видно, что зависимость коэффициента – γ от параметра «а», имеет спад в средней части при малых n . Это может быть связано с высокой чувствительностью преобразования (9) к параметру a при малых размерах трансформирующей матри-

цы. По мере увеличения размера матрицы трансформации зависимость α от r становится практически линейной.

Корреляционная матрица с равнокоррелированными по модулю входными данными похожа на матрицу одинаковыми коэффициентами корреляции. Более сложные корреляционные матрицы получаются, если использовать модель с положительно коррелированными, но случайно коррелированными по значению зависимыми входными данными.

Для изменения коэффициента корреляции в одном столбце или строке корреляционной матрицы (при сохранении одинаковыми всех других коэффициентов корреляции) достаточно к одному из диагональных элементов матрицы преобразования прибавить положительное приращение $\Delta\alpha$. Например, может быть изменен второй элемент матрицы преобразования

$$\begin{bmatrix} a & 1 & 1 & 1 & 1 \\ 1 & a + \Delta\alpha & 1 & 1 & 1 \\ 1 & 1 & a & 1 & 1 \\ 1 & 1 & 1 & a & 1 \\ 1 & 1 & 1 & 1 & a \end{bmatrix} \times \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & b & r & r & r \\ b & 1 & b & b & b \\ r & b & 1 & r & r \\ r & b & r & 1 & r \\ r & b & r & r & 1 \end{bmatrix} \quad (13)$$

где b – сниженное значение коэффициентов корреляции, $b < r$.

Исходя вышеизложенного, можно показать, что имитация случайного распределения значений коэффициентов корреляции биометрических данных, может быть достигнута случайным выбором значений диагональных элементов матрицы преобразований с единицами вне диагонали:

$$\begin{bmatrix} a_1 & 1 & 1 & 1 & 1 \\ 1 & a_2 & 1 & 1 & 1 \\ 1 & 1 & a_3 & 1 & 1 \\ 1 & 1 & 1 & a_4 & 1 \\ 1 & 1 & 1 & 1 & a_5 \end{bmatrix} \times \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & r_{12} & r_{13} & r_{14} & r_{15} \\ r_{12} & 1 & r_{23} & r_{24} & r_{25} \\ r_{13} & r_{23} & 1 & r_{34} & r_{35} \\ r_{14} & r_{24} & r_{34} & 1 & r_{45} \\ r_{15} & r_{25} & r_{35} & r_{45} & 1 \end{bmatrix} \quad (14)$$

где a_i – случайные значения коэффициентов, которые всегда больше единицы, но меньше некоторого максимального числа.

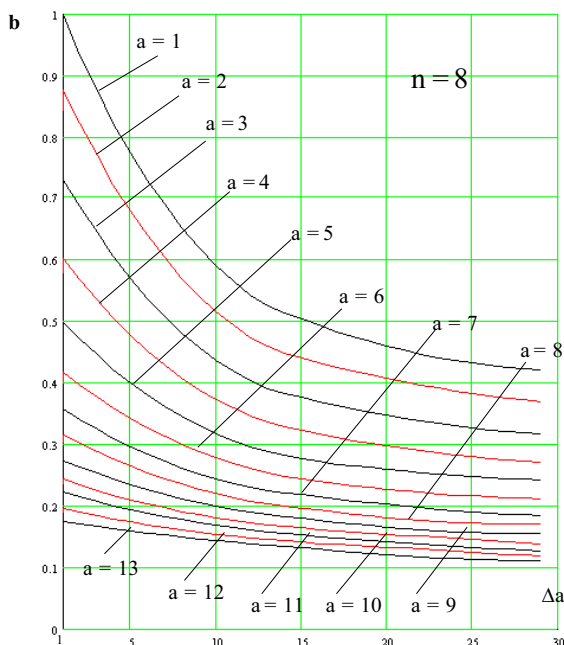


Рисунок 16 – Связь заниженного коэффициента корреляции в строке и столбце корреляционной матрицы с регулируемым параметром « Δa » для матрицы размерности 8×8

На рисунке 16 показана номограмма связи заниженного коэффициента корреляции в строке и столбце корреляционной матрицы с регулируемым параметром « Δa » для матрицы размерности 8×8 . Из данных рисунка 16 видно, что значение заниженного коэффициента корреляции в строке и столбце становится меньше при $a=6-13$. Из этого получаются соответствующие значения коэффициентов корреляции $r_{i,j}$.

Из вышеприведенных рисунков видно, что дисперсия меры Хемминга, как и сам закон распределения значений меры Хемминга множества «Все Чужие», сильно зависит от связанности (коррелированности) разрядов случайных кодов. Возникает вопрос о том, какова эта связь. Необходимо путем численных экспериментов идентифицировать эту связь.

Для идентификации вида связи, например, $\sigma(r)$ проведены проверки на чувствительность зависимости к знаку коэффициентов парной корреляции. Для этой цели сравнивались значения $\sigma(r)$ и вид функций распределения значений $p(H)$ для всех положительных коэффициентов:

$$r_{1,2} = +0.5, r_{1,3} = +0.5, r_{1,4} = +0.5, \dots, r_{i,j} = +0.5;$$

для всех отрицательных коэффициентов:

$$r_{1,2} = -0.5, r_{1,3} = -0.5, r_{1,4} = -0.5, \dots, r_{i,j} = -0.5;$$

для коэффициентов поочередно изменяющих свой знак:

$$r_{1,2} = +0.5, r_{1,3} = -0.5, r_{1,4} = +0.5, \dots, r_{i,j} = \pm 0.5.$$

Результаты численного эксперимента показали, что знаки коэффициентов парной корреляции вообще не влияют на вид закона распределения значений и его дисперсию.

То есть с точностью до ошибки, определяемой конечной обучающей выборкой, выполняется тождество:

$$\sigma(r = +0.5) \equiv \sigma(r = -0.5) \equiv \sigma(r = \pm 0.5). \quad (15)$$

В точке $r = \pm 0.5$ форма закона распределения значений $p(H, p=0.5)$ равномерная и она продолжает оставаться равномерной (смотри рисунок 29) при любом распределении знаков коэффициентов парной корреляции. Неизменность формы плотности распределения значений – $p(n, H, p=0.5, m(|r|) = 0.5)$ проще всего констатировать именно при наблюдении равномерного закона распределения значений как самого простого.

При других значениях модулей коэффициентов корреляции соотношение (20) и неизменность формы $p(n, H, p=0.5, m(|r|) = 0.5)$ продолжает выполняться. Это означает, что тождество (20) может быть записано в более общем виде:

$$\sigma(+r) \equiv \sigma(-r) \equiv \sigma(\pm r) \equiv \sigma(|r|). \quad (16)$$

Аналогичное соотношение может быть записано и для функций распределения значений в центральной точке симметрии:

$$p(n, H, p=0.5, +r) \equiv p(n, H, p=0.5, -r) \equiv p(n, H, p=0.5, \pm r) \equiv p(n, H, p=0.5, |r|), \quad (17)$$

а так же в любой иной точке:

$$\begin{aligned} p(n, H, p, +r) &\equiv p(n, H, p, -r) \equiv p(n, H, p, \pm r) \equiv \\ &\equiv p(n, H, p, |r|). \end{aligned} \quad (23)$$

Подчеркнем, что все вышесказанное выполняется только для одинаковых по модулю коэффициентов корреляции, имеющих только неопределенный знак.

В случае, если модуль коэффициентов парной корреляции может принимать несколько значений, например $m(|r|) = 0.0, 0.25, 0.5$, то ситуация усложняется и мы будем получать разные значения дисперсий и разные функции плотности распределения значений в зависимости от вероятности состояний $m(|r|) = 0.0, 0.25, 0.5$. То есть, можно говорить только о том, что дисперсии и функции плотности распределения значений будут совпадать при совпадении плотностей распределения значений модулей коэффициентов парной корреляции.

На основе результатов численного моделирования получены номограммы биномиального зависимого закона распределения в интервале $0 < m(|r|) < 1.0$ при $p = m\left(\frac{H}{n}\right) = 0.50$. Номограммы биномиального зависимого закона распределения в интервале $0 < m(|r|) < 0.5$ при $p = m\left(\frac{H}{n}\right) = 0.50$ приведены на рисунке 17.

Установлено, что известные модели распределения коэффициентов парной корреляции между входными биометрическими данными преобразователя эквивалентны модели равнокоррелированных входных данных при условии, что случайные значения распределений коэффициентов корреляции зависимых данных имеют один знак и их математическое ожидание совпадает с детерминированным значением. В модели статистического описания преобразователя биометрия-код используются равнокоррелированные входные данные и вычисленное математическое ожидание модуля коэффициентов парной корреляции. Номограммы биномиального зависимого закона распределения в интервале $0.5 < m(|r|) < 1.0$ при $p = m\left(\frac{H}{n}\right) = 0.50$ приведены на рисунке 18.

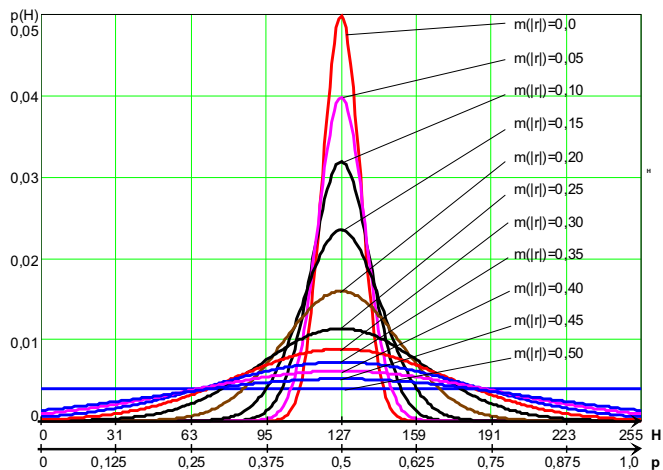


Рисунок 17 – Номограммы биномиального зависимого закона распределения в интервале $0 < m(r) < 0.5$ при $p = m\left(\frac{H}{n}\right) = 0,50$

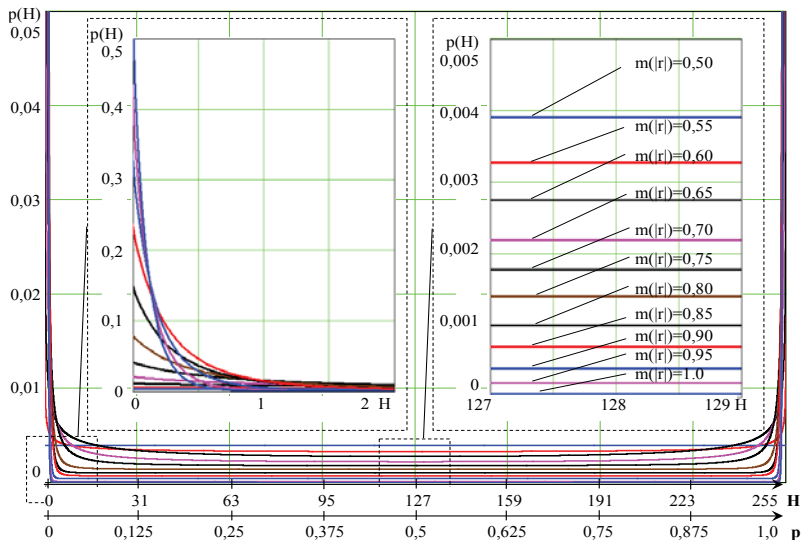


Рисунок 18 – Номограммы биномиального зависимого закона распределения в интервале $0.5 < m(r) < 1.0$ при $p = m\left(\frac{H}{n}\right) = 0,50$

Результаты моделирования дают значения функции $\alpha_{256} \left(m(r) \cdot m \left(\frac{H}{n} \right) \right)$ для 256-выходного преобразователя численно в виде таблиц биномиального зависимого распределения в интересующих нас точках по корреляции.

При необходимости предложенная методика может быть использована для восстановления функции $\alpha_n \left(m(r) \cdot m \left(\frac{H}{n} \right) \right)$ при любых значениях длины выходного кода преобразователя. Для этого повторяется численный эксперимент по схеме рисунка 10 при соответствующем значении длины выходного кода преобразователя.

4 СИНТЕЗ ТАБЛИЦ БИНОМИАЛЬНОГО ЗАВИСИМОГО ЗАКОНА РАСПРЕДЕЛЕНИЯ МЕРЫ ХЭММИНГА КОДОВ-ОТКЛИКОВ «ЧУЖОЙ» ПРЕОБРАЗОВАТЕЛЯ БИОМЕТРИЯ-КОД

Число эмулируемых биометрических образов «Все Чужие» при численном моделировании и получении таблиц биномиального зависимого закона составляет 100,000,000 образов. Интервал значений параметра $m\left(\frac{H}{n}\right)$ составляет $m\left(\frac{H}{n}\right) = 0,3-0,5$ с шагом 0,05. Интервал значений параметра $m(|r|)$ составляет $m(|r|)=0,0-0,5$ с шагом 0,05. Полученные в результате моделирования значения плотности вероятности были обработаны с помощью процедуры двумерного сглаживания случайных выбросов. Пример таблицы плотности распределения значений меры Хэмминга кодов-откликов «Чужой» обученного 256-выходного преобразователя биометрия-код при $p=m\left(\frac{H}{n}\right)=0,5$ и $m(|r|)=0.15$ приведен на рисунке 19.

Синтезированные таблицы биномиального зависимого распределения позволяют при заданных параметрах математического ожидания нормированной меры Хэмминга и математического ожидания коэффициентов парной корреляции для 256 – выходного преобразователя биометрия-код получать оценку вероятности ложной аутентификации «Чужого». Синтезированные таблицы в совокупности с методикой повышения достоверности автоматизированного прогнозирования вероятности ложной аутентификации «Чужого» представляют собой инструментарий, позволяющий администратору безопасности автоматизированной системы осуществлять выбор парольного слова для конкретного пользователя системы защиты информации.

При моделировании используются большие и сверхбольшие статистические базы «Все Чужие» объемом до $N=10^{12}$ образов и более, разработанные по методикам, предложенным в [47, 48]. Возможно также использование сжатых баз с сохранением значе-

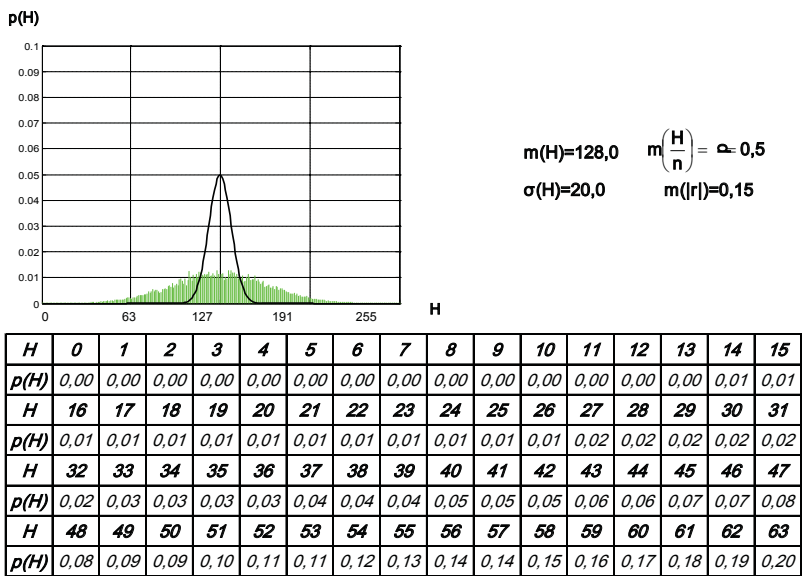


Рисунок 19 – Пример таблицы плотности распределения значений меры Хэмминга кодов-откликов «Чужой» обученного 256-выходного преобразователя биометрия-код при $p = m\left(\frac{H}{n}\right) = 0,5$ и $m(|r|) = 0.15$

ний статистических моментов, например, в пределах первых пяти моментов [62-68]. Объем базы в этом случае может составлять $N=10^4$ образов. С этой целью в [48, 66] под каждый тип биометрических образов тестовой выборки были сформированы критерии представительности: статистические характеристики среднестатистического пользователя по некоторому биометрическому параметру (математическое ожидание, среднееквадратическое отклонение, коэффициенты корреляции и другие статистические моменты); статистические характеристики среднего по некоторой группе пользователя, выделенной по некоторому биометрическому параметру (математическое ожидание, среднееквадратическое отклонение, коэффициенты корреляции и другие статистические моменты); представительность по численности групп пользователей, классифицированных по некоторому биометрическому па-

раметру. В качестве биометрических параметров использованы: стабильность написания слова-пароля, уникальность и относительная информативность биометрических образов, а также их стойкости к атакам подбора [48].

При моделировании преобразователя биометрия-код используется мера Хэмминга кодов «Свой» и «Все Чужие». Происходит понижение размерности задачи и переход от работы с множеством возможных состояний 256 битных кодов к более компактному множеству расстояний Хемминга между некоторым заданным кодом и кодами исследуемого нейросетевого преобразователя [34]. В работе [9] показано, что возможно построение модификации меры Хэмминга, позволяющая получить более точные результаты при сравнении кодов образов «Чужой» и множества «Все Чужие» или сравнения двух разных образов «Чужой 1» и «Чужой 2». Это дает возможность не зная кода «Свой», получать оценку вероятности ложной аутентификации «Чужого» по синтезированным таблицам биномиального зависимого распределения.

Как некоторую разновидность меры Хемминга следует отметить её нормированный вариант [9], который получается делением значений меры на её максимально возможное значение. После нормировки все значения меры Хемминга попадают в интервал от 0.0 до 1.0, что позволяет сравнивать между собой распределения мер Хемминга, полученные для нейросетевых преобразователей с разным числом выходов.

Еще одной интересной модификацией меры Хемминга h является её дополнение инверсной мерой $\bar{h}$. Если обычная мера Хемминга описывает кодовое расстояние между предъявленным кодом и заранее известным кодом «Свой», то инверсная мера описывает расстояние Хемминга между инверсией кода «Свой» и предъявленным кодом. Очевидно, что обычная и инверсная меры Хемминга связаны очень простым соотношением:

$$h = \max(h) - \bar{h} . \quad (18)$$

Обычная мера Хемминга построена на подсчете совпавших разрядов сравниваемых кодов. Инверсная мера (18) построена на

учете числа несовпавших разрядов сравниваемых между собой кодов. Заметим, что при вычислении инверсной меры Хемминга можно инвертировать любой из сравниваемых кодов. При вычислении линейных функционалов в пространстве значений кодов Хемминга инверсная мера является избыточной, нет смысла её использовать. При вычислении нелинейных функционалов инверсная мера Хемминга начинает играть важную роль при многомерных статистических исследованиях нейросетевых преобразователей биометрия-код, которые по своей природе являются нелинейными и, следовательно, их нельзя адекватно описать в пространстве только линейных функционалов.

Одним из важных нелинейных функционалов является функционал подобия распределений двух мер Хемминга или, что тоже самое, функционал расхождения распределений:

$$\eta_{1,2} = \frac{1}{2} \int_0^{\max(h)} |p_1(h) - p_2(h)| dh, \quad (19)$$

где $p_1(h)$ – первая из сравниваемых плотностей распределения значений; $p_2(h)$ – вторая из сравниваемых плотностей распределения значений.

При больших длинах кодов можно пренебрегать дискретным характером сравниваемых плотностей распределения значений меры Хемминга. Для малых длин кодов необходимо учитывать дискретный характер плотности распределения значений сравниваемых мер Хемминга:

$$\eta_{1,2} = \frac{1}{2} \sum_{h=0}^{\max(h)} |P_1(h) - P_2(h)|, \quad (20)$$

где $P_1(h)$ – первая из сравниваемых вероятностей появления значения меры Хемминга h ; $P_2(h)$ – вторая из сравниваемых вероятностей появления значения меры Хемминга h .

По своему результату вычисления по формулам (19) и (20) полностью совпадают и находятся в интервале от 0.0 до 1.0. В случае, если $\eta_{1,2} = 0.0$, то сравниваемые статистики полностью совпадают или являются подобными (когда длины сравниваемых кодов раз-

личны). Если $\eta_{1,2} = 1.0$, то сравниваемые статистики полностью не совпадают (не имеют общих значений меры Хемминга).

Следует подчеркнуть, что мера Хемминга и ее статистики могут быть вычислены только в том случае, если задан код, с которым происходит сравнение других кодов, например, код «Свой». В этом плане мера Хемминга между двумя образами «Чужой-1» и «Чужой-2» не может быть корректно определена, пока не найден код хотя бы одного центра двух сравниваемых множеств. Для корректного вычисления кода центра образа «Чужой» используется специальная процедура «проявки» его кода.

Если подавать на вход обученного нейросетевого преобразователя биометрия-код образы «Чужой», то на выходах преобразователя будут появляться коды, разряды которых будут иметь разную стабильность. Часть разрядов будет обладать очень высокой стабильностью, то есть в них состояние «0» или состояние «1» будет появляться с вероятностью от 0.9 до 1.0. Обычно таких разрядов мало. Разрядов с вероятностью появления знака «0» или «1» от 0.8 до 0.9 у образа «Чужой», как правило, оказывается больше. Для кодов «Чужой» хорошо обученного преобразователя биометрия-код ситуация, когда нестабильных разрядов больше чем стабильных, наиболее предпочтительна. У «идеального» преобразователя биометрия-код стабильных разрядов вообще не должно быть, то есть все разряды любого образа «Чужой» должны иметь равные вероятность появления состояний «0» и «1». Очевидно, что для любого разряда кода «Чужой» мы можем вычислить показатель его стабильности – s , который изменяется от 0.0 до 1.0 и вычисляется подобно нормированной мере Хемминга:

$$s = \frac{1}{2 \cdot \max(h)} \sum_{i=1}^{\max(h)} (P_i(0,1) - 0.5), \quad (21)$$

где $P_i(0,1)$ – вероятность появлений одного из наиболее часто встречающихся состояний «0» или «1» в i -том разряде нестабильного кода.

Для биометрических кодов образа «Свой» показатель стабильности (21) очень близок к единице. Для идеального преобразователя биометрия-код любой образ «Чужой» должен давать

$s \approx 0.0$. Для реальных нейросетевых преобразователей биометрия-код показатель средней стабильности близок к 0.3. Также, как и мера Хемминга, мера стабильности – s не учитывает какой из разрядов и обладает той или иной стабильностью. То есть, также как и мера Хемминга, мера стабильности может быть эффективно использована при статистическом исследовании нейросетевых преобразователей высоких и сверхвысоких размерностей. Эта мера подобна мере Хемминга тем, что позволяет резко снизить размерность вычислений и исследовать статистические свойства выходных кодов нейросетевых преобразователей.

Так как меры подобны и имеют одинаковое число компонент, на которых строятся (и та и другая меры строятся на анализе каждого разряда, сравниваемых кодов), мы можем построить третью меру, являющуюся их комбинацией. Формально новая мера строится как векторное произведение компонент двух предыдущих мер с последующим их суммированием:

$$sh_{1,2} = \sum_{i=1}^{\max(h)} (P(i_2) - 0.5) \cdot (i_1 \oplus i_2), \quad (22)$$

где i_1 – значение i -го разряда первого, заранее известного стабильного кода; i_2 – значение i -го разряда второго нестабильного кода, проявленное как наиболее часто встречающееся в примерах нестабильного кода; $P(i_2)$ – вероятность появления состояния i_2 в i -том разряде нестабильного кода.

Комбинационная мера (22) отражает стабильность совпадающих разрядов у сравниваемых кодов, не учитывая их расположение в коде. Очевидно, что можно построить аналогичную меру для оценки стабильности несовпавших у сравниваемых кодов разрядов:

$$s\bar{h}_{1,2} = \sum_{i=1}^{\max(h)} (P(i_2) - 0.5) \cdot (\bar{i}_1 \oplus i_2), \quad (23)$$

где $\bar{i}_1$ – инверсия i -го разряда первого стабильного кода, с которым сравнивается значение этого же i -го разряда нестабильного кода.

Прямая мера стабильности Хемминга (22) и инверсная мера стабильности Хемминга (23) дополняют друг друга и полностью совпадают, когда совпадают между собой значения прямой и инверсной мер Хемминга ($h = \bar{h}$) и когда стабильности совпавших и несовпавших разрядов сравниваемых кодов в среднем одинаковы. При необходимости меры (22) и (23) могут быть нормированы по аналогии с нормировкой (21). Так же очевидно, что для мер стабильности Хемминга могут быть построены плотности распределения их значений. На рисунке 20 приведены примеры двух распределений значений обычной меры Хемминга и меры стабильности кодов Хемминга для образов «Чужой-1», построенные относительно кода «Свой».

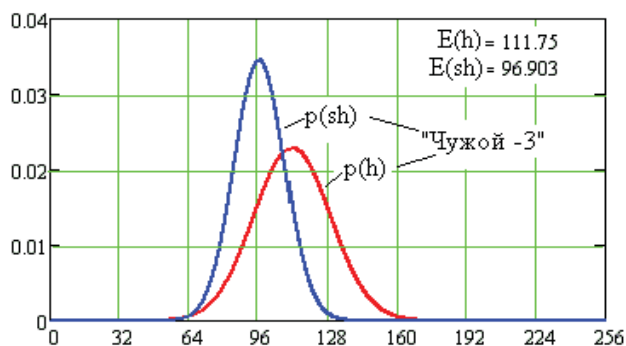


Рисунок 20 – Распределения значений обычной меры Хемминга – $p(h)$ и меры стабильности кодов Хемминга $p(sh)$ для образов «Чужой-3», построенные относительно кода «Свой»

Из данных рисунка 20 видно, что распределение значений меры Хемминга с учетом стабильности разрядов сравниваемых кодов всегда имеет меньшее значение математического ожидания и меньшее значение среднеквадратического отклонения. Эта общая закономерность. Мера Хемминга, взвешенная с учетом стабильности сравниваемых между собой кодов, дает более точные оценки вероятностей ошибок второго рода нейросетевых преобразователей биометрия-код.

Одним из важных свойств меры стабильности расстояний Хемминга является то, что они могут быть построены для сравнения двух нестабильных кодов. Например, если мы хотим сравнить между собой нестабильный код «Чужой-4» и «Чужой-5», то мы можем воспользоваться следующим выражением:

$$\sqrt{sh}_{4,5} = \frac{1}{N} \sum_{j=1}^N \sum_{i=1}^{\max(h)} \sqrt{(P(i_4) - 0.5)(P(i_5) - 0.5) \cdot (i_{4,j} \oplus i_{5,j})}, \quad (24)$$

где N- число пар кодов «Чужой-4» и «Чужой-5» использованных при вычислениях; $P(i_4)$ – вероятность появления наиболее вероятного значения в i-том разряде кода «Чужой-4»; $P(i_5)$ – вероятность появления наиболее вероятного значения в i-том разряде кода «Чужой-5»; $i_{4,j}$ – значение i-того разряда кода «Чужой-4» в j-той паре сравниваемых кодов; $i_{5,j}$ – значение i-того разряда кода «Чужой-5» в j-той паре сравниваемых кодов.

Примеры распределений значений обычной меры Хемминга и взвешенной двумя показателями стабильности меры Хемминга приведены на рисунке 21.

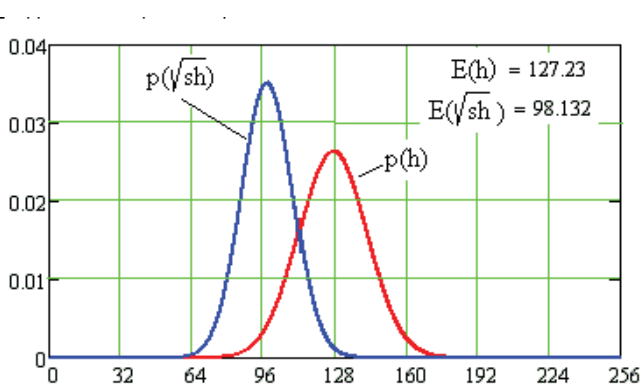


Рисунок 21 – Примеры распределений значений обычной меры Хемминга и взвешенной двумя показателями стабильности

Если сравнивать между собой данные рисунков 20 и 21, то становится очевидным большая эффективность сравнения меж-

ду собой двух нечетких кодов по формуле двойного взвешивания (24). Однократное взвешивание меры Хемминга показателями нестабильности одного из кодов всегда дает меньший сдвиг плотности распределения значений относительно исходной плотности распределения значений.

Таким образом, двойное взвешивание показателями нестабильности разрядов сравниваемых между собой кодов дает гораздо более точные результаты при статистических исследованиях. Естественно, что меры двойного взвешивания типа (24) могут быть обобщены на инверсные коды и пронормированы. Все это позволяет существенно расширить инструментарий статистического исследования нейросетевых преобразователей биометрия-код.

Проведенные эксперименты показывают, что при использовании любого варианта меры Хэмминга удастся получать оценки вероятности ложной аутентификации «Чужого» преобразователя биометрия-код с 256 выходами с относительной ошибкой предсказания менее одного бита. В исследованиях при моделировании использована мера Хэмминга без взвешивания мерой стабильности. При абсолютно достоверном «знании» закона распределения значений можно оценить погрешности статистических измерений вероятности P_2 .

Приведем примеры определения вероятности ложной аутентификации «Чужого» преобразователем биометрия-код по синтезированным таблицам биномиального зависимого закону распределения.

Оценим вероятность пропуска «Чужого» при $m\left(\frac{H}{n}\right) = 0,5$; $m(|r|) = 0,11$ по биномиальному зависимому закону распределения для преобразователя с 256 выходами. Находим по таблицам соответствующее распределение и значения плотностей вероятностей. По значению в точке $k=n$ получаем оценку вероятности ложной аутентификации «Чужого» как вероятность угадывания всех бит кода. В данном случае эта оценка равна $P_{\text{мод}} = 10^{-7,8}$. Ошибка оценки составляет менее одного бита.

Оценим вероятность пропуска «Чужого» при $m\left(\frac{H}{n}\right) = 0,45$; $m(|r|) = 0,15$ по биномиальному зависимому закону распределе-

ния для преобразователя с 256 выходами. Находим по таблицам соответствующее распределение и значения плотностей вероятностей. При $k=n$ получаем оценку вероятности ложной аутентификации «Чужого» как вероятность угадывания всех бит кода $P_{\text{мод}} = 10^{-6,1}$. Ошибка оценки составляет менее одного бита.

Полный вариант таблиц биномиального зависимого закона содержится в приложении 1 к монографии. Синтезированные таблицы плотности распределения значений меры Хэмминга кодов-откликов «Чужой» нейросетевого преобразователя биометрия-код с 256 выходами являются основой для разработки методики повышения достоверности автоматизированного прогнозирования вероятности ложной аутентификации «Чужого».

5 ОЦЕНКА РИСКОВ ЛОЖНОЙ АУТЕНТИФИКАЦИИ «ЧУЖОГО» ЧЕРЕЗ СРЕДНЕКВАДРАТИЧЕСКОЕ ОТКЛОНЕНИЕ НОРМИРОВАННОЙ МЕРЫ ХЭММИНГА И МАТЕМАТИЧЕСКОЕ ОЖИДАНИЕ МОДУЛЯ КОЭФФИЦИЕНТОВ ПАРНОЙ КОРРЕЛЯЦИИ

Наряду с численным моделированием закона распределения меры Хэмминга кодов-откликов «Чужой» обученного нейросетевого преобразователя биометрия-код необходимо получить аналитическое описание вероятности ложной аутентификации «Чужого». Для этого предложено перейти от математического ожидания модуля коэффициентов парной корреляции и среднеквадратического отклонения нормированной меры Хэмминга кодов-откликов «Чужой» к качеству принимаемых преобразователем решений. Оценку качества принимаемых преобразователем решений проводим согласно выражению:

$$q = \frac{m(H)}{\sigma(H)} = \frac{m\left(\frac{H}{n}\right)}{\sigma\left(\frac{H}{n}\right)}. \quad (25)$$

Целесообразно использовать предложенную в работе [9] оценку снизу вероятности ложной аутентификации «Чужого» на основе нормального закона распределения как функции от качества входных данных преобразователя. При использовании данной оценки был сделан переход от качества входных данных преобразователя, к качеству принимаемых преобразователем решений q .

Предложенная оценка снизу вероятности ложной аутентификации «Чужого» как функции от качества принимаемых преобразователем решений q вычисляется следующим образом:

$$P_2 = 0,5 - \Phi_0(q) . \quad (26)$$

Так как при расчетах по выражению (26) получаются очень малые значения порядка 10 в минус 12 степени, вычисления удобнее вести в двоичных логарифмах. При этом вероятность ложной аутентификации «Чужого» P_2 преобразуется в длину эквивалентного по стойкости к атакам подбора выходного кода преобразователя g . Предложенная оценка снизу длины эквивалентного по стойкости к атакам подбора выходного кода преобразователя как функции от качества принимаемых преобразователем решений q вычисляется следующим образом:

$$g_{\text{норм}}(q) = -\log_2(P_2) = -\log_2(0.5 - \Phi_0(q)) . \quad (27)$$

График предложенной в работе оценки снизу длины эквивалентного по стойкости выходного кода преобразователя g как функции от качества принимаемых преобразователем решений q в соответствии с выражением (27) приведен на рисунке 22 (нижняя кривая).

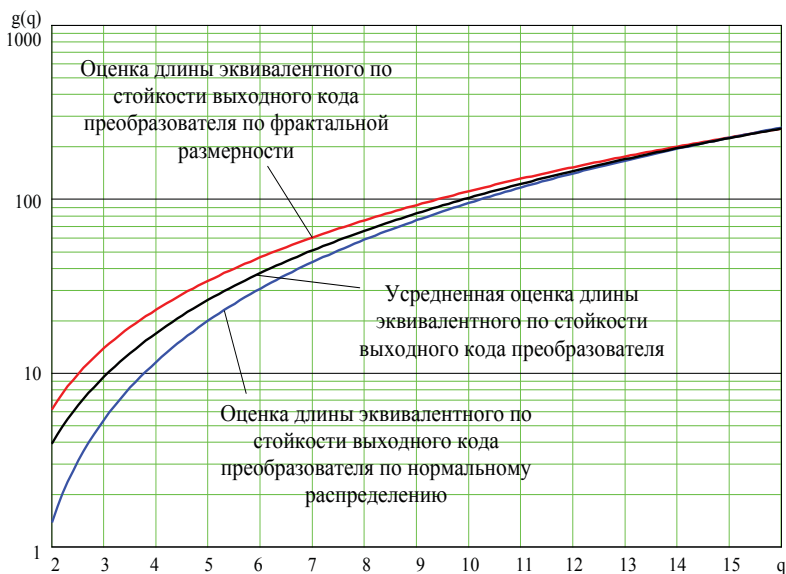


Рисунок 22 – График оценки длины эквивалентного по стойкости выходного кода преобразователя от качества принимаемых преобразователем решений $g(q)$

При вычислениях по выражению (3.3) ошибка оценки длины эквивалентного по стойкости выходного кода преобразователя g составляет порядка 8% от длины выходного кода преобразователя, то есть от 256 бит. Например, при математическом ожидании и среднеквадратическом отклонении нормированной меры Хэмминга кодов-откликов «Чужой», равным $m\left(\frac{H}{n}\right)=0,5$ и $\sigma\left(\frac{H}{n}\right)=0,056$ качество принимаемых преобразователем с 256 выходами решений составляет $q=8,929$. Оценка снизу длины эквивалентного по стойкости выходного кода преобразователя составляет $g(q)=96,6$.

Оценка снизу вероятности ложной аутентификации «Чужого» составляет $P_2=10^{-14,6}$.

Предложенная оценка длины эквивалентного по стойкости к атакам подбора выходного кода преобразователя g как функции от качества принимаемых преобразователем решений q на основе нормального закона распределения дает ошибку порядка 8% от длины выходного кода преобразователя, то есть от 256 бит. Поэтому в процессе выполнения возникла необходимость в повышении достоверности оценки длины эквивалентного по стойкости выходного кода преобразователя биометрия-код g . С этой целью была получена вторая оценка длины эквивалентного по стойкости выходного кода преобразователя g .

Установлено, что каждый из реальных преобразователей биометрия-код имеет свое предельное значение длины эквивалентного по стойкости выходного кода преобразователя g , которое получается меньше длины выходного кода преобразователя n . Получить оценки понижения длины эквивалентного по стойкости выходного кода преобразователя g математическими методами, применяемыми в классической криптографии, не удастся.

Поэтому возникла необходимость в привлечении аппарата теории фрактальной размерности для получения оценки длины эквивалентного по стойкости выходного кода преобразователя g [69]. С этой целью предложено ввести коэффициент понижения длины эквивалентного по стойкости выходного кода преобразователя биометрия-код. Предложенный в работе коэффициент

понижения длины эквивалентного по стойкости выходного кода преобразователя вычисляется следующим образом:

$$\alpha = \frac{g_{\text{фракт}}(q)}{n} . \quad (28)$$

При вычислениях по выражению (28) получаются дробные фрактальные значения длины эквивалентного по стойкости подбора выходного кода преобразователя. У биномиального распределения появляется дополнительная степень свободы – коэффициент понижения длины эквивалентного по стойкости выходного кода преобразователя α . Факториалы, которые определены только при целых значениях аргумента, заменяются на гамма-функции, позволяющие вычислять нецелые дробные фрактальные значения длины эквивалентного по стойкости выходного кода преобразователя g . Предложенная в работе оценка вероятности угадывания k разрядов из общего числа n разрядов выходного кода, соответствующая биномиальному зависимому закону распределения, составит:

$$P\left(n, m\left(\frac{H}{n}\right), \alpha\right) = \frac{\Gamma(\alpha n + 1)}{\Gamma(\alpha k + 1)\Gamma(n - k + 1)} \left(m\left(\frac{H}{n}\right)\right)^{\alpha k} \left(1 - m\left(\frac{H}{n}\right)\right)^{\alpha(n-k)} \quad (29)$$

Предложенная оценка длины эквивалентного по стойкости выходного кода преобразователя g как функции от качества принимаемых преобразователем решений q на основе теории фрактальной размерности вычисляется следующим образом:

$$g_{\text{фракт}}(q) = 0.0625q^2 . \quad (30)$$

График оценки длины эквивалентного по стойкости выходного кода преобразователя g как функции от качества принимаемых преобразователем решений q в соответствии с выражением (30) приведен на рисунке 22 (верхняя кривая).

Установлено, что вычисления по выражению (30) дают оценку сверху длины эквивалентного по стойкости выходного кода

преобразователя g . Получаемые на практике значения длины эквивалентного по стойкости выходного кода преобразователя g реальных преобразователей находятся в области, ограниченной полученными оценками (27) и (30). Поэтому предложено вычислять усредненную оценку длины эквивалентного по стойкости выходного кода преобразователя g . Усредненная оценка длины эквивалентного по стойкости выходного кода преобразователя g как функции от качества принимаемых преобразователем решений q вычисляется следующим образом:

$$g(q) = \frac{1}{2} g_{\text{норм}}(q) + \frac{1}{2} g_{\text{фракт}}(q) . \quad (31)$$

График усредненной оценки длины эквивалентного по стойкости выходного кода преобразователя g как функции от качества принимаемых преобразователем решений q в соответствии с выражением (31) приведен на рисунке 22 (средняя кривая). При вычислениях по выражению (31) ошибку вычисления длины эквивалентного по стойкости выходного кода преобразователя удается снизить с 8% до 1,2% от 256 бит.

Для повышения достоверности оценки длины эквивалентного по стойкости выходного кода преобразователя предложено сравнивать оценку (31) со значениями длины эквивалентного по стойкости выходного кода реальных преобразователей биометрия-код. График полученной в работе ошибки $\Delta g(q)$ как функции от качества принимаемых преобразователем решений q приведен на рисунке 23 (толстая кривая).

Для повышения достоверности оценки длины эквивалентного по стойкости выходного кода преобразователя g введен поправочный полином, вычисляемый следующим образом:

$$s(q) = -0,480 + 0,590q - 0,115q^2 + 0,005q^3 . \quad (32)$$

График зависимости (32) приведен на рисунке 23 (тонкая кривая). Оценка вероятности ложной аутентификации «Чужого» преобразователем биометрия-код P_2 как функции от качества

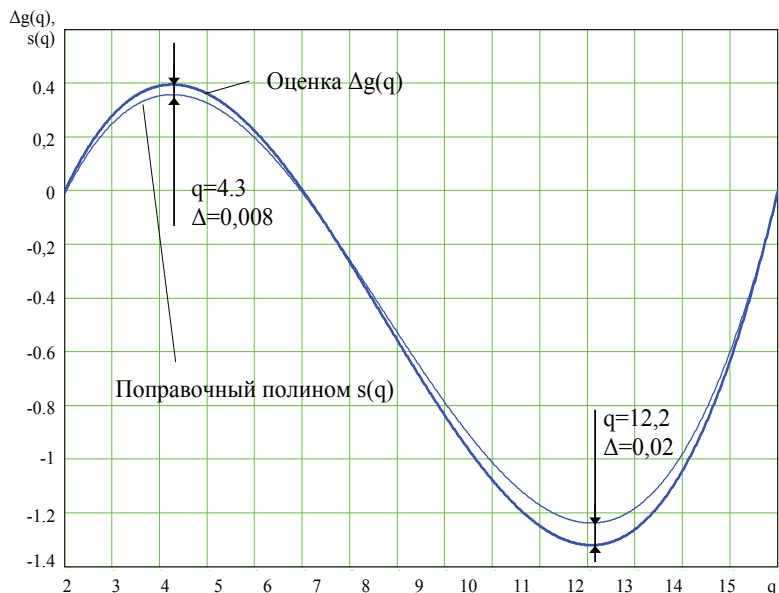


Рисунок 23 – Графики ошибки оценки длины эквивалентного по стойкости выходного кода преобразователя $\Delta g(q)$ и поправочного полинома $s(q)$

принимаемых преобразователем решений q , соответствующая биномиальному зависимому закону распределения, вычисляется следующим образом:

$$P_2(q) = 2^{g(q)+s(q)} = 10^{\frac{g(q)+s(q)}{3,3}}. \quad (33)$$

При вычислениях по выражению (33) ошибка вычисления длины эквивалентного по стойкости выходного кода преобразователя удастся снизить с 1,2% до 0,05% от 256 бит.

Полученные таблицы биномиального зависимого закона распределения и оценка рисков нарушения информационной безопасности по выражениям (32), (33) позволили разработать методику повышения достоверности автоматизированного прогнозирования вероятности ложной аутентификации «Чужого» нейро-

сетевым преобразователем биометрия-код при прогнозировании рисков средства высоконадежной биометрико-нейросетевой аутентификации. Эта методика позволила сократить приблизительно в 20 раз время выбора парольного слова для конкретного пользователя администратором безопасности защищенной автоматизированной системы. В отличие от ранее предложенных методик, рассчитанных на разработчика автоматизированной системы, разработанная методика ориентирована на администратора безопасности автоматизированной системы при выборе парольного слова для конкретного пользователя [70-72]. При этом снижается ошибка вычисления длины эквивалентного по стойкости выходного кода преобразователя с 8% до 0.05% от 256 бит. Алгоритм повышения достоверности автоматизированного прогнозирования вероятности ложной аутентификации «Чужого» обученным нейросетевым преобразователем биометрия-код по предложенной методике приведен на рисунке 24.

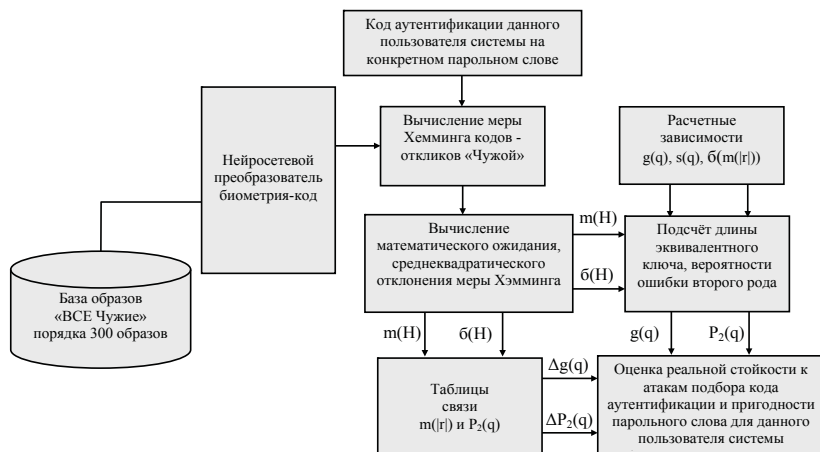


Рисунок 24 – Алгоритм повышения достоверности автоматизированного прогнозирования вероятности ложной аутентификации «Чужого» обученным нейросетевым преобразователем биометрия-код при управлении параметрами защиты информации

По методике повышения достоверности автоматизированного прогнозирования вероятности ложной аутентификации «Чужого» можно оценивать риски нарушения информационной безопасности при подборе сочетания «пользователь – парольное слово» и определять вероятность ложной аутентификации «Чужого». Например, при математическом ожидании и среднеквадратическом отклонении нормированной меры Хэмминга кодов-откликов

«Чужой», равным $m\left(\frac{H}{n}\right)=0,5$ и $\sigma\left(\frac{H}{n}\right)=0,06$ качество прини-

маемых преобразователем с 256 выходами решений составляет $q=8,367$. Оценка снизу длины эквивалентного по стойкости выходного кода преобразователя составляет $g(q)=78,9$. Оценка вероятности ложной аутентификации «Чужого» составляет $P_2=10^{-12,7}$.

Предложенная оценка длины эквивалентного по стойкости к атакам подбора выходного кода преобразователя g как функции от качества принимаемых преобразователем решений q дает ошибку порядка 0,05% от длины выходного кода преобразователя – 256 бит.

6 ИСПОЛЬЗОВАНИЕ ТАБЛИЦ БИНОМИАЛЬНОГО ЗАВИСИМОГО ЗАКОНА РАСПРЕДЕЛЕНИЯ ПРИ ИЗМЕРЕНИИ ЭНТРОПИИ ВЫСОКИХ ПОРЯДКОВ

Предложенный энтропийный подход является обобщением, описанной в 5 разделе методики, когда вычисляется вероятность ложной аутентификации «Чужого» через среднеквадратическое отклонение нормированной меры Хэмминга и математическое ожидание коэффициентов парной корреляции [73, 74]. При энтропийной оценке рисков нарушения информационной безопасности сначала сравниваются не два статистических момента, а форма закона распределения.

Для этого строится высокоразмерный наблюдатель энтропии, использующий частную меру Хэмминга при вычислении энтропии сочетаний бит выходного кода преобразователя [40]. Частная мера Хэмминга нужна при определении расстояния между кодами в ситуации, когда для сравнения берется не весь код, а только часть его длины. Определяется энтропия пар, троек, четверок, ..., сочетаний из 255 бит, сочетаний из 256 бит выходного кода через вычисление частных мер Хэмминга соответственно на кодах длины 2, 3, 4, ..., 255, 256 бит. Энтропия определяется по синтезированным таблицам биномиального зависимого закона с помощью нормирующей функции. В качестве такой функции может использоваться любая функция, задающая однозначный переход от гистограммы распределения к значению энтропии при задан-

ных параметрах $m\left(\frac{H}{n}\right)$ и $\sigma\left(\frac{H}{n}\right)$. Например, энтропия может на-

ходиться путем взвешивания по нормальному распределению и оценки фрактальности преобразователя биометрия-код по выражению (31), либо более точно оценена с учетом корректирующего полинома – выражение (32).

Проверка эффективности работы наблюдателя энтропии проведена на примерах определения энтропии текстов на русском и английском языках. Установлено, что с помощью высокоразмерного наблюдателя энтропии можно проводить измерение энтро-

пии преобразователя биометрия-код при наличии эталонов формы закона распределения в виде таблиц биномиального зависимого закона.

Одним из путей обеспечения гарантий конфиденциальности является доверенное тестирование стойкости обученного нейросетевого преобразователя биометрия-код к попыткам злоумышленников по извлечению из него конфиденциальной информации. При доверенном тестировании пользователь должен доверять проверяющему и сообщить ему свой код личного ключа. Далее, зная личный код пользователя, размещенный в нейросетевом преобразователе, проверяющий может вычислять значения меры Хэмминга между кодами «Свой» и «Чужой» и использовать эту информацию для ускоренного тестирования [67, 75]. Заметим, что все технологии, построенные на наблюдении распределения значений меры Хэмминга, оказываются неработоспособны, если пользователь не доверяет тестирующему человеку или тестирующему автомату свою тайну.

Рассмотрим задачу тестирования стойкости преобразователя биометрия-код к атакам подбора для ситуации, когда неизвестен размещенный в нем код и биометрический образ. Эта задача решается крайне просто в случае, когда преобразователь биометрия-код идеален (все его выходы независимы или некоррелированы). Тогда стойкость преобразователя биометрия-код определяется только числом его выходов, а каждый выход дает ровно один бит информации.

В том случае, когда между выходами нейронной сети появляются корреляционные связи, задача также может быть решена через вычисление энтропии, приходящейся на один выход нейросети. В частности, подобная задача легко может быть решена для текстов русского языка. При 8-битной кодировке кириллицы каждый одиночный бит русскоязычной фразы будет содержать примерно 0.56 бита реальной информации и 0.42 мнимой информации. Эти данные получены через вычисление энтропии одиночного символа русского языка [40] и делением этой величины на 8 (при 8-разрядной кодировке). Если учитывать пары разряды кодировок, принадлежащие паре двух символов в русском тексте,

то значение энтропии уменьшится. На рисунке 25 приведена кривая падения значения энтропии при учете 1, 2, 3, 4, ..., s разрядов в 1, 2, 3, 4, ..., s символах для русского языка. Очевидно, что кривая изменения энтропии $E(s)$ одного разряда 8-ми битных кодировок русских текстов оказывается много ниже прямой, соответствующей идеальному «белому шуму» в произвольной кодировке.

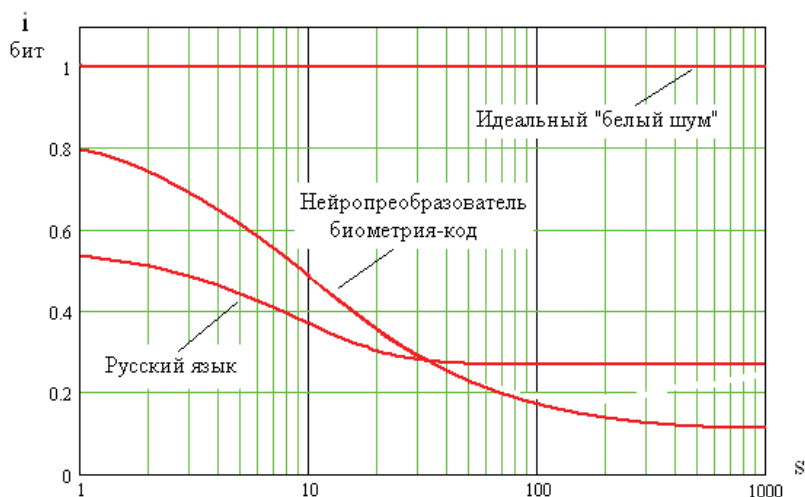
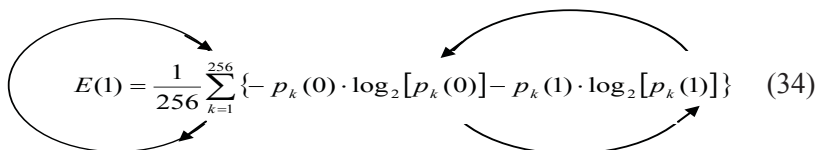


Рисунок 25 – Зависимость энтропии «белого шума», русского языка и нейросетевого преобразователя биометрия-код от числа учитываемых сочетаний символов (выходов)

Очевидно, что по аналогии с классическим учетом корреляционных связей русскоязычных кодировок или по аналогии с нейросетевым учетом этих связей, можно рассчитать энтропию нейросетевого преобразователя биометрия-код. Для этой цели необходимо осуществлять тестирование исследуемого нейросетевого преобразователя случайными биометрическими образами «Чужой».

По аналогии с расчетом энтропии одного бита одиночных символов естественного языка $E(s=1)$ для преобразователя биометрии в 256 разрядный код может быть использовано следующее выражение:

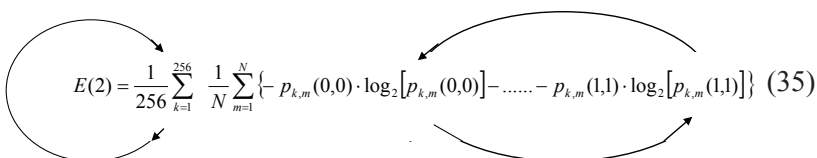


$$E(1) = \frac{1}{256} \sum_{k=1}^{256} \{-p_k(0) \cdot \log_2[p_k(0)] - p_k(1) \cdot \log_2[p_k(1)]\} \quad (34)$$

В выражении (34) k – номер выходного разряда нейросети; $p_k(0)$ – вероятность состояния «0» k -го выходного разряда нейросети; $p_k(1)$ – вероятность состояния «1» k -го выходного разряда нейросети.

В выражении (34) использовано условное отображение двух замкнутых колец. Правое кольцо соответствует обходу всех возможных состояний выходного разряда (их только два «0» и «1»). Левое кольцо соответствует обходу всех выходов нейросети с номерами k от 1 до 256.

По аналогии с (34) запишем формулу для вычисления энтропии с учетом всех возможных состояний 256 пар выходов нейросети:



$$E(2) = \frac{1}{256} \sum_{k=1}^{256} \frac{1}{N} \sum_{m=1}^N \{-p_{k,m}(0,0) \cdot \log_2[p_{k,m}(0,0)] - \dots - p_{k,m}(1,1) \cdot \log_2[p_{k,m}(1,1)]\} \quad (35)$$

В выражении (35) правое кольцо учитывает все 4 состояния пары разрядов с номерами k, m «0,0», «0,1», «1,0», «1,1». Левое кольцо отражает обход по k номерам выходов нейросети от 1 до 256.

Номера m выбираются случайно в интервале от 1 до 256 при выполнении условия $m \neq k$. Особо следует обратить внимание на то, что в выражении (35) отсутствует третье кольцо по всем возможным состояниям m . Это позволяет сделать вычисления (35) легко реализуемыми на практике для $N \leq 16$.

Очевидно, что выражение (35) легко может быть обобщено на тройки, четверки, пятерки разрядов $s = 3, 4, 5, 6, 7, \dots, 16$. При группе с числом анализируемых выходов менее 16 и числе

входных случайных примеров образов «Чужой» порядка 10000 особых требований к вычислительной технике не предъявляется. При учете состояний групп с числом выходов от 16 до 256 реализация полного правого кольца становится проблематичной. В связи с этим могут быть использованы три пути решения задачи вычислительной оценки собственной энтропии высокоразмерного нейросетевого преобразователя биометрия-код:

1. привлечение специальной высокопроизводительной вычислительной техники;

2. прореживание полей состояний, по кольцу которых производятся вычисления;

3. использование высокоразмерных наблюдателей распределения значений расстояний Хэмминга между образами «Чужой-1», «Чужой-2», ..., «Чужой-N» и базой образов «Все Чужие».

Наибольший интерес вызывает третий путь решения задачи, так как он является самым экономичным с точки зрения потребления вычислительных ресурсов. Очевидно, что задача подбора образа «Чужой-1» по своей сути близка к задаче подбора образов «Свой». То есть быстрые алгоритмы тестирования [1], построенные на использовании распределения меры Хемминга кодов «Свой» и «Все Чужие» теоретически могут быть заменены на быстрые алгоритмы, использующие распределения мер Хемминга:

«Чужой-1» и «Все Чужие»;

«Чужой-2» и «Все Чужие»;

.....

«Чужой-N» и «Все Чужие».

Чем лучше решает задачу защиты информации нейросетевой преобразователь биометрия-код, тем сложнее подбирать неизвестный образ «Свой» или неизвестный образ «Чужой-1». Между этими задачами существует явная пропорциональность. Тот факт, что распределение меры Хемминга «Свой» и «Чужой» близки между собой, иллюстрируется рисунком 26. Если вычислять энтропию преобразователя биометрия-код по мере Хэмминга «Чужой-2» вместо меры «Хэмминга» «Свой», то получим некоторую ошибку. Практика показала, что каждое из распределений меры Хемминга «Чужой-1», «Чужой-2», ..., «Чужой-N» будет

давать свою ошибку, и все эти ошибки имеют случайный характер. Усреднение энтропий, полученных по нескольким образам «Чужой», дает хорошее приближение энтропии преобразователя биометрия-код для образа «Свой». Именно это обстоятельство и позволяет тестировать преобразователь биометрия-код, не зная кода образов «Свой».

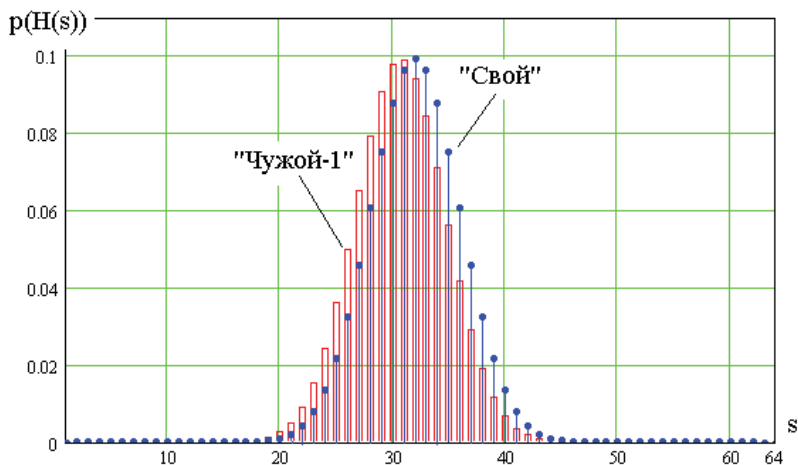
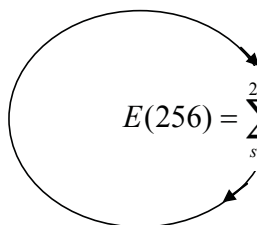


Рисунок 26 – Распределение значений мер Хэмминга для образов «Свой» и «Чужой-1» преобразователя биометрия-код с 64 выходами

Принципиальным преимуществом перехода от расчета энтропии по полю возможных состояний выходных кодов (34) и (35) к использованию дискретного поля возможных состояний меры Хэмминга является обход «проклятия размерности» [76, 77]. Пройти по кольцу всех кодов 2^{256} и вычислить вероятности их появления практически невозможно. Наоборот, пройти по кольцу 256 возможных состояний значений меры Хэмминга и рассчитать вероятности их появления технически не трудно. Расчет энтропии для преобразователя биометрия-код с 256 выходами осуществляется следующим образом:



$$E(256) = \sum_{s=1}^{256} \frac{1}{N} \sum_{m=1}^N -p_{s,m}(H(s)) \cdot \log_2[p_{s,m}(H(s))] \quad (36)$$

где s – дискретная переменная, описывающая возможные значения состояния меры Хэмминга; m – номер образа «Чужой», относительно которого вычисляется мера Хэмминга; $p_{s,m}(H(s))$ – вероятность появления кода с расстоянием Хэмминга – s .

Вероятность $p_{s,m}(H(s))$ может быть вычислена путем взвешивания оценки по нормальному распределению и оценки фрактальности преобразователя биометрия-код по выражению (31), либо более точно оценена с учетом корректирующего полинома – выражение (32).

На рисунке 27 показаны результаты вычисления частной меры Хэмминга по сочетаниям из 64 битов выходного кода преобразователя 1. Частная мера Хэмминга аналогична общей мере Хэмминга при равенстве длины сравниваемых кодов и размерности вычисления расстояния Хэмминга.

Частная мера Хэмминга определяет расстояние между кодами когда для сравнения берется не вся длина кода а только часть его длины. Определяется энтропия пар, троек, четверок, ..., сочетаний из 255 бит, сочетаний из 256 бит выходного кода через вычисление частных мер Хэмминга соответственно на кодах длины 2, 3, 4, ..., 255, 256 бит.

Экспериментально установлено, что частная мера Хэмминга топологически эквивалентна общей мере Хэмминга при условии проведения процедуры ортогонализации или при использовании кодов обладающих свойством псевдо ортогональности в пространстве кодов Хэмминга заданной длины.

На рисунке 28 показаны результаты вычисления частной меры Хэмминга по сочетаниям из 64 битов выходного кода преобразователя 2. Из данных рисунка 28 видно, что распределение смеще-

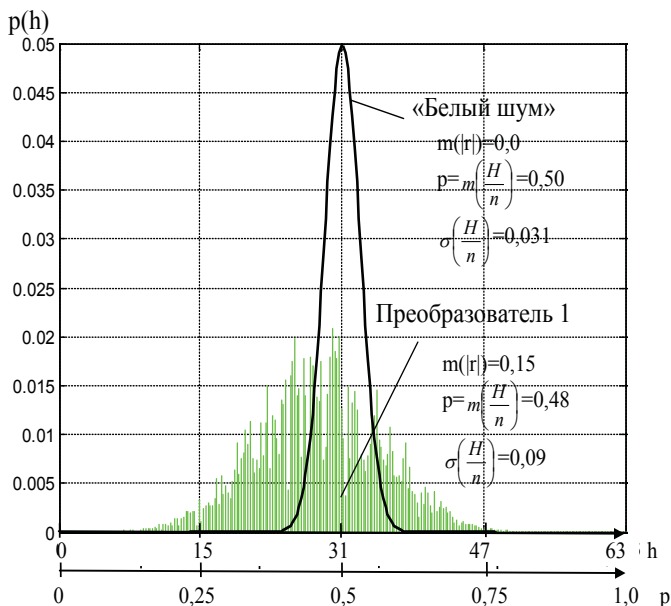


Рисунок 27 – Распределение частот появления кодов Хэмминга для последовательности из 64 битов выходного кода преобразователя 1 при $m(r)=0,15$; $p=m\left(\frac{H}{n}\right)=0,48$; $\sigma\left(\frac{H}{n}\right)=0,09$

но влево по сравнению с распределением преобразователя 1. Это говорит о том, что стойкость к атакам подбора кода преобразователя 1 выше по сравнению с соответствующим значением стойкости для преобразователя 2. Распределения меры Хэмминга при 256 бит выходного кода получаются как при тестировании преобразователя биометрия-код. Для оценки энтропии используется взвешивание вероятностей значений меры Хэмминга по параметру математическое ожидание нормированной меры Хэмминга. Чем дальше находится распределение относительно середины, тем меньше значение энтропии Хэмминга. Наоборот, чем больше энтропия преобразователя, тем ближе находится распределение к середине, то есть к характерному распределению для идеального преобразователя биометрия-код с единичной нормированной энтропией.

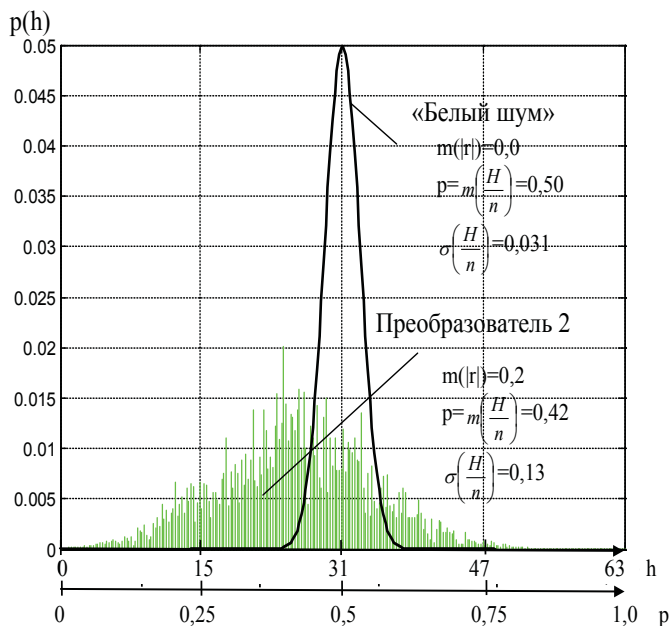


Рисунок 28 – Распределение частот появления кодов Хэмминга для последовательности из 64 битов выходного кода преобразователя 2 при

$$m(r)=0,20; p=m\left(\frac{H}{n}\right)=0,42; \sigma\left(\frac{H}{n}\right)=0,13$$

Из данных рисунка 29 видно, что идеальный «белый шум» всегда обладает максимально возможной энтропией (размерность наблюдателя не имеет значения). Энтропия преобразователя 1 с параметрами $m(r)=0,15; p=m\left(\frac{H}{n}\right)=0,48; \sigma\left(\frac{H}{n}\right)=0,09$ проходит выше кривой энтропии преобразователя 2 с параметрами $m(r)=0,20; p=m\left(\frac{H}{n}\right)=0,42; \sigma\left(\frac{H}{n}\right)=0,13$. Это говорит о том, что преобразователь 1 способен обеспечивать большую защиту от угроз нарушения информационной безопасности, чем преобразователь 2. Вероятность ложной аутентификации преобразователя 1 по результатам энтропийной оценки составляет

$P_2=10^{-8,6}$. Вероятность ложной аутентификации преобразователя 2 составляет $P_2=10^{-6,2}$. На рисунке 29 кривые энтропии для обоих преобразователей не находят пологого участка при $s < 255$ бит. Информативность кодов (обратный топологический эквивалент энтропии) рассмотренных преобразователей не полностью исчерпывается в этом диапазоне по s . Она падает до некоторого предельного значения информативности. У рассмотренных преобразователей это предельное значения информативности составляет порядка 500 бит, как это видно, например, из рисунка 25.

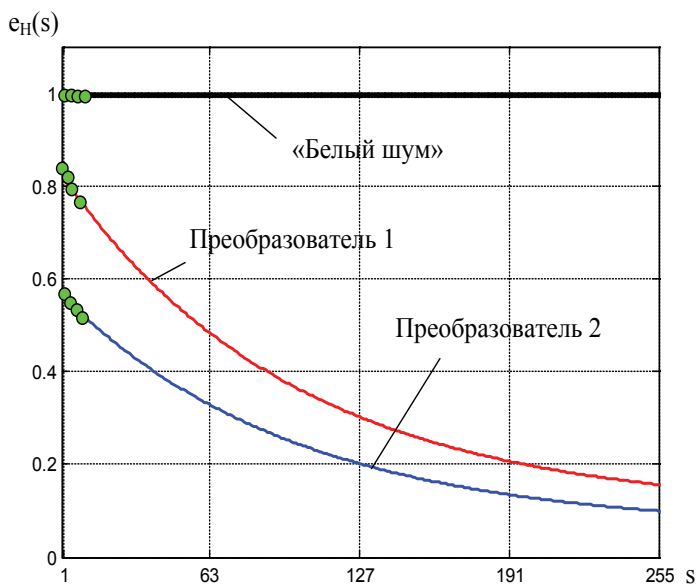


Рисунок 29 – Изменение относительного значения энтропии преобразователя биометрия-код, приходящееся на 1 бит их кодов для наблюдателей разной размерности (s -число одновременно наблюдаемых бит кодов Хэмминга)

Оценка рисков ложной аутентификации «Чужого» проводится через среднеквадратическое отклонение нормированной меры Хэмминга и математическое ожидание коэффициентов парной корреляции при условии знания кода «Свой».

Если код «Свой» неизвестен, то вероятность ложной аутентификации «Чужого» определяется с помощью высокоразмерного наблюдателя энтропии преобразователя биометрия-код с 256 выходами на базе 10000 образов. Высокоразмерный наблюдатель энтропии настраивается путем усреднения энтропий, полученных по нескольким образам «Чужой» и множеству «ВСЕ Чужие» по параметрам среднеквадратическое отклонение нормированной меры Хэмминга и математическое ожидание коэффициентов парной корреляции.

Проверить эффективность работы наблюдателя энтропии можно на примере вычисления энтропии кодов букв русского и английского языка на базе 100000 примеров кодов.

Рассмотрим классическую задачу измерения энтропии, какого-либо естественного языка. Если пользоваться классическим подходом, то потребуются тексты на исследуемом языке достаточно большого размера. В частности, при оценках энтропии русского языка по одиночным буквам потребуется текст, имеющий примерно 500 знаков. При учете вероятности появления пар букв потребуется уже текст с 50 000 знаков. Учет вероятностей появления сочетаний по 4-м буквам требует текст 5 000 000 знаков. Для учета сочетаний из 8 букв потребуются очень большие русскоязычные тексты из 500 000 000 знаков.

Таким образом, получается технологический барьер, для расчета энтропии русского языка (для сочетаний 16, 32, 64, 128 букв). Во-первых, в природе отсутствуют оцифрованные русскоязычные тексты нужных размеров, во-вторых, расчеты по большим и сверхбольшим базам русских текстов будут занимать значительное время (сотни лет машинного времени при условии привлечения суперкомпьютеров).

Сходные проблемы возникают в высоконадежной биометрии, когда на выходе нейросетевого преобразователя биометрия-код появляется своеобразный искусственный язык с существенными корреляционными связями между кодами различных биометрических образов. При этом проблема малых размеров тестовой выборки в биометрии оказывается намного острее, чем в лингвистике. В связи с этим по аналогии с измерением энтропии пре-

образователя биометрия-код целесообразно перейти от классических процедур расчета энтропии по вероятностям появления кодов букв в тексте к расчету вероятности появления значений расстояний Хэмминга между исследуемым кодом и другими возможными кодами языка.

Рассмотрим подробнее расчет энтропии русского языка через коды Хэмминга. Для определенности выберем часто встречающееся в языке сочетание пары букв «но», которое в кириллической кодировке будет соответствовать 16 разрядному коду. Очевидно, что расстояние Хэмминга между 16-ти разрядными кодами может принимать значение от 0 до 15. Если теперь взять начало классического романа И.А.Гончарова «Обрыв» размером 89 700 знаков и рассчитать на этом тексте относительные частоты появления кодов Хэмминга для сочетания букв «но» мы получим дискретное распределение, отображенное на рисунке 30.

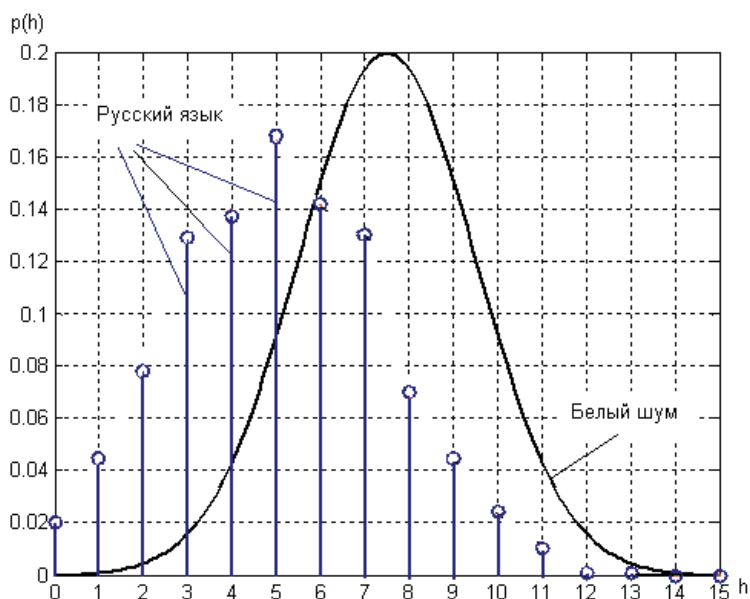


Рисунок 30 – Дискретное распределение значений кодов Хэмминга для сочетания букв «но» на русском тексте и «белом шуме»

Из частот появления кодов Хэмминга для сочетания букв «но» может быть рассчитана вероятность появления этого сочетания в русском языке. В нашем случае вероятность может быть вычислена по следующей формуле:

$$P_{\text{но}} = h_{0,\text{но}} \cdot \left(\sum_{i=1}^{15} h_{i,\text{но}} \right)^{-1}, \quad (36)$$

где $h_{i,\text{но}}$ – частота появления кода Хэмминга со значением i для сочетания букв «но».

Очевидно, что по формуле (36) может быть вычислена вероятность появления в тексте любых сочетаний пар букв кириллицы. Общее число возможных сочетаний пар букв в русском составляет $C_{32}^2 \approx 32^2$. Если по каждому j -тому сочетанию пар букв из возможных 32^2 сочетаний букв русского языка выполним операцию (36), то далее можно вычислить классическую энтропию языка пар букв:

$$E_2 = \sum_{j=1}^{32^2} -P_j \cdot \log_2(P_j). \quad (37)$$

Расчет по полному полю возможных состояний пар букв технически выполним для русского языка на тексте из 50 000 знаков, причем эти расчеты, сделанные как по кодам Хэмминга, так и по классической методике, полностью тождественны.

На обычной вычислительной технике обход кольца из 32^2 состояний занимает около 2 часов машинного времени.

Так же тождественными оказываются расчеты энтропии четверок возможных сочетаний букв, сделанные по классике и через коды Хэмминга:

$$E_4 = \sum_{j=1}^{32^4} -P_j \cdot \log_2(P_j). \quad (38)$$

Естественно, что обход кольца из 32^4 состояний потребует уже порядка 16 часов машинного времени и исходного текста из 5 000 000 букв.

Очевидно, что попытки рассчитать энтропию русского языка по 8, 16, 32 группам возможных сочетаний букв технически неосуществимы, если ориентироваться на классический способ расчета по выражениям (37), (38). Если же отказаться от классики и воспользоваться кодами Хэмминга, то задача оказывается технически выполнимой. Эта новая возможность обусловлена тем, что распределение значений кодов Хэмминга можно наблюдать для любых самых экзотических сочетаний букв, которые в реально языке встречаются архиредко. Происходит топологический переход от вычисления вероятности появления определенной комбинации символов к построению распределения кодов языка относительно этой комбинации.

Очевидно, что описанную выше процедуру можно повторить для других возможных сочетаний групп из 32 букв русского текста. При этом для каждого анализируемого сочетания группы из 32 букв в состоянии оценить вероятность его появления в тексте. Это в свою очередь означает, что можно рассчитать несколько тысяч оценок вероятностей и по ним построить приближенное значение энтропии:

$$E_{32} \cong \sum_{j=0}^{9999} -P_j \cdot \log_2(P_j) . \quad (39)$$

Возможность отказа при расчете энтропии от вычисления вероятности каждого из возможных сочетаний группы из 32 букв обусловлена тем, что энтропия по своей сути является своеобразным усреднением частных энтропий отдельных кодов.

То есть, усреднение по всему кольцу возможных состояний кодов вполне может быть заменено на усреднение по части кольца возможных состояний по выражению (39). Более того, если выбирать учитываемые коды не случайно, то удастся существенно снизить погрешность из-за усреднения по части кольца возможных состояний.

Одним из приемов снижения погрешности является выбор учитываемых кодов сочетаний букв, обладающих своеобразной статистической симметрией (псевдо ортогональностью). Удовлетворительные результаты получаются, если при расчетах выбираются коды, имеющие по отношению ко всем уже учтенным кодам расстояние Хэмминга близкое к половине максимально возможного расстояния. То есть при вычислении по формуле (39) желательно иметь коды с взаимным расстоянием Хэмминга близким к 128.

Достижение свойств псевдо ортогональности [44] при вычислении энтропии Хэмминга аналогично процедурам ортогонализации при синтезе алгоритмов быстрого обучения нейронных сетей. В [9] показано, что основной проблемой синтеза алгоритмов быстрого обучения нейронных сетей с большим числом входов и выходов является обеспечение независимости (некоррелированности) выходных данных при тестировании нейронной сети множеством случайных образов «Все Чужие». В монографиях [1, 9] показано, что нейроны однослойной или многослойной сети преобразователя биометрия-код могут быть обучены путем максимизации качества принимаемых ими решений.

Геометрическим смыслом алгоритма обучения через максимизацию качества является выталкивание множества образов «Свой» из центра гиперсферы «Все Чужие» на ее периферию при подборе весовых коэффициентов обучаемого нейрона. Очевидно, что выталкивание множества образов «Свой» на периферию гиперсферы может осуществляться по разным траекториям. На рисунке 31 приведены две возможные траектории выталкивания множества образов «Свой» из центра в две разные точки круга «Все чужие», образовавшиеся при обучении двух нейронов с двумя общими входами V_1, V_2 .

Из рисунка 31 видно, что разделяющие линии двух нейронов проходят через центр круга «Все Чужие» и перпендикулярны друг к другу (взаимно ортогональны). Это обстоятельство приводит к выполнению требований стандарта [39]. Выходы нейрона-1 и нейрона-2 некоррелированы (независимы) и их состояния «1», «0» равновероятны.

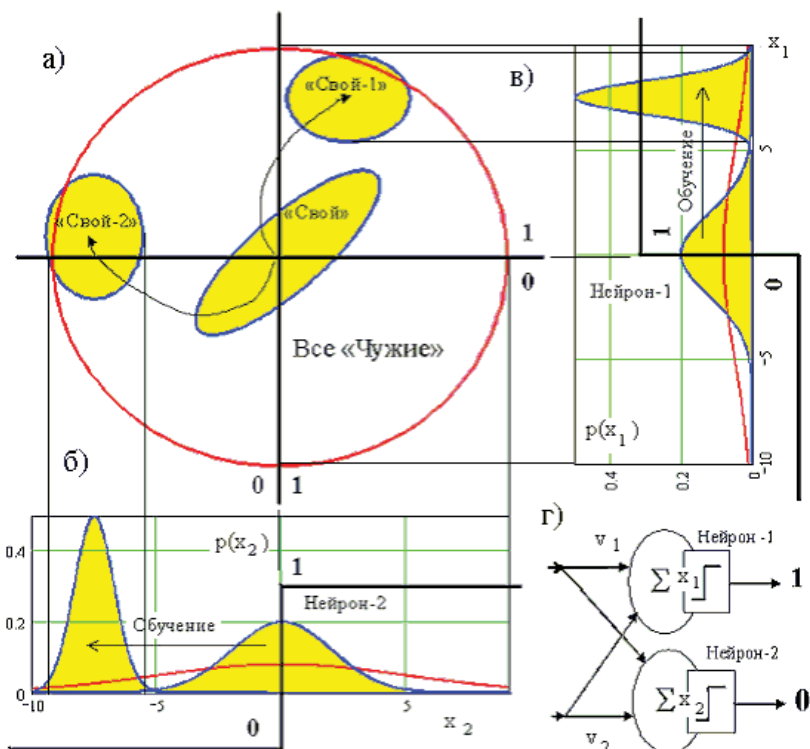


Рисунок 31 – Пример выполнения условий ортогонального обучения двух нейронов при их обучении алгоритмом выталкивания «Своего» из центра круга Все «Чужие»

Предположим, что решается задача обучения только для нейрона-1 и найдены два значения весовых коэффициентов « a_{11} » и « a_{12} », которые входят в следующее уравнение сумматора нейрона-1:

$$X_1 = a_{11} \cdot V_1 + a_{12} \cdot V_2 \quad (40)$$

В принципе можно повторить процедуру и искать весовые коэффициенты « a_{21} » и « a_{22} » нейрона-2 независимо от уже найденных весовых коэффициентов нейрона-1.

Обычно поиск весовых коэффициентов нейрона требует от 100 до 1000 итераций и занимает достаточно много времени. Сократить время настройки нейрона-2 можно путем простого вращения круга «Все Чужие» на четверть оборота.

Как известно операция вращения на угол $\pi/2$ системы координат сводится к простой смене местами весовых коэффициентов с изменением знака одного из них. То есть уравнение сумматора второго нейрона будет записываться через весовые коэффициенты нейрона-1 следующим образом:

$$X_2 = a_{12} \cdot V_1 - a_{11} \cdot V_2 . \quad (41)$$

Следует подчеркнуть, что операция вращения круга «Все Чужие» относительно его центра и операция независимого итерационного поиска весовых коэффициентов « a_{21} » и « a_{22} » нейрона-2 близки, но не тождественны. Наблюдается только примерное соответствие найденных коэффициентов:

$$\begin{cases} a_{21} \approx a_{12} \\ a_{22} \approx -a_{11} \end{cases} . \quad (42)$$

Проблема в том, что процедуры обучения нейронов и взаимной процедуры ортогонализации этих решений разделены. Сначала находят сами решения [8, 9], а уже потом пытаются сделать их независимыми, уменьшая корреляцию в группе соседних нейронов. Очевидно, что чем больше размерность отыскиваемых нейросетевых решений, тем сложнее оказывается задача их примитивной ортогонализации.

Например, если решается 24-х мерная задача (24 одинаковых входа у 24 нейронов), то необходимо найти 24 разных решений задачи, затем вычислить 24-мерную матрицу взаимных корреляций этих решений. Затем необходимо попытаться уменьшить значения модулей коэффициентов корреляции матрицы, например, итерационно корректируя каждое из независимо найденных 24 нейросетевых решений. Средствами современной вычислитель-

ной техники решить такую задачу удастся за несколько суток машинного времени.

Для того, чтобы сделать ортогональные алгоритмы обучения нейронных сетей быстрыми, необходимо сменить тактику обучения. Новая более эффективная тактика сводится к синтезу множества изначально ортогональных нейросетевых решений. По этой тактике ищется первое n -мерное решение (находится первая n -мерная разделяющая гиперплоскость), далее по этому первому решению находятся $(n-1)$ дополняющие его ортогональные решения. Дополняющие решения находятся путем синтеза $(n-1)$ дополнительных разделяющих гиперплоскостей ортогональных к первой найденной разделяющей гиперплоскости. Синтез взаимно ортогональных гиперплоскостей является тривиальной задачей, ее решение осуществляется путем умножения исходного вектора найденных весовых коэффициентов на соответствующие матрицы взаимно ортогональных вращений гипертесеры «Все «Чужие».

При вычислении энтропии Хэмминга выражения (40), (41), (42) могут быть использованы для достижения свойств псевдо ортогональности сравниваемых кодов. Ортогонализация в данном случае является одним из приемов понижения размерности вычислений при расчете энтропии Хэмминга. Еще одним способом понижения размерности вычислений является переход от сравнения двух моментов к сравнению большего числа моментов для повышения качества оценки рисков. Пытаясь понизить размеры базы «Все Чужие» при энтропийной оценке рисков получаем ситуацию, когда требуется учитывать не два статистических момента, а форму закона распределения при использовании эталонов в виде таблиц биномиального зависимого закона.

Повысить точность вычислений можно отказавшись от гипотезы нормальности закона плотности распределения значений кодов Хэмминга. В работе [9] показано, что гипотеза нормальности выполняется только для идеального «белого шума». В этом случае распределение частот кодов Хэмминга точно описывается биномиальным законом. Для зависимых (коррелированных) кодов расчеты следует проводить с учетом этого дополнительного фактора [34].

Изложенный выше подход к оценке энтропии русского и английского языков позволяет снять проблему конечной выборки исследуемого языка и отсутствия суперкомпьютеров. Нет нужды использовать всю выборку и большую вычислительную мощность, так как при переходе от бесконечного пространства всех возможных состояний к конечному множеству состояний меры Хэмминга можно получать оценки энтропии Хэмминга и вероятности ложной аутентификации «Чужого» на выборке порядка 100000 примеров кодов языка. При условии псевдо ортогональности сравниваемых кодов размер выборки удастся снизить до нескольких сотен примеров кодов языка. Распределение частот появления кодов Хэмминга для последовательности из 32 пар на русском языке при $m(|r|)=0,13$; $p=m\left(\frac{H}{n}\right)=0,38$; $\sigma\left(\frac{H}{n}\right)=0,07$ показано на рисунке 32.

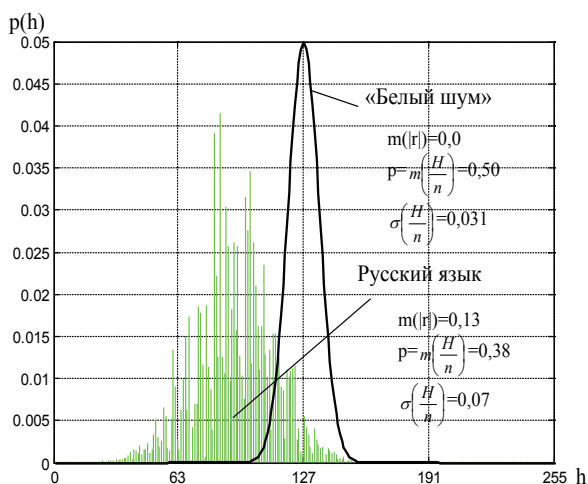


Рисунок 32 – Распределение частот появления кодов Хэмминга для последовательности из 32 пар на русском языке

при $m(|r|)=0,13$; $p=m\left(\frac{H}{n}\right)=0,38$; $\sigma\left(\frac{H}{n}\right)=0,07$

Из данных рисунка 32 видно, что коды Хэмминга имеют распределение значений с математическим ожиданием 78,3 и среднеквадратическим отклонением 19,4. В первом приближении можно считать распределение кодов Хэмминга нормальным. В этом случае можно рассчитать вероятность появления в гипотетическом русском тексте 16-ти кратного повторения «но». Эта вероятность составит величину $2.8 \cdot 10^{-5}$. Распределение частот появления кодов Хэмминга для последовательности из 32 пар на английском языке при $m(|r|)=0,19$; $p=m\left(\frac{H}{n}\right)=0,27$; $\sigma\left(\frac{H}{n}\right)=0,12$ показано на рисунке 33.

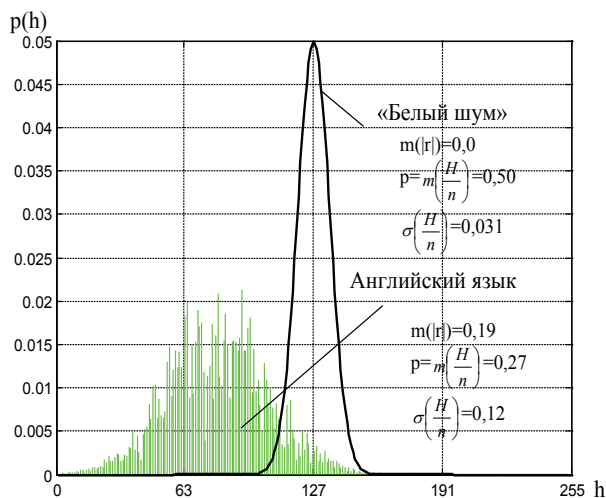


Рисунок 33 – Распределение частот появления кодов Хэмминга для последовательности из 32 пар на английском языке

$$\text{при } m(|r|)=0,19; p=m\left(\frac{H}{n}\right)=0,27; \sigma\left(\frac{H}{n}\right)=0,12$$

Из рисунка 33 видно, что коды Хэмминга имеют распределение значений с математическим ожиданием 69,12 и среднеквадратическим отклонением 30,72. В первом приближении можно считать распределение кодов Хэмминга нормальным. В этом случае можно рассчитать вероятность появления в гипотетическом

русском тексте 32-ти кратного повторения «но». Эта вероятность составит величину $3,4 \cdot 10^{-4}$. Английский язык имеет меньшее значение энтропии, чем русский примерно в полтора раза. Изменение относительного значения энтропии русского и английского языков, приходящееся на 1 бит их кодов для наблюдателей разной размерности (s -число одновременно наблюдаемых бит кода языка) показано на рисунке 34.

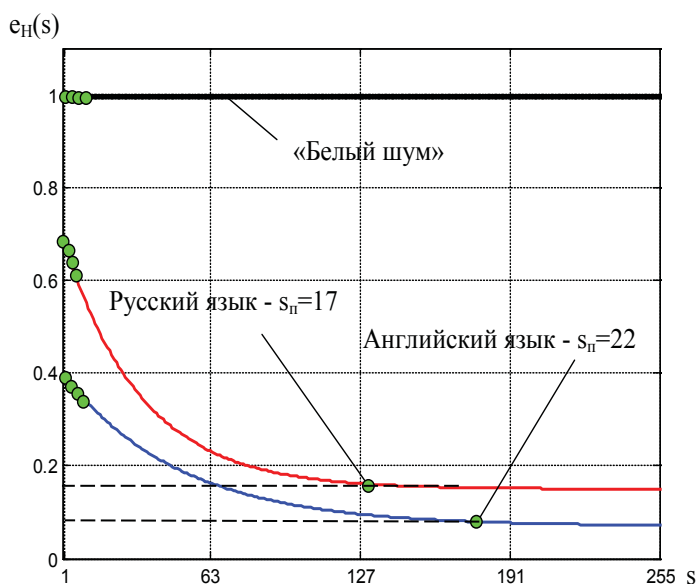


Рисунок 34 – Изменение относительного значения энтропии русского и английского языков, приходящееся на 1 бит их кодов для наблюдателей разной размерности (s -число одновременно наблюдаемых бит кода языка)

Из данных рисунка 34 видно, что идеальный «белый шум» всегда обладает максимально возможной энтропией (размерность наблюдателя не имеет значения). Энтропия русского языка перестает изменяться для групп с числом букв более 17 (размерность наблюдателя более $17 \cdot 8 = 136$). Как следствие показатель экспоненты энтропии русского языка составляет $17/3 = 6,4$ буквы (45,3 бита).

Для английского языка стабильный участок начинается после 22 буквы (показатель затухания экспоненты составляет 7,1 буквы или 56,8 бита). Результаты получены для классического английского текста Twain, Mark «Following the Equator» длиной 76 120 знака. То есть англоязычные пароли со смыслом должны иметь большую среднюю длину примерно в полтора раза, так как они имеют энтропию около 1,5 бит на один символ и меньше.

В конечном итоге переход от классических процедур вычисления энтропии к процедурам, использующим статистику распределения кодов Хэмминга, позволяет снять ранее существовавшие ограничения на размерность наблюдаемых групп букв языка. Удастся построить наблюдателей, учитывающих группы с 32 буквами и более. Подобные высокоразмерные наблюдатели способны работать с текстами достаточно малых размеров. На данный момент причин, по которым нельзя построить наблюдатели энтропии более высоких размерностей, нет. Для лингвистических исследований размерность наблюдателей энтропии нет смысла строить выше, чем 17 знаков (136 бит) для русского языка и 22 знака (176 бит) для английского языка.

Совершенно иная ситуация возникает при энтропийном исследовании нейросетевых преобразователей биометрия-код. Для исследования их выходных кодов могут потребоваться наблюдатели энтропии кодов Хэмминга размерностью 512 бит и выше.

ЗАКЛЮЧЕНИЕ

Авторами предложена модель статистического описания выходных состояний обученного нейросетевого преобразователя биометрия-код на основе биномиального зависимого закона распределения.

Установлено что используемая зарубежными исследователями гипотеза классического биномиального закона распределения независимых данных в качестве модели статистического писания преобразователя биометрия-код дает погрешности оценки стойкости от 70% при корреляции 0,15 до 100% при корреляции 1,0. Основной причиной является то, что зарубежные авторы не различают идеальный и реальный преобразователь, считая корреляции на выходах реального преобразователя незначительными [3].

Тестирование нейросетевого преобразователя биометрия-код с коррелированными связями между битами его выходного кода показывает, что гипотеза независимости данных полученных путем воспроизведения человеком своих личных биометрических образов не работает и дает искаженные результаты из-за естественной коррелированности биометрии человека и корреляции вносимой преобразователем биометрия-код. При увеличении числа входов и выходов преобразователя и использовании гипотезы независимости биометрических данных происходит накопление ошибки оценки вероятности ложной аутентификации «Чужого» с ростом корреляции битов выходного кода по сравнению с гипотезой независимости биометрических данных. При корреляции равной 0,5 топологическое независимое от размерности преобразователя превращение биномиального зависимого распределения в равномерное распределение статистических значений.

При корреляции большей 0,5 происходит перерождение биномиального зависимого распределения в двумодальный закон с минимальной плотностью распределения значений в центре.

Установлено, что при корреляции близкой к единице, ошибки оценки вероятности ложной аутентификации «Чужого» зарубежными авторами, по сравнению с оценками, полученными в рабо-

те, приближаются к 100% по сравнению с предлагаемой авторами моделью.

Разработанные модели и методики использованы в Российской Федерации при выполнении комплексного проекта по теме: «Разработка и подготовка производства телекоммуникационного оборудования, разработка программного сетевого, прикладного и специального обеспечения для создания цифровых сетей связи с персонализированным доступом» в соответствии с Постановлением Правительства РФ №218 от 09.04.2010 г. (Договор № 13G25.31.0039 от 07.09.2010 г.).

При выполнении научного проекта в Республике Казахстан по теме: №753. МОН.ГФ.13.13 «Исследование вариантов реализации и разработка действующего лабораторного образца ON-LINE системы биометрического обезличивания электронных историй болезней для медицинского учреждения», а также в ряде заявок на конкурсы инновационных научных проектов в России и Казахстане, научных публикациях и патентах [78-86].

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1 Волчихин В.И., Иванов А.И., Фунтиков В.А., Назаров И.Г., Язов Ю.К. Нейросетевая защита персональных биометрических данных. // – М.: Радиотехника, 2012. – 157 с. ISBN 978-5-88070-044-8. – Режим доступа: <http://www.radiotec.ru/catalog.php?cat=bo10&itm=170>

2 Смит Э.Р. Аутентификация от паролей до открытых ключей. – М.: издательство Вильямс, 2002. – 424 с.

3 Руд, Б. Руководство по биометрии. – М.: Техносфера, 2007. – 368 с. ISBN: 978-5-94836-109-3.

4 Воронина Н., Прохоров А., Семко Ю. Биометрические пароли. «Компьютер пресс». – 2002. – № 3. – С. 22-27.

5 Ефимов О.В., Иванов А.И. Программные хранители паролей с биометрическим доступом «Современные технологии безопасности». – 2002. – №2. – С. 30-33.

6 Ушмаев О.С. Биометрические технологии и защита информации. Мир измерений. – 2009. – №4 (98). – С. 4-10.

7 Чморра А.Л. Маскировка ключа с помощью биометрии «Проблемы передачи информации». – 2011. – №2 (47). – С. 128-143.

8 Волчихин В.И., Иванов А.И., Фунтиков В.А. Быстрые алгоритмы обучения нейросетевых механизмов биометрико-криптографической защиты информации: Монография. / – Пенза: Изд-во Пензенского государственного университета, 2005. – 273 с.

9 Ахметов Б.С., Волчихин В.И., Иванов А.И., Малыгин А.Ю. Алгоритмы тестирования биометрико-нейросетевых механизмов защиты информации: Монография. / – Алматы: КазНТУ имени К.И.Сатпаева, 2013. – 152 с. ISBN 978-101-228-586-4. – Режим доступа: <http://portal.kazntu.kz/files/publicate/2014-01-04-11940.pdf>

10 Иванов А.И. Нейросетевые алгоритмы биометрической идентификации личности. Книга 15, серии «Нейрокомпьютеры и их применение». – М.: Радиотехника, 2004. – 144 с. ISBN 5-93108-048-1. – Режим доступа: <http://www.radiotec.ru/catalog.php?cat=bo10&itm=92>

11 ГОСТ Р 52633.0–2006. Защита информации. Техника защиты информации. Требования к средствам высоконадежной биометрической аутентификации. – М.: Стандартинформ, 2007. – 24 с.

12 Волчихин В.И., Иванов А.И., Надеев Д.Н., Малыгин А.Ю. и др. Нейросетевое преобразование биометрического образа человека в код его личного криптографического ключа. – М.: Радиотехника (ИПРЖ), книга №29, научная серия «Нейрокомпьютеры и их применение», 2008.

– 87 с. ISBN 978-5-88070-168-1. – Режим доступа: <http://www.ozon.ru/context/detail/id/5022408/>

13 Dodis, Y. Extractors: How to Generate Strong Keys from Biometrics and Other Noisy Data / Y. Dodis, L. Reyzin, A. Smith. – 2004. – Режим доступа: www.cs.bu.edu/~reyzin/fuzzy.html

14 Juels, A. A fuzzy commitment scheme. In Sixth ACM Conference on Computer and Communications Security / A. Juels, M. Wattenberg. – New York : ACM Press, 1999. – P. 28-36.

15 Jain, A. K. Biometrics: A Tool for Information Security / A. K. Jain, A. Ross, S. Pankanti // IEEE transactions on information forensics and security. – 2006. – Vol. 1. – №2. – P. 125-143.

16 Uludag, U. Fuzzy Vault for Fingerprints. AVBPA 2005 : audio- and video-based biometric person authentication / U. Uludag, S. Pankanti, A. K. Jain. – Hilton Rye Town NY, 2005. – Vol. 3546. – P. 310-319.

17 Yang, S. Secure fuzzy vault based fingerprint verification system / S. Yang, I. Verbauwhede // In Thirty-Eighth Asilomar Conference on Signals, Systems, and Computers (2004). – 2004. – Vol. 1. – P. 577-581.

18 Nagar, S. Chaudhury. Biometrics based Asymmetric Cryptosystem Design Using Modified Fuzzy Vault Scheme / S. Nagar // 18th International Conference on Pattern Recognition (ICPR'06). – 2006. – P. 537-540.

19 Boyen, X. Reusable cryptographic fuzzy extractors, CCS / X. Boyen. ACM Press. – 2004. – P. 82-91. – Режим доступа: <http://ai.stanford.edu/~xb/ccs04/slides/index.html> – presentation slides

20 Hao, F. Combining Crypto with Biometrics Effectively / F. Hao, R. Anderson, J. Daugman // IEEE Transactions on Computers. – 2006. – Vol. 55. – №9. – P. 1081-1088. – Режим доступа: <http://www.cl.cam.ac.uk/TechReports/>

21 Sahai, A. Fuzzy identity based encryption / A. Sahai, B. Waters // in Proceedings of EUROCRYPT'05 on Advances in Cryptology, LNCS 3494. – Springer-Verlag, 2005. P. 457-473.

22 Материалы сайта ISO/IEC JTC 1/SC 37 – Биометрия. Режим доступа: http://www.iso.org/iso/iso_technical_committee?commid=313770

23 Ахметов Б.С., Надеев Д.Н., Урнев И.В. Сериков И.В. Аппроксимация биномиального зависимого закона композициями нормального, равномерного, арксинусного распределения значений. – М.: Радиотехника «Нейрокомпьютеры: разработка, применение», 2012. – №3. – С. 17-20.

24 Ахметов Б.С., Волчихин В.И., Куликов С.В., Малыгина Е.А. Моделирование длинных биометрических кодов, воспроизводящих корреляционные связи выходных данных нейросетевого преобразователя

«Нейрокомпьютеры: разработка, применение» – М.: Радиотехника, 2012. – №3. – С. 40-43.

25 Ахметов Б.С., Иванов А.И., Урнев И.В., Сериков И.В., Газин А.И. Оценка значений числа степеней свободы статистик описания выходного кода преобразователя биометрия-код при использовании распределения χ^2 . – Алматы: Изд-во КазНТУ имени К.И.Сатпаева. – Режим доступа: <http://portal.kazntu.kz/files/publicate/2013-04-05-elbib.pdf>

26 Иванов А.И. Перечень основных уязвимостей, возникающих при применении открытых неизменяемых биометрических образов в гражданских приложениях защиты персональной информации. Современные технологии безопасности. – 2007. – №4 (23). – С. 22-24.

27 Малыгин А.Ю., Воячек С.А. Оценка надежности и безопасности систем высоконадежной биометрической аутентификации личности // Вестник Костромского государственного университета им. Н.А. Некрасова. – 2007. – №1. – С. 88-92.

28 Малыгин А.Ю. Быстрые алгоритмы тестирования высоконадежной биометрической защиты. – М.: ВИМИ «Вопросы защиты информации», 2007. – №4 (79). – С. 8-12.

29 Малыгин А.Ю. Балансировка малых баз биометрических образов перед их тиражированием. – М.: ВИМИ «Вопросы защиты информации», 2007. – № 4(79). – С. 13-15.

30 Малыгин А.Ю. Формирование больших и сверхбольших баз биометрических образов для тестирования средств высоконадежной биометрической защиты – М.: ВИМИ «Вопросы защиты информации», 2007. – №4 (79). – С. 17-21.

31 Волчихин В.И., Артемов И.И., Безяев С.Б., Малыгин А.Ю., Алисов В.А. Особенности тестирования биометрико-нейросетевой аутентификации пользователей телекоммуникационной системы Министерства образования и науки Российской Федерации. – М.: Радиотехника «Нейрокомпьютеры: разработка, применение», 2012. – №3. – С. 49-51.

32 Малыгин А.Ю., Иванов А.И. О необходимости учета корреляционных связей при проведении тестирования многомерных биометрико-нейросетевых преобразователей. /Информационно-измерительная техника: межвузовский сборник научных трудов. – Пенза: Изд-во ПГУ, 2008. Вып. 33. – С. 110-114.

33 Надеев Д.Н., Иванов А.И. Связь коэффициента парной корреляции выходов нейропреобразователя биометрия-код со среднеквадратическим отклонением меры Хемминга образов «Чужие». – М.: Радиотехника «Нейрокомпьютеры: разработка, применение», 2007. – №12. – С. 27-29.

34 Надеев Д.Н. Синтез таблиц вероятности ошибок первого и второго рода для неидеальных биометрико-нейросетевых преобразователей с 256 выходами «Нейрокомпьютеры: разработка, применение». – М.: Радиотехника, 2007. – №12. – С. 30-31.

35 Надеев Д.Н., Фунтиков В.А., Иванов А.И. Связь энтропии выходных состояний нейросетевых преобразователей биометрия-код с коэффициентами парной корреляции «Нейрокомпьютеры: разработка, применение» – М.: Радиотехника, 2012. – №3. – С. 74-77.

36 Надеев Д.Н. Аналитическое описание оценки вероятности появления коллизий у нейросетевого преобразователя биометрия-код. – М.: Радиотехника «Нейрокомпьютеры: разработка, применение», 2009. – №6. – С. 53-55.

37 Малыгин А.Ю., Надеев Д.Н., Иванов А.И. Оценка погрешностей определения статистических моментов, обусловленных отсутствием представительной выборки // Труды научно-технической конференции «Безопасность информационных технологий». – Пенза: Изд-во ПНИЭИ, 2005. Том 6. – С. 60-61.

38 Малыгин А.Ю., Иванов А.И., Надеев Д.Н. Сокращение объемов тестовых выборок за счет знания закона распределения выходных состояний преобразователей биометрия-код // Труды научно-технической конференции «Безопасность информационных технологий». – Пенза: Изд-во ПНИЭИ, 2005. Том 6. – С. 62-64.

39 ГОСТ Р 52633.5-2011 «Защита информации. Техника защиты информации. Автоматическое обучение нейросетевых преобразователей биометрия-код доступа». – М.: Стандартиформ, 2011.

40 Фунтиков В.А. Энтропийный подход к тестированию стойкости нейросетевого преобразователя биометрия-код к атаке подбора. – М.: Радиотехника «Нейрокомпьютеры: разработка, применение», 2009. – №6. – С. 56-58.

41 Иванов А.И. Оценка корректности быстрых алгоритмов тестирования качества работы нейросетевых преобразователей биометрия-код. – Минск. «Электроника», 2013. – №6. – С. 43-45.

42 Ахметов Б.С., Захаров О.С., Иванов А.И., Картбаев Т.С., Малыгин А.Ю., Огнев И.В. Метод оценки вероятностей появления ошибок нейросетевых преобразователей биометрия-код, использующий очень малые тестовые выборки // Вестник КазНТУ им. К.И. Сатпаева. – Алматы, 2013. – №3 (97). – С. 279-283.

43 Иванов А.И. Нейросетевая защита личных биометрических данных. «Мир измерений». – 2009. – №4 (98). – С. 18-21.

44 Иванов А.И. Ортогонализация алгоритмов быстрого обучения искусственных нейронных сетей большой и сверхбольшой размерности. –

М.: Радиотехника. «Нейрокомпьютеры: разработка, применение», 2009. – №6. – С. 11-13.

45 Иванов А.И. Проект национального биометрического стандарта ГОСТ Р 52633.5, или Нейросетевой обход «проклятия размерности» очень плохих данных / «Защита информации. INSIDE», – 2010. – №5. – С. 48-56.

46 Иванов А.И. Быстрые алгоритмы оценки качества нейросетевых решений на малых тестовых выборках / «Защита информации. INSIDE». – 2011. – №6. – С. 43-46.

47 Малыгин А.Ю., Федулаев В.В., Надеев Д.Н. Требования к синтетическим базам биометрических образов и генераторам для их формирования. – М.: Радиотехника «Нейрокомпьютеры: разработка, применение», 2007. – №12. – С. 60-64.

48 Малыгин А.Ю., Волчихин В.И., Федулаев В.В., Безяев А.В. Оценка размеров технически реализуемых баз биометрических образов, необходимых для корректного тестирования высоконадежных нейросетевых преобразователей. – М.: Радиотехника. «Нейрокомпьютеры: разработка, применение», 2007. – №12. – С. 52-54.

49 Иванов А.И. Оценка остаточных корреляционных связей при тестировании нейросетевых преобразователей биометрия-код. – М.: Радиотехника. «Нейрокомпьютеры: разработка, применение», 2007. – №12. – С. 25-26.

50 Надеев Д.Н., Иванов А.И., Малыгин А.Ю. Статистическое описание идеального нейросетевого преобразователя биометрия-код: «информационный» вечный двигатель. – М.: Радиотехника. «Нейрокомпьютеры: разработка, применение», 2007. – №12. – С. 15-17.

51 Akhmetov B.S., Doszhanova A.A., Ivanov A.I., Kartbayev T.S., Malygin A.Yu. Dynamic biometric authentication of personality in open information spaces. // Word Academy of science, engineering and technology (WASET). – Singapore, 2013. issue 79. – P. 129-138.

52 Малыгин А.Ю., Надеев Д.Н., Иванов А.И. Две причины не идеальности нейросетевых преобразователей биометрия-код по выходным случайным состояниям кодов «Чужие». – М.: ВИМИ «Вопросы защиты информации», 2008. – №2 (81). – С. 19-21.

53 Надеев Д.Н., Иванов А.И. Модуль-нормальная связь коэффициентов парной корреляции со стойкостью нейросетевой защиты. – М.: ВИМИ «Вопросы защиты информации», 2008. – №3. – С. 35-38.

54 Ахметов Б.С., Иванов А.И., Малыгин А.Ю., Картбаев Т.С., Надеев Д.Н., Огнев И.В. Энтропийно-корреляционный подход к расчету вероятности совместного появления большого числа зависимых событий.

// Вестник Казахстанско-Британского технического университета. – Алматы, 2013. – №2 (25). – С. 54-58.

55 Ахметов Б.С., Иванов А.И., Малыгин А.Ю., Картбаев Т.С., Надеев Д.Н., Огнев И.В. О связи высокоразмерной энтропии и высокоразмерной корреляции с математическим ожиданием модулей коэффициентов парной корреляции. // Вестник ВКГТУ. – Усть-Каменогорск, 2013. – Ч.1. – С. 240-244.

56 Андреев Д.Ю., Иванов А.И., Захаров О.С., Хозин Ю.В. Модификация меры Хемминга через взвешивание мерой стабильности выходных данных нейросетевых преобразователей биометрия-код. – М.: Радиотехника. «Нейрокомпьютеры: разработка, применение», 2009. – №6. – С. 49-52.

57 Надеев Д.Н., Иванов А.И. Оценка аддитивной погрешности замены биномиального закона распределения на нормальный закон при исследованиях преобразователей биометрия-код ключа доступа // Труды научно-технической конференции «Безопасность информационных технологий». – Пенза: Изд-во ПНИЭИ, 2005. – Том 6, – С. 8-9.

58 Захаров О.С., Надеев Д.Н., Иванов А.И., Тришин А.В. Контроль работоспособности механизма центрирования входных данных преобразователя биометрия-код // Труды научно-технической конференции «Безопасность информационных технологий». – Пенза: Изд-во ПНИЭИ, 2005. – Том 6. – С. 13-15.

59 Иванов А.И., Надеев Д.Н. Оценка вероятностей состояний выходных разрядов преобразователя биометрия-код: моделирование закона распределения // Труды научно-технической конференции «Безопасность информационных технологий». – Пенза: Изд-во ПНИЭИ, 2005. – Том 6. – С. 41-45.

60 Иванов А.И. Захаров С.М. Идеальная машина обучения нейросетевых приложений, технически реализуемые стратегии наращивания уровня сложности нейросетевого искусственного интеллекта. – М.: Радиотехника «Нейрокомпьютеры: разработка, применение», 2007. – №12. – С. 9-11.

61 Капитуров Н.В., Хозин Ю.В., Иванов А.И. Оценка потенциальных возможностей идеальной машины обучения, способной усваивать «плохие» и «очень плохие» данные. – М.: Радиотехника «Нейрокомпьютеры: разработка, применение», 2007. – №12. – С. 12-14.

62 Надеев Д.Н. Моделирование биномиального зависимого закона распределения значений вероятностей ошибок нейросетевых преобразователей для высоконадежной биометрической защиты. – М.: ВИМИ «Вопросы защиты информации», 2008. №3. – С. 31-35.

63 Волчихин В.И., Малыгин А.Ю., Иванов А.И., Фунтиков В.А. Законодательно метрологические предпосылки для осуществления корректных измерений стойкости нейросетевых преобразователей биометрия-код к атакам подбора. – М.: Радиотехника «Нейрокомпьютеры: разработка, применение», 2007. – №12. – С. 73-79.

64 Волчихин В.И., Малыгин А.Ю., Лупанов М.Ю., Семенов В.А. Планирование численных экспериментов по тестированию высоконадежных биометрико-криптографических средств защиты. – М.: ВИМИ «Вопросы защиты информации», 2005. – №4. – С. 17-22.

65 Олейник Ю.И., Огнев И.В., Малыгин А.Ю. Пути экономии ресурсов при ускоренном тестировании средств высоконадежной нейросетевой биометрии. – М.: Радиотехника «Нейрокомпьютеры: разработка, применение», 2009. – №6. – С. 68-71.

66 Фунтиков В.А. Оценка «нано» и «пико» вероятностей появления редких событий биометрической аутентификации на малых тестовых выборках. – М.: Радиотехника «Нейрокомпьютеры: разработка, применение», 2012. – №3. – С. 21-24.

67 Захаров С.М., Иванов А.И. Обучение искусственных нейронных сетей на предельно малых выборках из одного-двух примеров. – М.: Радиотехника «Нейрокомпьютеры: разработка, применение», 2007. – №12. – С. 31-33.

68 Федулаев В.В., Малыгин А.Ю., Надеев Д.Н., Иванов А.И. Экспресс тестирование нейросетевых преобразователей после их обучения. – М.: Радиотехника «Нейрокомпьютеры: разработка, применение», 2007. – №12. – С. 34-37.

69 Майоров А. В., Захаров О.С., Иванов А.И., Малыгин А.Ю., Федулаев В.В. Тестирование стойкости биометрико-нейросетевой защиты информации посредством искусственного ослабления. – М.: Радиотехника «Нейрокомпьютеры: разработка, применение», 2007. – №12. – С. 34-37.

70 Иванов А.И., Надеев Д.Н., Захаров О.С. Измерение показателя фрактальности преобразователей биометрия-код. Современные технологии безопасности. – 2007. – №4 (23) – С. 36-38.

71 Ахметов Б.С., Иванов А.И., Малыгин А.Ю., Картбаев Т.С., Мукапил К. // Таратылған ақпараттық жүйелердегі қолданушылардың биометриялық идентификациясының моделдері мен әдістері // Вестник КазНТУ им. К.И. Сатпаева. – Алматы, 2013. – №5 (99). – С. 33-37.

72 Захаров О.С., Хозин Ю.В., Иванов А.И., Капитуров Н.В. Оценка технических ограничений, накладываемых на размерность выходных кодов нейросетевых преобразователей рукописных паролей. – М.: ВИМИ «Вопросы защиты информации», 2008. – № 3. – С. 27-31.

73 Akhmetov B.S., Ivanov A.I., Malygin A.Yu., Kartbayev T.S., Mukapil K. Biometric Dynamic Personality Authentication in Open Information Space // International Journal of Computer Technology and Applications. – India, 2013. – Vol. 4, Issue 5. – P. 846-855.

74 Иванов А.И., Фунтиков В.А., Майоров А.В., Надеев Д.Н. Моделирование кодовых последовательностей с энтропией естественных и искусственных биометрических языков // Инфокоммуникационные технологии. – 2010. – Том 8. – №4. – С. 75-79. – Режим доступа: <http://ikt.psuti.ru>

75 Иванов А.И., Майоров А.В., Язов Ю.К. Энтропийно-корреляционная оценка хэширующих свойств нейросетевого преобразователя биометрия-код доступа. – М.: Радиотехника «Нейрокомпьютеры: разработка, применение», 2012. – №3. – С. 35-39.

76 Майоров А.В., Шашков Б.Д., Урнев И.В., Сауанова К.Т., Малыгин А.Ю. Повышение качества работы преобразователей биометрия-код с помощью регуляции избыточности // В сборнике трудов Международного симпозиума «Надежность и качество – 2012». – Пенза: Изд-во ПГУ, 2012. Том 1. – С. 290-292.

77 Иванов А.И. Нейросетевое преодоление «проклятия» размерности, выход на «благодать» высокой размерности биометрических данных // «Защита информации. ИНСАЙД». – 2007. – № 5. – С. 50-56.

78 Akhmetov B.S., Doszhanova A.A., Malygin A.Yu., Ivanov A.I., Kartbayev T.S. Biometric depersonalization electronic medical patients illness stories // The 6th International Conference on Public Safety, Republic of Poland, Poznan, 2013. – P. 150-153.

79 Ахметов Б.С., Досжанова А.А., Иванов А.И., Малыгин А.Ю. Биометрический аудит как средство усиления защиты ERP систем // Материалы II Международной научной конференции «Высокие технологии – залог устойчивого развития». – Алматы: Изд-во КазНТУ им. К.И. Сатпаева, 2013. – С. 287-288.

80 B. Akhmetov, A. Ivanov, I. Urnev, V. Funtikov, A. Malygin. Statistical description of output states of the neural network “biometrics-code” transformers // Progress in Electromagnetics Research Symposium» (PIERS, Moscow, 2012). Published by The Electromagnetics Academy 777 Concord Avenue, Suite 207, Cambridge. – P. 61-65.

81 Иванов А.И., Ахметов Б.С., Безяев В.С., Сериков И.В., Малыгин А.Ю. Сравнительная оценка уровня безопасности средств биометрической аутентификации // В сб. трудов II Международной НПК «Информационно-инновационные технологии: интеграция науки, образования и бизнеса». – Алматы: Изд-во КазНТУ им. К.И. Сатпаева, 2011. Том 1. – С. 382-386.

82 B. Akhmetov, V. Volchihin, A. Ivanov, T. Kartbayev, A. Malygin. Highly reliable multi-biometric authentication of human-being personality to support citizens interaction with E-government and E-business // 3rd world conference on information technology. 14-16 november, 2012, University of Barcelona, Faculty of Library and Information Science, Barcelon, Spain. – P. 74-81.

83 Волчихин В.И., Безяев В.С., Иванов А.И., Малыгин А.Ю. Разработка программного сетевого, прикладного и специального обеспечения для создания цифровых сетей связи с персонализированным доступом // В сб. 7 Саратовского салона изобретений, инноваций и инвестиций. – Саратов: Изд-во СГУ, 2012. – С. 20-21.

84 Иванов А.И., Кисляев С.Е., Гелашвили П.А. Искусственные нейронные сети в биометрии, медицине, здравоохранении. – Самара: ООО «Офорт», 2004. – 236 с.

85 Патент RU 2346397 «Способ защиты персональных данных биометрической идентификации и аутентификации» / Иванов А.И., Фунтиков В.А., Ефимов О.В.

86 Патент RU 2473125 «Способ классификации рисунков отпечатка пальца при аутентификации личности» / Волчихин В.И., Фунтиков В.А., Иванов А.И., Секретов М.В., Майоров А.В., Малыгин А.Ю. <http://elibrary.ru/item.asp?id=18859157>

Приложение А

Таблицы вероятностей биномиального зависимого закона распределения меры Хэмминга выходного кода многомерного высоконадежного нейросетевого преобразователя биометрия-код

Пояснения к таблицам

В таблицах приведены вероятности значений меры Хэмминга выходного кода многомерного высоконадежного нейросетевого преобразователя биометрия-код, определенные с помощью биномиального зависимого закона распределения.

Таблицы получены при числе выходных разрядов нейросетевого преобразователя биометрия-код (НПБК) $n=256$ (длина выходного кода НПБК $n=256$).

Диапазон изменений параметра вероятность появления состояния «1» в разряде кода НПБК p от 0,5 (нескомпрометированный ключ пользователя или атаки подбора при неизвестном ключе) до $p=0,3$ (частичная компрометация или тестирование методом ослабления ключа). Шаг по p равен 0,05. Значения по p , отраженные в таблицах – 0.5; 0.45; 0.4 ; 0.35; 0.3.

Диапазон изменений параметра корреляция между разрядами кода НПБК r от 0 (когда наблюдается совпадение биномиального зависимого закона с классическим биномиальным законом распределения) до $r=0,5$ (когда происходит топологическое преобразование биномиального зависимого закона распределения в равномерный закон). Шаг по r равен 0,05. Значения по r , отраженные в таблицах – 0; 0.01; 0.02; 0.03; 0.04; 0.05; 0.1; 0.15; 0.2; 0.25; 0.3; 0.35; 0.4; 0.45; 0.5

В таблицах использованы обозначения:

p – вероятность появления состояния «1» в разряде кода НПБК;

n – длина выходного кода НПБК;

r – корреляция между разрядами кода НПБК.

P_{mod} – значение вероятности биномиального зависимого закона ($\cdot 10^{-2}$);

E_{norm} – ошибка аппроксимации нормальным законом ($\cdot 10^{-3}$);

$E_{фракт}$ – ошибка фрактальной аппроксимации ($\cdot 10^{-3}$).

$P_{ид}$ – значение вероятности классического биномиального закона;

E – ошибка аппроксимации классическим биномиальным законом.

В каждой таблице указаны параметры биномиального зависимого закона, например:

Параметры биномиального зависимого закона $p = 0,3$; $n = 256$; $r = 0,01$

В каждой таблице приводится график функции плотности биномиального зависимого закона распределения при заданных параметрах p, n, r (рисунок А.1).

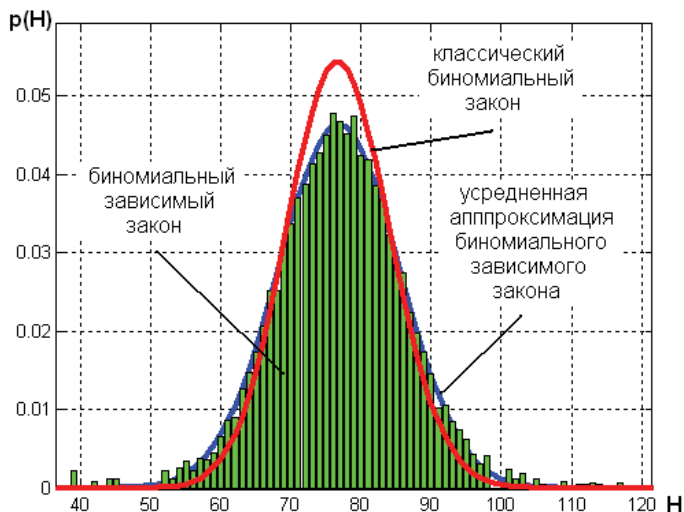


Рисунок А.1 – График функции плотности биномиального зависимого закона распределения при $p = 0,3$; $n = 256$; $r = 0,01$

Значения вероятностей биномиального зависимого закона распределения меры Хэмминга выходного кода НПБК при данных значениях параметров p, n, g приводятся в зависимости от длины угаданной «Чужим» части ключа НПБК m , как в таблице А.1. Диапазон изменений параметра длины угаданной «Чужим» части ключа НПБК m от 0 (оценивается вероятность ситуации, когда «Чужому» не удастся угадать ни один разряд выходного кода НПБК) до $m=n$ (оценивается вероятность ошибки пропуска «Чужого», когда «Чужому» удастся угадать всю длину ключа НПБК). Шаг по m равен 1. Значения по m , отраженные в таблицах – $m=0,1,2\dots256$.

**Таблица А.1 – Вероятности биномиального зависимого закона распределения меры Хэмминга
выходного кода НПКБ при $p = 0,3$; $n = 256$; $r = 0,01$ на половине длины ключа $n/2$ и длины ключа n**

H	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
$R_{\text{под}}$	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	
$E_{\text{ном}}$	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	
$E_{\text{фракт}}$	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	

H	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64
$R_{\text{под}}$	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,1	0,1	0,1	0,2	0,2	0,3	0,4	0,5	0,7	0,9	1,1	1,3	1,5	
$E_{\text{ном}}$	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	
$E_{\text{фракт}}$	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	-0,6	1,4	2,3	-1,1	0,3	-2,3	-3,0	1,0	-2,1	-1,0	2,1	0,5	-4,4	2,6

H	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96
$R_{\text{под}}$	1,8	2,1	2,4	2,7	3,1	3,4	3,7	4,0	4,2	4,4	4,5	4,6	4,6	4,6	4,5	4,3	4,1	3,9	3,6	3,3	2,9	2,6	2,3	2,0	1,7	1,4	1,2	1,0	0,8	0,5	0,4	
$E_{\text{ном}}$	1,6	-1,3	1,3	-0,2	-0,1	-2,8	-0,9	0,3	-2,4	-1,0	2,9	-2,8	-1,4	0,8	-3,9	-2,0	-1,1	0,2	-1,3	0,7	1,7	2,7	-3,4	-2,2	0,5	-2,6	-1,1	-0,8	-1,3	1,1	0,8	
$E_{\text{фракт}}$	2,9	-2,9	2,6	-0,3	-0,1	-5,0	-2,0	0,7	-5,4	-2,5	7,0	-5,2	3,5	1,9	-7,7	-3,7	-2,5	0,3	4,9	-2,8	1,3	3,9	5,0	-5,8	-4,2	1,1	-5,4	-2,7	-9,4	-2,3	2,5	1,8

H	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128
$R_{\text{под}}$	0,3	0,2	0,2	0,1	0,1	0,1	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	
$E_{\text{ном}}$	-0,3	3,3	-2,6	0,1	-0,9	3,3	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	
$E_{\text{фракт}}$	-0,6	6,1	-5,1	0,3	-1,8	7,7	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	

H	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160
$R_{\text{под}}$	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	
$E_{\text{ном}}$	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	
$E_{\text{фракт}}$	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	

H	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192
$R_{\text{под}}$	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	
$E_{\text{ном}}$	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	
$E_{\text{фракт}}$	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	

H	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224
$R_{\text{под}}$	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	
$E_{\text{ном}}$	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	
$E_{\text{фракт}}$	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	

H	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256
$R_{\text{под}}$	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	
$E_{\text{ном}}$	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	
$E_{\text{фракт}}$	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	

Примечания:

1. Моделирование зависимого биномиального закона и вычисление вероятностей ошибок пропуска «Чужого» проведено на 1,000,000 промерах;

2. Полученные в результате численного моделирования данные обработаны двумерной сглаживающей функцией;

3. Диапазон по вероятности появления состояния «1» в разряде кода НПБК p составляет (0,2-0,7), диапазон по корреляции между разрядами кода НПБК $r = (0,0-0,5)$;

4. Гистограмма биномиального зависимого распределения и вероятности ошибки пропуска «Чужого» 256-разрядного НПБК находится по таблице при заданных значениях.

Примеры использования таблиц биномиального зависимого закона при оценке вероятностей ошибок пропуска «Чужого» многомерного высоконадежного нейросетевого преобразователя биометрия-код

Первый пример соответствует ситуации, когда пароль пользователя не скомпрометирован, а корреляция между выходными разрядами преобразователя биометрия-код составляет 0,15.

Второй пример приводится при оценке вероятности пропуска «Чужого» в ситуации, когда пароль пользователя скомпрометирован, и смещение распределения происходит влево с параметром вероятности появления состояния «1» в разряде кода НПБК $p=0,3$, корреляция между выходными разрядами преобразователя биометрия-код равна 0,35. Значение вероятности пропуска «Чужого» оценивается с помощью таблиц и функции плотности распределения, выполняется оценка вероятности угадывания «Чужим» половины длины ключа $n/2$ и вероятности угадывания «Чужим» всей длины ключа n НПБК.

Рассмотрим первый пример. Оценим вероятность пропуска «Чужого» у многомерного высоконадежного нейросетевого преобразователя биометрия-код, когда параметры биномиального зависимого закона распределения равны $p = 0,5$; $n = 256$ и $r = 0,15$

График функции плотности биномиального зависимого закона на распределения при данных параметрах приведен на рисунке А.2.

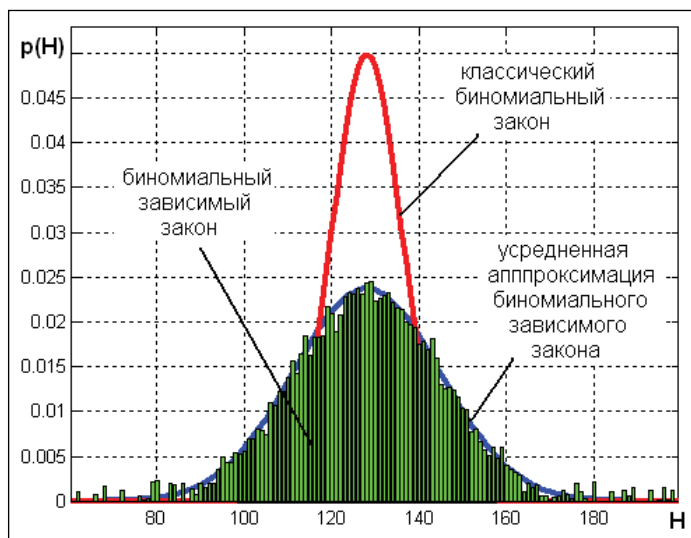


Рисунок А.2 – График функции плотности биномиального зависимого закона распределения при $p = 0,5$; $n = 256$; $r = 0,15$

В таблице А.2 приведены фрагменты таблиц биномиального зависимого закона распределения меры Хэмминга выходного кода НПБК при данных значениях параметров p , n , r на половине длины ключа $n/2$ и длины ключа n .

Таблица А.2 – Вероятности биномиального зависимого закона распределения меры Хэмминг
а выходного кода НПБК при $p = 0,5$; $n = 256$; $r = 0,15$

<i>H</i>	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256
<i>P_{код}</i>	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
<i>E_{норм}</i>	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
<i>E_{фракт}</i>	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0

Оценим вероятность пропуска «Чужого» многомерным НПБК при $p = 0,5$; $n = 256$; $r = 0,15$.

Половине длины ключа $n/2$ соответствует вероятность классического биномиального закона $P_{\text{мод}} = 10^{-14,8}$, вероятность биномиального зависимого закона определяется по таблице А.1 и равна $P_{\text{мод}} = 10^{-4,6}$.

Ошибка аппроксимации нормальным законом равна $E_{\text{норм}} = 10^{-0,9}$; ошибка фрактальной аппроксимации – $E_{\text{фракт}} = 10^{-1,4}$. Вероятность угадывания всей длины ключа n НПБК «Чужим» равна $P_{\text{мод}} = 10^{-77,1}$ по биномиальному закону. Соответствующее значение вероятности биномиального зависимого закона определяется по функции плотности распределения и равно $P = 10^{-11,3}$.

Ошибка аппроксимации нормальным законом равна $E_{\text{норм}} = 10^{-1,2}$; ошибка фрактальной аппроксимации – $E_{\text{фракт}} = 10^{-2,1}$.

Рассмотрим второй пример. Оценим вероятность пропуска «Чужого» у многомерного высоконадежного нейросетевого преобразователя биометрия-код, когда параметры биномиального зависимого закона распределения равны $p = 0,4$; $n = 256$ и $r = 0,35$

График функции плотности биномиального зависимого закона распределения при данных параметрах приведен на рисунке А.3.

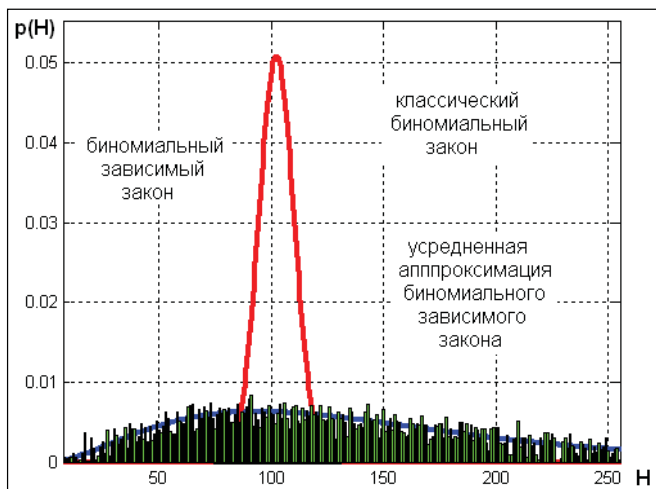


Рисунок А.3 – График функции плотности биномиального зависимого закона распределения при $p = 0,4$; $n = 256$; $r = 0,35$

Таблица А.3 – Вероятности биномиального зависимого закона распределения меры Хэмминг
а выходного кода НПКБ при $p = 0,4$; $n = 256$; $r = 0,35$

N	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128
$R_{\text{код}}$	0,6	0,6	0,6	0,6	0,6	0,6	0,6	0,6	0,6	0,6	0,6	0,6	0,6	0,6	0,6	0,6	0,6	0,6	0,6	0,6	0,6	0,6	0,6	0,6	0,6	0,6	0,6	0,6	0,6	0,6	0,6	
$E_{\text{норм}}$	-1,2	-0,3	-4,0	0,6	-1,7	-3,6	0,8	-1,5	0,3	-0,4	-2,3	-1,4	1,7	4,1	-0,9	0,9	-1,4	-2,0	-2,4	-0,5	2,3	-2,4	1,9	-1,9	2,2	1,0	0,1	-0,3	3,7	3,3	1,6	2,1
$E_{\text{фракт}}$	-2,6	-0,5	-3,4	1,2	-3,4	-3,4	1,6	-3,5	0,8	-0,9	-4,2	-3,4	3,1	10,0	-2,1	1,8	-2,6	-4,7	-5,9	-0,9	4,9	-5,3	4,3	-3,5	4,8	2,4	0,3	-0,6	8,3	6,7	2,8	4,0

N	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256
$R_{\text{код}}$	0,2	0,2	0,2	0,2	0,2	0,2	0,2	0,2	0,2	0,2	0,2	0,2	0,2	0,2	0,2	0,2	0,2	0,2	0,2	0,2	0,2	0,2	0,2	0,2	0,2	0,2	0,2	0,2	0,2	0,2	0,2	
$E_{\text{норм}}$	-1,7	2,1	3,5	-0,3	2,4	-2,6	0,6	-2,2	0,5	1,1	-1,4	0,9	-0,9	1,4	0,5	1,9	0,6	-1,1	-2,8	-0,8	1,0	-0,3	-0,4	-0,5	3,3	1,0	-1,3	3,2	5,0	0,0	3,1	2,0
$E_{\text{фракт}}$	-2,9	4,2	6,0	-1,5	5,8	-6,9	1,4	-4,8	1,4	2,2	-2,6	1,6	-2,3	2,6	1,1	3,1	1,2	-2,0	-5,9	-1,9	1,8	-0,6	-0,7	-1,1	8,0	2,1	-3,0	5,3	8,5	-0,1	6,5	3,6

В таблице А.3 приведены фрагменты таблиц биномиального зависимого закона распределения меры Хэмминга выходного кода НПБК при данных значениях параметров p, n, r на половине длины ключа $n/2$ и длины ключа n .

Оценим вероятность пропуска «Чужого» многомерным НПБК при $p = 0,4$; $n = 256$; $r = 0,35$. Половине длины ключа $n/2$ соответствует вероятность классического биномиального закона $P_{\text{мод}} = 10^{-14,2}$, вероятность биномиального зависимого закона определяется по таблице А.1 и равна $P_{\text{мод}} = 10^{-2,2}$. Ошибка аппроксимации нормальным законом равна $E_{\text{норм}} = 10^{-0,5}$; ошибка фрактальной аппроксимации – $E_{\text{фркт}} = 10^{-0,9}$. Вероятность угадывания всей длины ключа n НПБК «Чужим» равна $P_{\text{мод}} = 10^{-47,9}$ по биномиальному закону. Соответствующее значение вероятности биномиального зависимого закона определяется по функции плотности распределения и равно $P = 10^{-5,1}$. Ошибка аппроксимации нормальным законом равна $E_{\text{норм}} = 10^{-1,1}$; ошибка фрактальной аппроксимации – $E_{\text{фркт}} = 10^{-2,6}$.

Примечания:

1. Моделирование зависимого биномиального закона и вычисление вероятностей ошибок пропуска «Чужого» проведено на 1,000,000 промерах;
2. Полученные в результате численного моделирования данные обработаны двумерной сглаживающей функцией;
3. Диапазон по вероятности появления состояния «1» в разряде кода НПБК p составляет (0,2–0,7), диапазон по корреляции между разрядами кода НПБК r – (0,0–0,5);
4. Гистограмма биномиального зависимого распределения и вероятности ошибки пропуска «Чужого» 256-разрядного НПБК находится по таблице при заданных значениях.

**Ахметов Б.С., Надеев Д.Н., Фунтиков В.А.,
Иванов А.И., Малыгин А.Ю.**

ОЦЕНКА РИСКОВ ВЫСОКОНАДЕЖНОЙ БИОМЕТРИИ

Монография

Подписано в печать 08.08.2014.
Формат 60х84/16. Бумага офсетная. Гарнитура «Times».
Объем 6,75 печ. л. Тираж 500 экз. Заказ № 1359.

Отпечатано в ТОО «Издательство LEM»
050008, г. Алматы, ул. Ауэзова, 82, уг. ул. Мынбаева, тел./факс 375-51-33.