 

**МЕТОДЫ И СИСТЕМЫ ЗАЩИТЫ ИНФОРМАЦИИ, ИНФОРМАЦИОННАЯ БЕЗОПАСНОСТЬ/METHODS AND SYSTEMS OF INFORMATION PROTECTION, INFORMATION SECURITY**

**DOI: https://doi.org/10.60797/itech.2026.11.1 EDN: TKFNRG**

# МЕТОД ДЕКОДИРОВАНИЯ КОДОВ РИДА-СОЛОМОНА НА ОСНОВЕ КОРРЕКЦИИ ОШИБОК И СТИРАНИЙ ДЛЯ БИОМЕТРИЧЕСКИХ СИСТЕМ

Научная статья

**Ассанович Б.А.[1, *], Гаврилова И.Л.[2]**
[1] ORCID : 0000-0003-0418-9595;
[1, 2] Гродненский государственный университет имени Янки Купалы, Гродно, Беларусь

* Корреспондирующий автор (bas[at]grsu.by)

Предложена: 11.11.2025; Принята: 30.04.2026; Опубликована: 14.07.2026

**Аннотация**

В работе рассматривается биометрическая система (БС) аутентификации, основанная на схеме нечеткого обязательства (*fuzzy commitment*), предназначенная для защиты секретных ключей с использованием помехоустойчивых кодов. Предлагается усовершенствованный метод декодирования с исправлением ошибок и стираний для недвоичных кодов Рида-Соломона (РС), использующий вспомогательные данные двух типов, связанных с предыдущими измерениями и их вариациями. Метод базируется на многоуровневой оценке надежности символов с возможностью персонализации на основе профилей пользователей, с использованием вспомогательных данных и хеша ключа, что позволяет снизить избыточность кодирования при сохранении высокой надежности восстановления.

Предложены две оригинальные стратегии: *интеллектуальная*, включающая трехкомпонентную оценку надежности символов на основе исторических данных биометрических измерений с гибридным алгоритмом управления стираниями с расширенным поиском (*fall-back*), а также *каскадная*, последовательно применяющая фиксированный, адаптивный и приоритетный методы отбора стираний. Оценка метода проводилась с использованием синтетических данных и реальной биометрии человеческих лиц из базы данных UvA-NEMO для кодов РС длины 63 с различной корректирующей способностью: РС(63,21), РС(63,15) и РС(63,11).

Результаты экспериментов демонстрируют прирост успешности декодирования относительно базового алгоритма (без обработки стираний) от 9,5% для РС(63,11) до 43,1% для кода РС(63,21). Интеллектуальная стратегия с *fall-back* механизмом достигает 96,3% успешности для РС(63,21) при нагрузке в 4 раза меньшей, чем каскадная, а для РС(63,11) обеспечивает 99,8% успешности. При решении задачи аутентификации достигнуты значения FRR менее 1% для РС(63,11) и менее 2% для РС(63,15), что соответствует требованиям большинства практических приложений.

**Ключевые слова:** коды Рида-Соломона, декодирование с исправлением ошибок и стираний, биометрическая система, fuzzy commitment, fall-back механизм, вспомогательные данные, надежность символов, аутентификация.

# A METHOD FOR DECODING REED–SOLOMON CODES BASED ON ERROR AND ERASURE CORRECTION FOR BIOMETRIC SYSTEMS

Research article

**Assanovich B.A.[1, *], Gavrilova I.L.[2]**
[1] ORCID : 0000-0003-0418-9595;
[1, 2] Yanka Kupala State University of Grodno, Hrodna, Belarus

* Corresponding author (bas[at]grsu.by)

Suggested: 11.11.2025; Accepted: 30.04.2026; Published: 14.07.2026

**Abstract**

The work examines a biometric authentication system (BAS) based on a *fuzzy commitment* scheme, designed to protect secret keys using error-correcting codes. An improved decoding method with error and erasure correction for non-binary Reed–Solomon (RS) codes is suggested, using two types of auxiliary data related to previous measurements and their variations. The method is based on a multilevel assessment of symbol reliability, with the possibility of personalisation based on user profiles, utilising auxiliary data and a key hash; this allows for a reduction in coding redundancy while maintaining a high reliability of reconstruction.

Two original strategies are proposed: an *intelligent* strategy, involving a three-component evaluation of symbol reliability based on historical biometric measurement data, with a hybrid erasure control algorithm featuring an extended *fall-back* search; and a *cascading* strategy, which sequentially applies fixed, adaptive and priority erasure selection methods. The method was evaluated using synthetic data and real human facial biometric data from the UvA-NEMO database for 63-bit PC codes with varying correction capabilities: PC(63,21), PC(63,15) and PC(63,11).

The experimental results demonstrate an increase in decoding success rate relative to the baseline algorithm (without erasure coding) ranging from 9,5% for PC(63,11) to 43,1% for the PC(63,21) code. The intelligent strategy with a *fall-back* mechanism achieves a success rate of 96,3% for PC(63,21) with a computational load four times lower than that of the cascaded approach and provides a success rate of 99,8% for PC(63,11). When solving the authentication problem, FRR values

 

of less than 1% were achieved for RS(63,11) and less than 2% for RS(63,15), which meets the requirements of most practical applications.

**Keywords:** Reed–Solomon codes, error- and erasure-correction decoding, biometric system, fuzzy commitment, fall-back mechanism, auxiliary data, symbol reliability, authentication.

**Введение**

Традиционные биометрические системы (БС) используют как отдельные, так и каскадные схемы кодирования с применением одного или нескольких помехоустойчивых кодов для обеспечения устойчивости к естественным вариациям биометрических характеристик [1]. Однако известные подходы приводят к недостаточной достоверности или значительной избыточности [2], [3]. В данном исследовании демонстрируется, что использование расширенных возможностей кодов РС, в частности методов декодирования с исправлением ошибок и стираний, позволяет отказаться от второго каскада кодирования без потери достоверности.

Предлагаемая БС реализует модифицированную схему JW (Juels-Wattenberg) [4], которая выполняет авторизацию доступа для пользователя с использованием персонального ключа, связанного с его биометрией, имеющее название схемы нечеткого обязательства FC (Fuzzy Commitment). Это криптографический метод, который позволяет связать (зафиксировать) секретное значение с нечеткими, или «зашумленными», биометрическими данными. Применение кодов корректирующих ошибки ECC (Error Correcting Codes) позволяют восстановить исходный секретный ключ, даже если считываемые биометрические данные немного отличаются от тех, что были использованы при фиксации ключа. При регистрации биометрических данных генерируется секретное значение и создаются вспомогательные данные HD (Helper Data), состоящие из закодированного при помощи ECC ключа и «наложенного» на кодовое слово биометрического измерения. Вспомогательные данные, по сути, «подсказывают» [5] системе, как восстановить секретное значение на основе последующих, возможно, «зашумленных» считываний биометрических данных. Схема FC при регистрации сохраняет не само секретное значение, а лишь хеш ключа и «зашумлённую» разницу между ключом и биометрией. При аутентификации в БС, реализующей FC, удается восстановить созданный ключ благодаря корректирующей способности кода [6]. Таким образом, безопасная аутентификация обеспечивается проверкой хеша ключа, а конфиденциальность биометрии — отсутствием её хранения в открытом виде [7]. Стремясь преодолеть неотъемлемую проблему шумов, присущих биометрическим данным, и отойдя от концепции безопасного хранения эталонного шаблона, Додис и соавторы в 2004 году модифицировали FC и предложили конструкцию так называемого нечеткого экстрактора FE (Fuzzy Extractor), который позволяет извлекать из биометрического образа стабильный и равномерно распределенный криптографический ключ [8]. Несмотря на то, что фундамент нечётких экстракторов был построен достаточно давно, стремительная эволюция методов глубокого обучения (Deep Learning, DL) предъявляет новые требования к выбору признаков и модальностей.

Одной из последних работ, учитывающих эти изменения, стало исследование Курбатова с соавторами [9]. В этой работе представлена конструкция X-Lock и её модификация Keyed X-Lock, отличающаяся простотой реализации и применимостью на устройствах с ограниченными ресурсами. Однако эти схемы фактически реализуют ECC с повторениями и в обоих случаях имеют достаточно высокую вероятность отказа FRR (False Rejection Rate) для ключей большой длины, что ограничивает их практическую применимость. Авторы также рассмотрели конструкцию с двоичными кодами Гоппы, где при длине кода 512 бит максимальная исправляющая способность составляет 56 ошибок (10.9%), что недостаточно для работы с реальным уровнем биометрического шума. Важным вкладом работы являются экспериментальные оценки средних внутриклассового $d_+ \approx 0.20$ — 0.25 и межклассового расстояния $d_- \approx 0.48$–0.50 в современных нейросетевых моделях распознавания лиц (InsightFace, AdaFace и других), что определяет требования по учету уровня искажений биометрических признаков более 20%.

Для решения этой проблемы в современных БС широко применяются линейные коды блочного типа [10], БЧХ-коды [11] и коды РС [12], [3].

Алгебраический алгоритм декодирования кодов Рида-Соломона [12], известный как алгоритм Берлекэмпа-Мэсси (Berlekamp-Massey, BM), стал стандартом для декодирования кодов РС с жестким принятием решения. Алгоритм BM до сих пор широко применяется на практике и для его изучения и использования разработано множество программных комплексов, как например [13], что актуально для подготовки специалистов в области информационной безопасности.

В последние годы отмечается активное внедрение методов DL в задачу декодирования помехоустойчивых кодов [14], [15], [16]. Так, в исследовании X. An и др. [15] для оценки количества ошибок используется глубокая нейронная сеть (Deep Neural Network, DNN), которая динамически выбирает или оптимизирует алгоритм декодирования. Однако ограничением этого подхода является валидация на кодах малой длины, что ставит под вопрос его обобщающую способность для более длинных кодов. В работе W. Zhang и др. [16] результатом является трансформация классического стохастического итеративного алгоритма в DNN, что позволяет ввести адаптивные веса и избегать локальных оптимумов, повышая корректирующую способность. Критическим недостатком данного исследования также является его ориентация исключительно на короткий код РС(15,11) без демонстрации масштабируемости, что оставляет открытым вопрос о его вычислительной эффективности и устойчивости при увеличении длины кода.

Однако прямое применение передовых DL-методов с DNN к задачам биометрии сталкивается с вычислительными ограничениями при росте размерности полей Галуа и практическими ограничениями при развёртывании на мобильных и встроенных устройствах. В БС актуально применение «легковесных» алгебраических методов, которые эффективно используют информацию о надежности канала для маркировки стираний и коррекции ошибок, но при этом обладают предсказуемой вычислительной сложностью. Предлагаемый в данной работе подход фокусируется не на фундаментальном переосмыслении процесса декодирования с помощью DL, а на целевой оптимизации классической цепочки «оценка надежностей, маркировка стираний, алгебраическое декодирование» для специфических условий биометрических каналов.

 

В работе рассматривается биометрическая система, в которой секретный ключ K кодируется недвоичным кодом РС длины n символов, что позволяет улучшить эффективность декодирования без применения каскадных кодов [5]. Для этого вместо традиционного жёсткого декодирования (Hard Decision Decoding, HDD) предлагается метод мягкого декодирования (Soft Decision Decoding, SDD), базирующийся на оценке надёжности символов кода над целочисленным полем Галуа.

Новизна исследования определяется двумя предложенными стратегиями обработки биометрических данных: интеллектуальной и каскадной.

Интеллектуальная стратегия — с трехкомпонентной оценкой надежности на основе исторических профилей пользователей и гибридным управлением стираниями с расширенным поиском. Каскадная стратегия — с последовательным применением фиксированного, адаптивного и приоритетного методов отбора стираний. Данные стратегии обеспечивают высокую успешность декодирования при оптимальном балансе вычислительной нагрузки для различных типов кодов.

**Биометрическая система с двумя образцами вспомогательных данных**

Для создания биометрического ключа и формирования устойчивого биометрического шаблона в БС создается супервектор $Y = y^1 \cup y^2 \cup \quad y^M$ из отдельных векторов $y^i$, где M — количество обрабатываемых биометрических фреймов данных. Это может быть несколько отпечатков пальца либо видеокадров лица человека и т.п. Предлагаемая система реализует схему FC (ниже представлена ее формализация), используя один или два образца вспомогательных данных HD1 и HD2, которые выполняют принципиально различные функции. HD1 содержит данные, связанные с вариацией измерений, которые образуют персонализированную мета-информацию о стабильности биометрических признаков пользователя, применяемую адаптивным алгоритмом для оптимального выбора стираний. Информация HD2 представляет собой публичные криптографические данные, маскирующие секретный ключ. Инновация предлагаемого подхода заключается не в изменении схемы FC, а в использовании HD1 для аналитической оптимизации процесса коррекции ошибок, применяемого к HD2.

Важными характеристиками производительности БС являются вероятность ложного доступа FAR (False Acceptance Rate) и вероятность ложного отказа FRR (False Rejection Rate). Величина FRR зависит от вероятности ошибки на символ $p_{err}$ для биномиального распределения $X_i \sim Bin(n_i, p_{err})$ и фактически ограничена сверху вероятностью того, что в i-том блоке данных число ошибок превысит порог коррекции $t_i$ $FRR_i \sim P(X_i > t_i)$ [3].

При равномерном квантовании на L уровней вероятность «угадать» отдельный символ для злоумышленника равна $(p_g=1)/L$. Тогда, в рамках биномиальной модели для $Y \sim Bin(n, p_g)$, где Y — количество угаданных злоумышленником символов, вероятность FAR вычисляется как $FAR=P(Y \geq k)$, где k – порог восстановления ключа. Оптимальным для решаемой задачи признано квантование на L=64 уровня. Данный шаг квантования более чем на порядок меньше среднего абсолютного отклонения биометрических данных и гарантирует надёжное разделение различных реализаций одного признака.

Классическая модель биномиального канала предполагает, что вероятность ошибки символа p_err фиксирована и одинакова для всех символов, а ошибки статистически независимы. Однако, как показано в работе [18] и подтверждено экспериментальными данными, реальное распределение вещественных биометрических признаков существенно отличается от равномерного. На основе анализа тестовой выборки базы данных UvA-NEMO с применением критерия согласия Колмогорова-Смирнова установлено, что распределение признаков наилучшим образом аппроксимируется бета-распределением с параметрами α=7,223 и β=3,754. Выбор обоснован минимальным значением статистики (D=0.024, с p-value>0.05) по сравнению с альтернативными распределениями (см. зависимость распределений от надежности на рис. 1).

Переход от биномиальной модели к бета-распределению принципиально меняет оценку вероятности ошибки символа p_err при равномерном квантовании на L=64 уровня. Оценка ошибки квантования, отображающая вариативность биометрических данных, описывается на основе функции плотности вероятности f(x), которая аппроксимируется бета-распределением и может быть получена с учетом независимости двух измерений при попадании в интервал квантования

$$p_{err} = 1 - \sum_{i=1}^{L} \left( \int_{(i-1)\delta}^{i\delta} f(x)dx \right)^2 .$$

В биномиальной модели с равномерным распределением признаков $p_{err}$ составляла около 0.25, тогда как для бета-распределения численное интегрирование даёт $p_{err}=0.174$.

Современные тенденции к криптографической стойкости биометрических систем определяют использование ключей длиной не менее 128 бит [5]. В данной работе были выбраны три основных кода для эффективной реализации БС в предположении, что используется квантование с уровнем L=6 бит:

– код РС(63,21), для которого достаточно M=1 кодового слова (длина ключа: 21×6=126 бит, близко к 128 бит).

– код РС(63,15), в случае которого необходимо 2 кодовых слова для длины ключа 2×15×6=180 бит (более 128 бит).

– код РС(63,11), который требует M=2 кодовых слова (длина ключа: 2×11×6=132 бита).

Выбор между ними определяется компромиссом между вычислительной сложностью, особенностями применения в режиме исправления ошибок и стираний, в параметрах FAR и FRR. Для кода РС(63,21) без использования стираний переход от биномиальной модели ($p_{err}=0,22$) к бета-распределению ($p_{err}=0,174$) снижает FRR с 1,02% до 0.19%. А для кода РС(63,15) и РС(63,11) при $p_{err}=0,174$ величина FRR составляет $3,4\times10^{-6}$ и $1,2\times10^{-7}$ соответственно.

При фиксированной общей вероятности искажения символа $p_{err}$ переход от ошибок к стираниям позволяет снизить «стоимость» искажений с $2p_{err}$ до $2p_e+p_s$, что эквивалентно увеличению запаса исправляющей способности на $np_s$, где n — длина кода, а $p_s$ — вероятность стирания символа и $p_e$ — вероятность ошибки, которые в сумме дают $p_{err}$.

 

Доля успешного декодирования кода с учетом стираний, например РС(63,21), вычисляется сумма трехчленного распределения по всем допустимым комбинациям числа ошибок $N_e$=e и числа стираний $N_s$=s:

$$P_{\text{success}} = \sum_{e=0}^{n} \sum_{s=0}^{n-e} 1[2e + s \le 42] \cdot \frac{n!}{e!s!(n-e-s)!} \cdot p_e^e p_s^s \left(1 - p_e - p_s\right)^{n-e-s},$$

где 1[2e+s≤42] — индикаторная функция, которая «включает» те слагаемые, которые допустимы с точки зрения корректирующей способности кода.

При введении декодирования со стираниями с параметрами $p_e$=0,10, $p_s$=0,074, сумма которых дает $p_{err}$=0,174, FRR для кода РС(63,21) сокращается до $2{,}4\times10^{-7}$, а для кодов РС(63,15) и РС(63,11) уменьшается до значений $2{,}3\times10^{-10}$ и $10^{-12}$ соответственно. Параметр FAR также претерпевает изменения для бета-распределения с $p_{err}$=0,174 и L=64 FAR и составляет менее $10^{-20}$ для всех кодов.

При бета-распределении код РС(63,21) остаётся вполне приемлемым для многих приложений с FRR=0,19%. Однако для критических применений, требующих FRR менее $10^{-6}$, безальтернативным становится использование двух кодовых слов РС(63,11) или РС(63,15). Декодирование со стираниями в сочетании с бета-распределением позволяет достичь $FRR=2{,}4\times10^{-7}$ даже для кода РС(63,21). Это демонстрирует, что правильный учёт статистических характеристик источника может компенсировать недостаточную исправляющую способность кода.

**Декодирование кодов Рида-Соломона**

Известно, что результатом HDD-декодирования кода с n символами является либо декодированное сообщение, либо индикация ошибки декодирования. Однако код РС позволяет исправлять не только ошибки, но и стирания. Если $N_s$ символов кода помечены как стертые, а оставшиеся n-$N_s$ символов содержат $N_e$ ошибок, алгоритм BM находит правильное кодовое слово при выполнении условия, связанного с кодовым расстоянием d [14]:

$$N_s + 2N_e \le 2t < d$$

Если $N_s$=0, декодер используется как декодер только для коррекции ошибок, а если $0<N_s\le d-1$, декодер называется декодером ошибок-и-стираний (Error-and-Erasure Decoding, EED). Однако для применения принципа EED необходимо определить, какие символы являются ошибочными или стертыми.

В распоряжении декодера имеются исторические данные HD1, формируемые на основе абсолютных отклонений предшествующих биометрических измерений от среднего значения (или от центра квантования):

$$D_i = \left| y_i^j - \mu_i \right|,$$

где $y_i^j$ — i-ый элемент в j-том биометрическом образце, $\mu_i$ — его базовое значение.

Таким образом, отклонения $D_i$ могут служить дополнительной информацией для реализации EED на этапе верификации в биометрической системе. Для этого в обрабатываемом кодовом слове необходимо пометить $N_s$ элементов из общего числа символов n как стирания и сформировать вектор стирания на основе интегральной метрики надежности, получаемой путем адаптивной агрегации трех компонент, описанной ниже.

В основу выбора стратегии стираний положим аналитический подход, основанный на работе Вебера и др. [19], в которой параметры кода d и количество ошибок $N_e$ заменяются относительными долями $\delta=d/n$, $\varepsilon=N_e/n$. Дискретный вектор надёжности символов аппроксимируется непрерывной функцией распределения $h(\phi)$, $\phi\in(0,1]$. Затем определяется оптимальная доля стираний $\tau^*$, максимизирующая корректирующую способность кода на основе интегральной «накопленной ненадежности» $\Delta_h(\tau)$. После этого значение $\tau^*$ преобразуется в пороговый уровень надежности $h_{threshold}$ и все символы с надежностью ниже этого порога объявляются стираниями.

На основе сформированного вектора стирания выполняется процедура EED. Если декодирование оказывается неудачным, активируется механизм расширенного поиска (fall-back), в рамках которого количество стираний последовательно увеличивается. Процесс продолжается до тех пор, пока либо декодирование не завершится успешно с восстановлением ключа K, либо не будут исчерпаны все итерации.

**Математическая модель БС и процедура верификации ключа**

Пользовательский ключ K длины k кодируется кодом РС с параметрами (n, k) над полем $GF(2^m)$, в результате чего формируется кодовое слово:

$$C = RS\_\text{encode}(K), C \in GF(2^m)^n.$$

На этапе регистрации биометрический образец пользователя подвергается квантованию [17], в результате чего формируется вектор $Z\in\{0, 1, \ldots, 2^m\text{-}1\}^n$. Побитовое сложение по модулю 2 кодового слова с квантованными данными дает публичные вспомогательные данные HD2:

$$W = C \oplus Z.$$

Параллельно формируется вектор исторических разностей D, статистические характеристики которого (квантили $Q_1$, $Q_3$, коэффициент вариации $CV_i$) составляют вспомогательные данные HD1.

На этапе верификации пользователь предъявляет новое биометрическое измерение, которое квантуется в вектор Z'. Восстановление искаженного кодового слова выполняется как:

$$C' = W \oplus Z' = C \oplus \left(Z \oplus Z'\right).$$

Разность Z⊕Z^' представляет собой зашумленный вектор, указывающий позиции, в которых произошли изменения квантованных значений.

Для организации эффективного декодирования вычисляется вектор надежности R=(R(1),R(2),…,R(n)) где $R(i)\in[0,1]$ характеризует степень доверия к корректности i-того символа. Методика вычисления R(i) подробно рассматривается в следующем разделе.

На основе вектора R формируется множество стираний:

$$\varepsilon = \{i \mid R(1) \le h\},$$

где h — пороговое значение, определяемое выбранной стратегией декодирования. Затем выполняется EED — декодирование искаженного кодового слова C' с учетом позиций стираний:

 

$$K' = EED\_decode\left(C', \varepsilon\right).$$

Успешным считается декодирование, при котором восстановленный ключ совпадает с исходным ключом.

**Нахождение компонент надежности**

Для реализации эффективного EED критически важной является точная оценка надежности символов. Предлагается трехкомпонентный метод оценки надежности на основе исторических разностей.

Компонента 1: Первая компонента отражает тот факт, что символы, демонстрирующие стабильность в предыдущих измерениях, должны получать более высокую оценку надёжности. Она вычисляется как:

$$R_1(i) = \frac{1}{1+D_i},$$

где $D_i$ — историческая разность для i-того символа, определяемая как среднее отклонение его значений от среднего (или от центра квантования) на обучающей выборке, что на практике приводит фактически к тем же результатам.

Компонента 2: Вторая компонента учитывает относительную стабильность символа с учетом его среднеквадратичного отклонения $\sigma_i$, нормированного на значение центра квантования $\mu_i$.

Эта компонента, обратная к коэффициенту вариации $CV_i=\sigma_i/\mu_i$, определяется как

$$R_2(i) = \frac{1}{1+CV_i}.$$

Компонента 3: Третья компонента представляет собой дискретную оценку, основанную на положении исторической разности $D_i$ относительно квантилей её эмпирического распределения:

$$R_3(i) = \begin{cases} 0.9, & D_i \le Q_1 \\ 0.6, & Q_1 < D_i \le Q_3 \\ 0.3, & D_i > Q_3 \end{cases}$$

где $Q_1$ и $Q_3$ — первый и третий квантили распределения исторических разностей. Такой подход позволяет разделить символы на три категории: стабильные (высокая надёжность), промежуточные (средняя надёжность) и нестабильные (низкая надёжность).

В отличие от простого усреднения, предлагается изучить динамическое формирование весовых коэффициентов для каждой компоненты, учитывающее накопленный опыт взаимодействия с пользователем. Процесс адаптивной агрегации компонент надежности может быть реализован в четыре этапа:

Таблица 1 - Этапы адаптивной агрегации компонент надежности

DOI: https://doi.org/10.60797/itech.2026.11.1.1

| Этап | Измерения | Механизмы | Результат |
|---|---|---|---|
| 1 | a=1 | Разные веса $w_1=w_2=w_3=1/3$ | Нейтральный статус |
| 2 | a=2 | Веса по согласованности $w_j=c_j/(\sum c_k)$ | персонализация |
| 3 | a≥3 | Экспоненциальное сглаживание $w_j=0{,}9w_j+0{,}1e_j$ | Плавная оптимизация |
| 4 | $a \to \infty$ | Сходимость к оптимальным значениям | Стабилизация |

*Примечание: здесь 0.9 и 0.1 – коэффициенты экспоненциального сглаживания веса*

При первом измерении используется равное взвешивание. После второго измерения вычисляется согласованность каждого источника, отражающая стабильность его оценок между двумя последовательными измерениями, где верхние индексы 1 и 2 обозначают номер измерения (метку времени):

$$c_j = 1 - \frac{1}{n}\sum_{i=1}^{n}\left|R_j^1(i) - R_j^2(i)\right|,$$

отражающая стабильность оценок между двумя последовательными измерениями. Веса для второго измерения формируются пропорционально согласованности:

$$w_j = c_j / \left(c_1 + c_2 + c_3\right).$$

Начиная с третьего измерения, включается механизм экспоненциального сглаживания с коэффициентом памяти p. После каждого успешного декодирования вычисляется эффективность источников:

$$e_j = \frac{1}{n}\sum_{i=1}^{n}\left[v_i R_j(i) + \left(1 - v_i\right)\left(1 - R_j(i)\right)\right]$$

где $v_i=1$ если символ i декодирован правильно. Обновление весов производится по формуле $w_j^{(new)}=0.9w_j^{(old)}+0.1e_j$ с последующей нормировкой. Коэффициент p=0,9 обеспечивает время полужизни информации около 7 измерений, что гарантирует устойчивость к случайным выбросам.

 

Итоговая надежность символа вычисляется как взвешенная сумма:

$$R(i) = w_1 R_1(i) + w_2 R_2(i) + w_3 R_3(i) \ .$$

**Оптимизация стратегии стираний**

Полученный вектор надежностей R используется для аналитической оптимизации доли стираний. Сначала строится эмпирическая функция распределения надежностей h(φ) путем сортировки R(i) по возрастанию. Полученное распределение аппроксимируется бета-распределением Beta(α,β). Для аппроксимирующего распределения $I_\varphi(\alpha,\beta)$ вычисляется оптимальная доля стираний $\tau^*$, максимизирующая корректирующую способность [19]:

$$\tau^* = \arg\max_{\tau\in[0,\delta]} \Delta_h(\tau) \ , \ \ \Delta_h(\tau) = \frac{1}{\delta}\left[1 - \frac{\alpha}{\alpha+\beta} + 2\int_{\tau}^{\tau+\epsilon(\tau)} I_\varphi(\alpha,\beta)d\varphi\right] \ ,$$

где τ — доля стираний, а λ — параметр декодирования кода РС, имеющий λ=2 (для ВМ случая) и определяющий нормированный радиус исправления ошибок $\epsilon(\tau)=(\delta-\tau)/\lambda$.

Далее методом прямого поиска по отсортированному массиву надежностей определяется пороговое значение $h_{threshold}=R_{(\lceil\tau^*n\rceil)}$, и все символы с $R(i)\le h_{threshold}$ маркируются как стирания.

**Две стратегии декодирования РС кодов с анализом надежностей символов**

Искажённое кодовое слово, возникающее из-за различий биометрических образцов после вычитания HD2, в большинстве случаев успешно декодируется, позволяя восстановить ключ K, вычислить его хеш и сравнить с эталонным. Достоверность (успешность декодирования) определяется применяемым кодом и алгоритмом декодирования.

Для кода РС с кодовым расстоянием d условие успешного декодирования определяет нагрузку на декодер Load, при которой возможно выполнение декодирование EDD: $Load\le 2N_e+N_s$ и которая может иметь значение $Max_{Load}=d-1$.

Предложим 2 стратегии маркировки стираний, представленные в виде алгоритма 1 и алгоритма 2.

***1. Алгоритм 1.* Адаптивное декодирование с интеллектуальным управлением**

Вход: вектор надежностей R, полученное кодовое слово $C_{est}$, профиль пользователя profile

Выход: восстановленный ключ $K_{decoded}$

1. Вычисление скорректированных метрик с учетом границ квантования.

Для каждого символа i:

Определить ближайший центр $\mu_{target}$ и соседний центр $\mu_{neighbor}$

Вычислить базовое расстояние $base_{dist}=|y_i-\mu_{target}|$

Вычислить расстояние до соседнего центра $boundary_{dist}=|y_i-\mu_{neighbor}|$

Если нет превышения порога $boundary_{dist}<threshold$,

то вводится штраф с коэффициентом $k_p$:

$$\text{penalty} = k_p \cdot \left(\text{threshold} - \text{boundary}_{\text{dist}}\right)$$

Иначе penalty=0

Скорректированное расстояние: $D_i=base_{dist}+penalty$

Компоненты надежности:

$$R_1(i) = 1/\left(1 + D_i\right)$$
$$R_2(i) = 1/\left(1 + CV_i\right)$$

$R_3(i)$ — квантильная классификация по $D_i$ и $(Q_1, Q_3)$ из HD1

2. Адаптивная агрегация компонент

В зависимости от номера измерения a:

Если a=1: $R_{final}(i)=(R_1(i)+R_2(i)+R_3(i))/3$

Если a=2:

Вычислить согласованности $c_j=1-1/n\ \sum|R_j^1(i)-R_j^2(i)|$

Вычислить веса $w_j=c_j/(c_1+c_2+c_3)$

$$R_{\text{final}}\ (i) = w_1 R_1(i) + w_2 R_2(i) + w_3 R_3(i)$$

Если a≥3:

Загрузить текущие веса $w_j$ из профиля

$$R_{\text{final}}\ (i) = w_1 R_1(i) + w_2 R_2(i) + w_3 R_3(i)$$

3. Оптимизация стратегии стираний

Отсортировать $R_{sorted}=sort(R_{final})$

Аппроксимировать распределение R_sorted бета-распределением Beta(α,β) методом моментов

Найти оптимальную долю стираний $\tau^*=\arg\max_r \Delta_h(\tau)$

Определить количество стираний $k=\lceil\tau^*\cdot n\rceil$

Вычислить порог $h_{threshold}=R_{sorted}\ (k)$

4. Основная попытка декодирования

Сформировать множество стираний $erasures=\{i|\ R_{final}(i)\le h_{threshold}\}$

Если количество $N_s$: $|erasures|\le 2t$:

Выполнить декодирование $K_{decoded}=EED_{decode}(C_{est}, erasures)$

При успехе: обновить профиль и вернуть $K_{decoded}$

5. Расширенный поиск (резервный механизм)

6. Получить индексы в порядке возрастания надежности

$$sorted\_indices = argsort(R_{final})$$

Для $k_{test}=k+1$ до 2t:

$$erasures = sorted\_indices[1 : k_{test}]$$

 

$$K_{decoded} = EED\_decode(C_{est},\ erasures)$$

При успехе: обновить профиль и вернуть $K_{decoded}$
Обновление профиля при успехе
Вычислить эффективность $e_j$ каждого источника
Обновить веса: $w_j=0.9\cdot w_j+0.1\cdot e_j$
Нормировать веса
Обновить квантили ($Q_1$, $Q_3$) с учетом новых $D_i$
8. Возврат ошибки декодирования в случае неудачи всех попыток

**2. Алгоритм 2: Каскадное декодирование с равными весами**

Вход: вектор надежностей R, полученное кодовое слово $C_{est}$, профиль пользователя profile
Выход: восстановленный ключ $K_{decoded}$
1. Вычисление базовых компонент надежности
Для каждого символа i:
$R_{base}(i)=1/(1+base_{dist}(i))$ — расстояние до центра интервала
$R_2(i)=1/(1+CV\_i))$ — коэффициент вариации
$R_3(i)$ — квантильная классификация по $base_{dist}(i)$ и ($Q_1$, $Q_3$) из HD1
Итоговая надежность с равными весами:

$$R_{\text{final}}(i) = \left(R_{\text{base}}(i) + R_2(i) + R_3(i)\right)/3$$

2. Каскадный перебор стратегий
Уровень 1: Фиксированная стратегия (Fixed)

$$\text{erasures} = \text{find}\left(R_{\text{final}} < 0.35\right)$$
$$\text{errors} + \text{find}\left(0.35 \le R_{\text{final}} < 0.65\right)$$

Если для $N_e$, $N_s$: 2|errors|+|erasures|≤Max_Load:

$$K_{\text{decoded}} = EED\_\text{decode}\left(C_{\text{est}}, \text{erasures}\right)$$

При успехе: обновить профиль и вернуть $K_{decoded}$
Уровень 2: Адаптивная стратегия с усечением нагрузки (Adaptive limited)
Вычислить квартили $Q_1(R)$ = quantile($R_{final}$, 0.25); $Q_3(R)$ = quantile($R_{final}$, 0.75)
Кандидаты на ошибки: $all_{errors}$ = find($Q_1(R)<R_{final}\le Q_3(R)$)
Кандидаты на стирания: $all_{erasures}$ = find($R_{final}\le Q_1(R)$)
Сортировать кандидаты по возрастанию надежности
Усечь нагрузку: удалять самые ненадежные символы, пока 2|errors|+|erasures|≤Max_Load

$$K_{\text{decoded}} = EED\_\text{decode}\left(C_{\text{est}}, \text{erasures}\right)$$

При успехе: обновить профиль и вернуть $K_{decoded}$
Уровень 3: Стратегия приоритетного стирания (Erasure priority)
Определить максимально возможное число стираний

$$k_m = \lfloor \text{Max\_Load}/2 \rfloor$$

Отсортировать индексы по возрастанию надежности:
$sorted_{indices}$=argsort($R_{final}$)
Выбрать k' худших символов: erasures= $sorted_{indices}$[1:k']
$K_{decoded}$=EED_decode($C_{est}$, erasures)
При успехе: обновить профиль и вернуть $K_{decoded}$
3. Обновление профиля
Обновить квантили $Q_1$, $Q_3$ с учетом новых измерений
4. Возврат ошибки декодирования в случае неудачи всех трех попыток.

Представленные алгоритмы реализуют принципиально разные подходы к декодированию. Стратегия с интеллектуальным управлением делает акцент на максимально точную оценку надёжности каждого символа, используя штрафы за близость к границам квантования и динамическую адаптацию весов трёх компонент на основе истории пользователя. Дополнительно применяется аналитическая оптимизация доли стираний $\tau^*$ через бета-аппроксимацию и резервный расширенный поиск fall-back.

Каскадная стратегия, напротив, использует упрощённые метрики без штрафов и с равными весами, но компенсирует это интеллектуальным перебором трёх базовых методов декодирования в оптимальном порядке. Сначала применяется быстрая фиксированная стратегия с эмпирически подобранными порогами 0,35/0,65. В случае неудачи задействуется адаптивная стратегия с усечением нагрузки, позволяющая уложиться в ограничения кода. Последним резервом служит стратегия приоритетного стирания, гарантирующая допустимую нагрузку ценой потери части информации.

**Экспериментальные результаты**

Перед непосредственным сравнением стратегий декодирования на основе статистического анализа был реализован этап предварительного обучения параметров штрафной функции penalty, учитывающей близость символа к границе интервала квантования. Из 15 возможных попарных комбинаций 6 биометрических измерений для каждого пользователя выделялось 7 обучающих пар и 8 тестовых. Для каждого символа фиксировались три величины: расстояние до соседнего центра $boundary_{dist}$, расстояние до своего центра $base_{dist}$ и бинарный факт ошибки (перескока в другой интервал) $error_{occurred}$.

 

Процесс обучения включал последовательную реализацию следующих шагов: дискретизацию пространства значений $boundary_{dist}$ на интервалы, оценку условной вероятности ошибки в каждом интервале, поиск порогового значения, при котором риск ошибки превышает 20%, и расчёт коэффициента штрафа $k_p$, обеспечивающего желаемое снижение надёжности для символов в опасной зоне, определяемой пороговым значением threshold. Полученные таким образом параметры (threshold и $k_p$) использовались в адаптивной стратегии с интеллектуальным управлением для коррекции метрики $R_1$.

После завершения этапа обучения был проведён эксперимент по сравнению трёх стратегий декодирования: фиксированной (Fixed), адаптивной с интеллектуальным управлением (Adaptive+штраф+веса) и комбинированной каскадной. Оценка производилась для трёх кодов РС с различной избыточностью – РС(63,21), РС(63,15) и РС(63,11) — при пяти уровнях шума, соответствующих доле ошибочных символов от 16% до 46%. Данные, полученные при 95% уровне доверия, сведены в таблицу 2 и показаны на рисунке 1.

Таблица 2 - Сравнительный анализ трёх стратегий декодирования

DOI: https://doi.org/10.60797/itech.2026.11.1.2

| Код РС | Шум | Фиксированная стратегия, % | Адаптивная стратегия с интеллектуальным управлением, % | Комбинированная каскадная стратегия, % |
|---|---|---|---|---|
| РС(63,21) | 0,50 | 100,0 [99,3-100,0] | 99,5 [98,7-100,0] | 100,0 [99,3-100,0] |
| | 0,75 | 86,0 [82,6-89,4] | 80,5 [76,7-84,3] | 89,0 [85,9-92,1] |
| | 1,00 | 50,0 [45,1-54,9] | 46,5 [41,6-51,4] | 52,0 [47,1-56,9] |
| | 1,25 | 25,5 [21,4-29,6] | 20,5 [16,8-24,2] | 30,0 [25,7-34,3] |
| | 1,50 | 12,0 [9,1-14,9] | 8,5 [6,1-10,9] | 12,5 [9,6-15,4] |
| РС(63,15) | 0,50 | 99,5 [98,7-100,0] | 99,0 [98,0-100,0] | 100,0 [99,3-100,0] |
| | 0,75 | 89,0 [85,9-92,1] | 87,5 [84,2-90,8] | 90,0 [86,9-93,1] |
| | 1,00 | 70,0 [65,5-74,5] | 69,0 [64,4-73,6] | 75,0 [70,8-79,2] |
| | 1,25 | 48,5 [43,6-53,4] | 48,5 [43,6-53,4] | 52,5 [47,6-57,4] |
| | 1,50 | 20,0 [16,3-23,7] | 17,5 [14,0-21,0] | 22,5 [18,7-26,3] |
| РС(63,11) | 0,50 | 100,0 [99,3-100,0] | 99,5 [98,7-100,0] | 100,0 [99,3-100,0] |
| | 0,75 | 96,5 [94,5-98,5] | 93,5 [90,9-96,1] | 98,5 [97,1-99,9] |
| | 1,00 | 81,0 [77,2-84,8] | 80,5 [76,7-84,3] | 85,0 [81,5-88,5] |
| | 1,25 | 54,0 [49,1-58,9] | 60,5 [55,7-65,3] | 57,5 [52,6-62,4] |
| | 1,50 | 36,0 [31,5-40,5] | 35,5 [31,0-40,0] | 39,5 [34,9-44,1] |

Усреднение результатов по всем трём кодам и пяти уровням шума (см. рисунок 1) показывает, что комбинированная каскадная стратегия обеспечивает наилучшие показатели с общей успешностью 65.2%, превосходя фиксированную стратегию (62.5%) на 2.7 процентных пункта и адаптивную с интеллектуальным управлением (60.4%) на 4.8 процентных пункта. При этом фиксированная стратегия, несмотря на свою простоту, демонстрирует более высокую эффективность, чем сложная адаптивная схема со штрафами и обучаемыми весами. Фиксированная стратегия применяет неизменные пороги для маркировки стираний и ошибок, что обеспечивает простоту и предсказуемость без этапа обучения.

Для кода с наименьшей избыточностью РС(63,21) комбинированная стратегия достигает средней успешности 56.7%, что на 1.9% выше фиксированной и на 5.6% выше адаптивной. Наибольший выигрыш наблюдается при средних уровнях шума: при σ=1.0 успешность комбинированной каскадной стратегии составляет 52.0% против 46.5% у адаптивной, при σ=1.0–30.0% против 20.5%. Это свидетельствует о способности каскадного подхода эффективно использовать резервные механизмы именно в тех условиях, где простые стратегии перестают справляться.

Для кода средней избыточности РС(63,15) средняя успешность комбинированной стратегии достигает 68.0%, что на 2.6% выше фиксированной и на 3.7% выше адаптивной. Максимальное преимущество зафиксировано при σ=1.0: 75.0% против 69.0% у адаптивной (+6.0%). При высоком уровне шума (σ=1.5) комбинированная стратегия сохраняет преимущество в 2.5% над фиксированной и 5.0% над адаптивной.

Наиболее мощный код РС(63,11) ожидаемо демонстрирует наилучшие абсолютные показатели: средняя успешность комбинированной стратегии составляет 76.1%, что на 2.6% выше фиксированной и на 2.2% выше адаптивной. При σ=1.25 наблюдается единственное исключение, где адаптивная стратегия незначительно превосходит комбинированную (60.5% против 57.5%), однако при максимальном шуме σ=1.5 комбинированная стратегия вновь выходит в лидеры с результатом 39.5% против 35.5% у адаптивной. Суммируя результаты по всем трём кодам, комбинированная каскадная стратегия обеспечивает устойчивое преимущество в 13 из 15 исследованных точек (87% случаев), а адаптивная стратегия с интеллектуальным управлением, несмотря на свою вычислительную сложность в среднем проигрывает всем стратегиям.

 

Для сравнения предлагаемых стратегий EDD декодирования кодов РС было также изучена обработка реальных измерений из биометрии человеческих лиц, взятых из базы данных UvA-NEMO [5] которые после квантования кодировались различными кодами РС длиной 63 символа. Выборка содержала 240 различных биометрических образцов, составленных из 6-ти измерений для 40 пользователей, что соответствовало объему сгенерированных синтетических данных.

Ключевым результатом стало то, что использование реальных данных показало существенно более высокую эффективность предложенных стратегий, чем предсказывали синтетические модели. Для кода РС(63,21) интеллектуальная стратегия достигла 84.2% успешности на реальных данных, что соответствует синтетическому шуму σ≈0.65, а каскадная стратегия (93.3%) – σ≈0.55. Еще более впечатляющие результаты получены для РС(63,11), где интеллектуальная стратегия обеспечила 99.8% успешности, что эквивалентно работе в условиях минимального шума.

На реальной биометрии отчетливо проявилось преимущество интеллектуальной стратегии, использующей индивидуальные профили пользователей (коэффициенты вариации $CV_i$ и квантили $Q_1$, $Q_3$). В то время как на синтетических данных преимущество интеллектуальной стратегии перед базовой было умеренным (46.5% против 50.0% для РС(63,21) при σ=1.0), а на реальных данных разрыв составил 43 процентных пункта (84.2% против 41.2%).

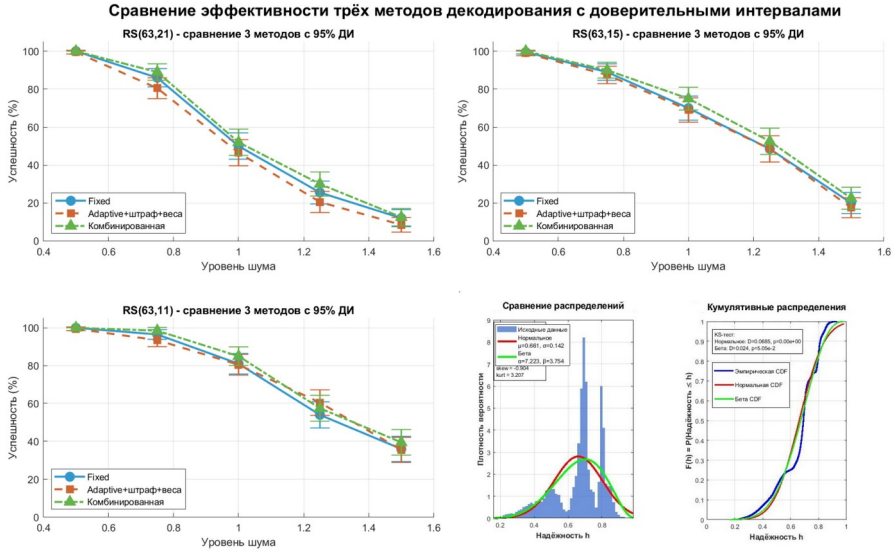


Рисунок 1 - Сравнение эффективности 3-х методов декодирования
DOI: https://doi.org/10.60797/itech.2026.11.1.3

Детальное исследование различных опций интеллектуальной стратегии позволило выявить следующие закономерности:

– адаптивная агрегация весов в зависимости от номера измерения не дает значимого преимущества (разница менее 1%), что позволяет использовать упрощенную версию с равными весами;

– использование центров квантования и средних значений дает сопоставимые результаты, оба подхода работоспособны;

– штраф за близость к границам квантования не влияет на итоговую успешность при качественном определении стираний.

Механизм резервирования *fall-back* с увеличением числа стираний при неудаче оказался критически важным, обеспечивая прирост успешности до 12% для слабых кодов.

**Практические рекомендации и оценка сложности предложенных алгоритмов**

На основе проведенного анализа может быть рекомендована следующая оптимальная стратегия для практического применения в биометрических системах:

– для высоконадежных систем целесообразно использовать код РС(63,11), обеспечивающий 99.8% успешности при минимальной нагрузке равной 13.

– для систем с ограниченными вычислительными ресурсами оптимален код РС(63,15) с 98.3% успешности и той же нагрузкой.

– интеллектуальная стратегия должна использовать равные веса компонент и обязательный fall-back механизм.

Важным аспектом практического применения разработанных алгоритмов является их вычислительная сложность, определяющая возможность реализации в системах реального времени с ограниченными ресурсами.

Базовая стратегия без использования стираний служит естественным ориентиром: она требует лишь однократного вызова декодера кода РС, что делает её минимальной по сложности. Фиксированная стратегия с порогами 0.35/0.65 добавляет к этому два линейных прохода по массиву надёжностей для идентификации стираний и ошибок (сложность O(n)), что увеличивает время выполнения примерно на 10-15% при незначительном росте потребления памяти.

 

Интеллектуальная адаптивная стратегия с персонализированными параметрами в своей теоретически полной реализации требует хранения профилей пользователей и нормализации массива надёжностей для аппроксимации бета-распределением. Однако, как показали эксперименты, вычисление оптимального τ является вычислительно затратным и не даёт значимого выигрыша по сравнению с использованием фиксированного значения τ. В упрощённой реализации, применённой в данном исследовании, интеллектуальная стратегия использует равные веса компонент и фиксированную долю стираний, что снижает её сложность.

Комбинированная каскадная стратегия, последовательно применяющая методы с фиксированными порогами, адаптивный с ограничением нагрузки и приоритетный по стираниям, демонстрирует оптимальный баланс между эффективностью и вычислительной сложностью. В лучшем случае (когда срабатывает fixed) она требует лишь линейного прохода по данным и одного декодирования; в худшем – выполняет до трёх попыток декодирования и две-три сортировки массива надёжностей. При этом каскадная стратегия не нуждается в предварительном обучении и может быть реализована с теми же требованиями к памяти, что и фиксированная.

Информация по оценке сложностей сведена в таблице 3.

Таблица 3 - Сводная таблица вычислительной сложности стратегий декодирования

DOI: https://doi.org/10.60797/itech.2026.11.1.4

| Стратегия | Память | Попыток декодирования (ср./макс) | Доп. вычисления | Относительное время* |
|---|---|---|---|---|
| Базовая | O(n) | 1,0/1 | нет | 1,0× |
| Фиксированная | O(n) | 1,0/1 | линейный проход O(n) | 1,1-1,2× |
| Интеллектуальная (упрощенная) | O(n) | 1,15/42 | 1 сортировка O(n logn)+fall-back | 2,0-2,5× |
| Интеллектуальная (полная) | O(n·m)** | 1,15/42 | 50×*betainv*+сортировка | 50-100× |
| Комбинированная | O(n) | 1,5/3 | 2-3 сортировки O(n logn) | 1,5-3,0× |

*Примечание: *относительно базовой с кратным (×) увеличением времени обработки; ** – количество пользователей (хранение индивидуальных профилей)*

Интеллектуальная стратегия в полной версии требует хранения индивидуальных профилей для каждого пользователя (память O(n·m), где m — число пользователей) и выполнения этапа обучения, в ходе которого методом полного перебора подбираются оптимальные весовые коэффициенты компонент надёжности. Интеллектуальная стратегия в упрощённой версии с фиксированным τ=0,2 и равными весами также требует формирования профиля пользователя (центры квантования, коэффициенты вариации $CV_i$ и квантили $Q_1$, $Q_3$), однако этот процесс не связан с пробными декодированиями и выполняется однократно по 5 эталонным измерениям с линейной вычислительной сложностью O(n).

На основе проведённого анализа вычислительной сложности и эффективности различных стратегий декодирования можно сформулировать следующие рекомендации для практической реализации.

Для большинства биометрических систем оптимальным выбором является комбинированная (каскадная) стратегия с кодом РС(63,15), которая обеспечивает 99.2% успешности декодирования при умеренной вычислительной нагрузке (в 1,5–3 раза выше базовой) и не требует предварительного обучения, что критически важно для сценариев с быстрой регистрацией новых пользователей.

В системах с жёсткими ограничениями по энергопотреблению или вычислительным ресурсам (мобильные устройства, встраиваемые системы) предпочтительна упрощённая интеллектуальная стратегия с фиксированным τ=0.2 в сочетании с кодом РС(63,11). Она демонстрирует 99,8% успешности при нагрузке всего 13, что в 2.5 раза меньше комбинированной стратегии, однако требует однократного формирования профиля пользователя по эталонным измерениям. Для слабых кодов, таких как РС(63,21), комбинированная стратегия остаётся безальтернативным лидером (93,3% против 84,3% у интеллектуальной), несмотря на более высокую нагрузку.

Полная версия интеллектуальной стратегии с оптимизацией доли стираний $\tau^*$ через бета-аппроксимацию признана нецелесообразной для практического применения из-за многократного роста вычислительной сложности в (50–100×) при отсутствии значимого выигрыша в успешности. Однако при наличии 10 и более эталонных измерений интеллектуальная стратегия с fall-back становится предпочтительной, достигая сопоставимой успешности с нагрузкой 13,6 против 55 у комбинированной.

Комбинированная и фиксированная стратегии не нуждаются в предварительном обучении и могут применяться непосредственно после получения эталонных измерений.

**Заключение**

Оценка метода декодирования проводилась с применением синтетических данных для кодов РС(63,21), РС(63,15) и РС(63,11), а также с использованием биометрии человеческих лиц, взятых из базы данных UvA-NEMO [5].

 

Проведенное исследование продемонстрировало, что использование реальных биометрических данных позволило не только подтвердить эффективность разработанных алгоритмов, но и выявить их преимущества, которые оставались незамеченными при работе только с синтетической информацией. В частности, на реальных данных интеллектуальная стратегия показала результаты, соответствующие синтетическому шуму $\sigma \approx 0.6$-$0.65$, что подтверждает адекватность выбранных моделей и значимость персонализации.

Разработанные методы обеспечивают прирост успешности декодирования относительно базового алгоритма без учета стираний от 9,5% для сильного кода РС(63,11) до 43,1% для слабого кода РС(63,21). Комбинированная (каскадная) стратегия достигла 93,3% успешности для РС(63,21), а интеллектуальная стратегия с fall-back механизмом — 96,3% при нагрузке в 4 раза меньшей.

Экспериментально установлено, что адаптивная агрегация весов компонент надежности не дает значимого выигрыша (разница менее 1%), что позволяет использовать упрощенную версию с равными весами. Фиксированная доля стираний $\tau=0.2$ оказалась близкой к оптимальной для всех рассмотренных кодов, а применение fall-back механизма обеспечивает дополнительный прирост успешности до 12% для слабых кодов без существенного увеличения средней нагрузки. При решении задачи аутентификации предложенные методы позволяют достичь FRR менее 1% для кода РС(63,11) и FRR менее 2% для РС(63,15), что соответствует требованиям большинства практических приложений.

На основе анализа результатов экспериментов на реальных биометрических данных и синтетических данных можно сформулировать следующие рекомендации по выбору кода и стратегии декодирования.

– Для систем, требующих максимальной надежности (успешность >99%, FRR<1%), оптимальным выбором является код РС(63,11) в сочетании с интеллектуальной стратегией (равные веса, только стирания, fall-back механизм), которая обеспечивает 99,8% успешности при минимальной нагрузке 13 и сохраняет эффективность даже при высоком уровне шума ($\sigma=1,0$ в синтетических экспериментах).

– Для систем с ограниченными вычислительными ресурсами, где важен баланс между надежностью и сложностью, рекомендуется код РС(63,15) с той же интеллектуальной стратегией, демонстрирующий 98.3% успешности на реальных данных при нагрузке 13. Эта конфигурация обеспечивает оптимальное соотношение эффективности и вычислительных затрат.

– Код РС(63,21) с каскадной стратегией целесообразно применять в сценариях с очень высоким уровнем шума (эквивалентным $\sigma \geq 1,25$ в синтетике), когда критически важна максимальная успешность (93.3% на реальных данных), а вычислительные ресурсы не ограничены. При наличии 10 и более эталонных измерений интеллектуальная стратегия с fall-back становится предпочтительной, достигая 95–97% успешности при нагрузке 13.6.

– Во всех случаях ключевыми элементами успеха являются формирование индивидуальных профилей пользователей по эталонным измерениям и обязательное применение fall-back механизма.

– Кроме того, в отличие от ранее исследованного на базе UvA-NEMO подхода с каскадными кодами [5], повышающего корректирующую способность ценой кратного роста сложности и длины ключа, предложенный метод интеллектуального управления стираниями решает ту же задачу за счет более точной оценки надежности символов на основе персонализированного профиля пользователя.

– Достигнутые результаты открывают ряд перспективных направлений для развития предложенного метода декодирования кодов РС с интеллектуальным управлением стираниями:

– Расширение области применения. Валидация подхода на других биометрических модальностях (отпечатки пальцев, радужная оболочка, голос) для создания многомодальных систем.

– Совершенствование статистического профилирования. Разработка методов коррекции оценок распределения при малых выборках с использованием байесовских подходов и оптимизация бета-аппроксимации для снижения вычислительной сложности.

– Архитектурные преобразования. Переход к распределённым архитектурам на основе схем разделения секрета Шамира и интеграция с блокчейн-технологиями для децентрализованного хранения биометрических данных [20], [21].

**Благодарности**

Авторы выражают благодарность рецензенту за полезные рекомендации в представлении работы.

**Конфликт интересов**

Не указан.

**Рецензия**

Все статьи проходят рецензирование. Но рецензент или автор статьи предпочли не публиковать рецензию к этой статье в открытом доступе. Рецензия может быть предоставлена компетентным органам по запросу.

**Acknowledgement**

The authors express their gratitude to the reviewer for their helpful suggestions on the presentation of the paper.

**Conflict of Interest**

None declared.

**Review**

All articles are peer-reviewed. But the reviewer or the author of the article chose not to publish a review of this article in the public domain. The review can be provided to the competent authorities upon request.

## Список литературы / References

1. Osorio Roig D. Stable hash generation for efficient privacy-preserving face identification / D. Osorio Roig, C. Rathgeb, C. Busch [et al.] // IEEE Transactions on Biometrics Behavior and Identity Science. — 2021. — Vol. 99. — P. 1–16. — DOI: 10.1109/TBIOM.2021.3100639.

 

2. Gilkalaye B.P. Euclidean-distance based fuzzy commitment scheme for biometric template security / B.P. Gilkalaye, A. Rattani, R. Derakhshani // 2019 7th International Workshop on Biometrics and Forensics (IWBF). — Cancun, 2019. — P. 1–6. — DOI: 10.1109/IWBF.2019.8739177.

3. Han Vinck A.J. Coding Concepts and Reed-Solomon Codes / A.J. Han Vinck. — Essen : Institute for Experimental Mathematics, 2013. — 206 p.

4. Juels A. A fuzzy commitment scheme / A. Juels, M. Wattenberg // ACM Conference on Computer and Communications Security. — Singapore, 1999. — P. 28–36.

5. Assanovich B. Authentication system based on biometric data of smiling face from stacked autoencoder and concatenated Reed-Solomon codes / B. Assanovich, K. Kosarava // PRIP 2021. Communications in Computer and Information Science. — 2022. — Vol. 1562. — P. 205–219. — DOI: 10.1007/978-3-030-98883-8_15.

6. Mohan D.D. Significant feature based representation for template protection / D.D. Mohan, N. Sankaran, S. Tulyakov [et al.] // 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). — Long Beach, 2019. — P. 2389–2396. — DOI: 10.1109/CVPRW.2019.00293.

7. Adamovic S. Fuzzy commitment scheme for generation of cryptographic keys based on iris biometrics / S. Adamovic, M. Milosavljevic, M. Veinovic [et al.] // IET Biom. — 2017. — Vol. 6. — № 2. — P. 89–96. — DOI: 10.1049/iet-bmt.2016.0061.

8. Dodis Y. Fuzzy extractors: How to generate strong keys from biometrics and other noisy data / Y. Dodis, R. Ostrovsky, L. Reyzin [et al.] // SIAM Journal on Computing. — 2008. — Vol. 38. — № 1. — P. 97–139. — DOI: 10.1137/060651380.

9. Kurbatov O. Unforgettable Fuzzy Extractor: Practical Construction and Security Model / O. Kurbatov, D. Zakharov, L. Antadze [et al.] // IACR Cryptology ePrint Archive. — 2025. — Art. 1799.

10. Chen L. Face template protection using deep LDPC codes learning / L. Chen, G. Zhao, J. Zhou [et al.] // IET Biometrics. — 2019. — Vol. 8. — № 3. — P. 190–197. — DOI: 10.1049/iet-bmt.2018.5156.

11. Assanovich B. Biometric database based on HOG structures and BCH codes / B. Assanovich, Yu. Veretilo // Proceedings of Information Technologies and Systems 2017 (ITS 2017). — Minsk, 2017. — P. 286–287.

12. Rao K.D. Channel Coding Techniques for Wireless Communications / K.D. Rao. — New Delh : Springer, 2015. — 394 p. — DOI: 10.1007/978-81-322-2292-7.

13. Шарипов Р.Р. Разработка программного комплекса реализации алгоритма Берлекэмпа-Месси на простых регистрах сдвига с линейной обратной связью для обучающихся по дисциплине «Криптография» / Р.Р. Шарипов, А.А. Кассирова // Computational Nanotechnology. — 2025. — Т. 12. — № 1. — С. 97–104. — DOI: 10.33693/2313-223X-2025-12-1-97-104.

14. Atieno L. An adaptive Reed-Solomon error-and-erasures decoder / L. Atieno, J. Allen, D. Goeckel [et al.] // Proceedings of the 2006 ACM/SIGDA 14th International Symposium on Field Programmable Gate Arrays (FPGA 2006). — Monterey, 2006. — P. 150–158. — DOI: 10.1145/1117201.1117224.

15. An X. High-Efficient Reed-Solomon Decoder Based on Deep Learning / X. An, Y. Liang, W. Zhang // 2020 IEEE Int. Symposium on Circuits and Systems (ISCAS). — Seville, 2020. — P. 1–5. — DOI: 10.1109/ISCAS45731.2020.9181187.

16. Zhang W. Iterative Soft Decoding of Reed-Solomon Codes Based on Deep Learning / W. Zhang, S. Zou, Y. Liu // IEEE Communications Letters. — 2020. — Vol. 24. — № 9. — P. 1991–994.

17. Immler V. New insights to key derivation for tamper-evident physical unclonable functions / V. Immler, K. Uppund // IACR Transactions on Cryptographic Hardware and Embedded Systems. — 2019. — Vol. 3. — P. 30–65. — DOI: 10.46586/tches.v2019.i3.30-65.

18. Иванов А.И. Оценка вероятности ошибок биометрической аутентификации на малых выборках, использующая гипотезу бета-распределения расстояний Хэмминга / А.И. Иванов, А.В. Безяев, А.В. Елфимов [и др.] // Специальная техника. — 2017. — № 1. — С. 48–51.

19. Weber J.H. Asymptotic Single-Trial Strategies for GMD Decoding with Arbitrary Error-Erasure Tradeoff / J.H. Weber, V.R. Sidorenko, C. Senger [et al.] // Problems of Information Transmission. — 2012. — Vol. 48. — P. 324–333. — DOI: 10.1134/S0032946012040023.

20. Богаченко Н.Ф. Схема разделения секрета между иерархически связанными участниками / Н.Ф. Богаченко // Математические структуры и моделирование. — 2022. — № 4 (64). — С. 117–121.

21. Утеев Г. Разработка децентрализованной системы идентификации личности по биометрическим данным с помощью технологии блокчейн и компьютерного зрения / Г. Утеев, Р.Ф. Гибадуллин // Международный научно-исследовательский журнал. — 2024. — № 4 (142). — DOI: 10.23670/IRJ.2024.142.6.

### Список литературы на английском языке / References in English

1. Osorio Roig D. Stable hash generation for efficient privacy-preserving face identification / D. Osorio Roig, C. Rathgeb, C. Busch [et al.] // IEEE Transactions on Biometrics Behavior and Identity Science. — 2021. — Vol. 99. — P. 1–16. — DOI: 10.1109/TBIOM.2021.3100639.

2. Gilkalaye B.P. Euclidean-distance based fuzzy commitment scheme for biometric template security / B.P. Gilkalaye, A. Rattani, R. Derakhshani // 2019 7th International Workshop on Biometrics and Forensics (IWBF). — Cancun, 2019. — P. 1–6. — DOI: 10.1109/IWBF.2019.8739177.

3. Han Vinck A.J. Coding Concepts and Reed-Solomon Codes / A.J. Han Vinck. — Essen : Institute for Experimental Mathematics, 2013. — 206 p.

4. Juels A. A fuzzy commitment scheme / A. Juels, M. Wattenberg // ACM Conference on Computer and Communications Security. — Singapore, 1999. — P. 28–36.

 

5. Assanovich B. Authentication system based on biometric data of smiling face from stacked autoencoder and concatenated Reed-Solomon codes / B. Assanovich, K. Kosarava // PRIP 2021. Communications in Computer and Information Science. — 2022. — Vol. 1562. — P. 205–219. — DOI: 10.1007/978-3-030-98883-8_15.

6. Mohan D.D. Significant feature based representation for template protection / D.D. Mohan, N. Sankaran, S. Tulyakov [et al.] // 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). — Long Beach, 2019. — P. 2389–2396. — DOI: 10.1109/CVPRW.2019.00293.

7. Adamovic S. Fuzzy commitment scheme for generation of cryptographic keys based on iris biometrics / S. Adamovic, M. Milosavljevic, M. Veinovic [et al.] // IET Biom. — 2017. — Vol. 6. — № 2. — P. 89–96. — DOI: 10.1049/iet-bmt.2016.0061.

8. Dodis Y. Fuzzy extractors: How to generate strong keys from biometrics and other noisy data / Y. Dodis, R. Ostrovsky, L. Reyzin [et al.] // SIAM Journal on Computing. — 2008. — Vol. 38. — № 1. — P. 97–139. — DOI: 10.1137/060651380.

9. Kurbatov O. Unforgettable Fuzzy Extractor: Practical Construction and Security Model / O. Kurbatov, D. Zakharov, L. Antadze [et al.] // IACR Cryptology ePrint Archive. — 2025. — Art. 1799.

10. Chen L. Face template protection using deep LDPC codes learning / L. Chen, G. Zhao, J. Zhou [et al.] // IET Biometrics. — 2019. — Vol. 8. — № 3. — P. 190–197. — DOI: 10.1049/iet-bmt.2018.5156.

11. Assanovich B. Biometric database based on HOG structures and BCH codes / B. Assanovich, Yu. Veretilo // Proceedings of Information Technologies and Systems 2017 (ITS 2017). — Minsk, 2017. — P. 286–287.

12. Rao K.D. Channel Coding Techniques for Wireless Communications / K.D. Rao. — New Delh : Springer, 2015. — 394 p. — DOI: 10.1007/978-81-322-2292-7.

13. Sharipov R.R. Razrabotka programmnogo kompleksa realizacii algoritma Berlekempa-Messi na prostyh registrah sdviga s linejnoj obratnoj svyaz'yu dlya obuchayushchihsya po discipline "Kriptografiya" [Development of a Software Package for the Implementation of the Algorithm Berlecamp – Messy on Simple Shift Registers with Linear Feedback for Students of the Discipline "Cryptography"] / R.R. Sharipov, A.A. Kassirova // Computational Nanotechnology. — 2025. — Vol. 12. — № 1. — P. 97–104. — DOI: 10.33693/2313-223X-2025-12-1-97-104. [in Russian]

14. Atieno L. An adaptive Reed-Solomon error-and-erasures decoder / L. Atieno, J. Allen, D. Goeckel [et al.] // Proceedings of the 2006 ACM/SIGDA 14th International Symposium on Field Programmable Gate Arrays (FPGA 2006). — Monterey, 2006. — P. 150–158. — DOI: 10.1145/1117201.1117224.

15. An X. High-Efficient Reed-Solomon Decoder Based on Deep Learning / X. An, Y. Liang, W. Zhang // 2020 IEEE Int. Symposium on Circuits and Systems (ISCAS). — Seville, 2020. — P. 1–5. — DOI: 10.1109/ISCAS45731.2020.9181187.

16. Zhang W. Iterative Soft Decoding of Reed-Solomon Codes Based on Deep Learning / W. Zhang, S. Zou, Y. Liu // IEEE Communications Letters. — 2020. — Vol. 24. — № 9. — P. 1991–994.

17. Immler V. New insights to key derivation for tamper-evident physical unclonable functions / V. Immler, K. Uppund // IACR Transactions on Cryptographic Hardware and Embedded Systems. — 2019. — Vol. 3. — P. 30–65. — DOI: 10.46586/tches.v2019.i3.30-65.

18. Ivanov A.I. Otsenka veroyatnosti oshibok biometricheskoi autentifikatsii na malikh viborkakh, ispolzuyushchaya gipotezu beta-raspredeleniya rasstoyanii Khemminga [Estimating the probability of biometric authentication errors on small samples using the hypothesis of beta distribution Hamming distances] / A.I. Ivanov, A.V. Bezyaev, A.V. Elfimov [et al.] // Spetsialnaya tekhnika [Special equipment]. — 2017. — № 1. — P. 48–51. [in Russian]

19. Weber J.H. Asymptotic Single-Trial Strategies for GMD Decoding with Arbitrary Error-Erasure Tradeoff / J.H. Weber, V.R. Sidorenko, C. Senger [et al.] // Problems of Information Transmission. — 2012. — Vol. 48. — P. 324–333. — DOI: 10.1134/S0032946012040023.

20. Bogachenko N.F. Skhema razdeleniya sekreta mezhdu ierarkhicheski svyazannimi uchastnikami [Scheme for sharing a secret between hierarchically related participants] / N.F. Bogachenko // Matematicheskie strukturi i modelirovanie [Mathematical structures and modeling]. — 2022. — № 4 (64). — P. 117–121. [in Russian]

21. Uteev G. Razrabotka detsentralizovannoi sistemi identifikatsii lichnosti po biometricheskim dannim s pomoshchyu tekhnologii blokchein i kompyuternogo zreniya [Development of the decentralized biometric identity verification system using blockchain technology and computer vision] / G. Uteev, R.F. Gibadullin // Mezhdunarodnii nauchno-issledovatelskii zhurnal [International Research Journal]. — 2024. — № 4 (142). — DOI: 10.23670/IRJ.2024.142.6. [in Russian]