УДК 003.26+004.56+004.627

Н. В. Попеня, Д. М. Романенко
Белорусский государственный технологический университет

МЕТОД АУДИОСТЕГАНОГРАФИИ ДЛЯ AAC-СЖАТЫХ АУДИОСИГНАЛОВ НА ОСНОВЕ ЭХО-КОДИРОВАНИЯ И АДАПТИВНОГО КЕПСТРАЛЬНОГО АНАЛИЗА

Статья посвящена разработке и исследованию метода аудиостеганографии, предназначенного для работы с аудиосигналами, подверженными AAC-сжатию. Предлагаемый метод основан на эхо-кодировании, при котором бинарная информация встраивается путем добавления к сигналу его ослабленной копии с одной из двух фиксированных временных задержек. Ключевым элементом является алгоритм извлечения данных, использующий адаптивный кепстральный анализ. Особенность анализа заключается в динамическом вычислении порога детектирования эха для каждого аудиоблока на основе статистических характеристик (медианы и стандартного отклонения) его кепстра в областях, свободных от ожидаемых пиков. Такой адаптивный подход обеспечивает надежность обнаружения скрытой информации в условиях значительных искажений, вносимых AAC-сжатием. При встраивании данных в «тихие» сегменты различных типов аудиоконтента и последующем сжатии кодеком AAC (256 кбит/с) достигнута высокая точность восстановления информации, характеризуемая коэффициентом битовых ошибок (BER) менее 6% для большинства тестовых сигналов. Проведен сравнительный анализ BER для сценариев извлечения из исходного сигнала, после сохранения в формате WAV без потерь и после AAC-сжатия. Обсуждается влияние характеристик сигнала на эффективность метода, анализируется характер возникающих ошибок и обосновывается необходимость применения кодов коррекции ошибок для обеспечения безошибочного извлечения.

Ключевые слова: аудиостеганография, эхо-кодирование, кепстральный анализ, адаптивный порог, AAC-сжатие, устойчивость к сжатию.

Для цитирования: Попеня Н. В., Романенко Д. В. Метод аудиостеганографии для AAC-сжатых аудиосигналов на основе эхо-кодирования и адаптивного кепстрального анализа // Труды БГТУ. Сер. 3, Физико-математические науки и информатика. 2025. № 2 (296). С. 110–119.
DOI: 10.52065/2520-6141-2025-296-14.

N. V. Popenya, D. M. Romanenko
Belarusian State Technological University

AN AUDIO STEGANOGRAPHY METHOD FOR AAC-COMPRESSED AUDIO SIGNALS BASED ON ECHO HIDING AND ADAPTIVE CEPSTRAL ANALYSIS

This article is dedicated to the development and research of an audio steganography method designed for operation with audio signals subjected to AAC compression. The proposed method is based on echo hiding, where binary information is embedded by adding an attenuated copy of the signal at one of two fixed time delays. A key element is the data extraction algorithm, which utilizes adaptive cepstral analysis. The distinctiveness of this analysis lies in the dynamic calculation of the echo detection threshold for each audio block, based on the statistical characteristics (median and standard deviation) of its cepstrum in regions free of expected peaks. This adaptive approach enhances the reliability of detecting hidden information under conditions of significant distortions introduced by AAC compression. The paper presents results of experimental validation, including Bit Error Rate (BER) assessment after AAC compression (256 kbps) on various audio content types. When embedding data into silent segments of various audio content types followed by AAC compression (256 kbps), a high data recovery accuracy was achieved, characterized by a Bit Error Rate (BER) of less than 6% for most test signals. A comparative BER analysis is provided for extraction scenarios from the original signal, after lossless WAV saving, and after AAC compression. The influence of signal characteristics on the method's effectiveness is discussed, the nature of occurring errors is analyzed, and the necessity of employing error correction codes to ensure error-free extraction is substantiated.

Keywords: audio steganography, echo hiding, cepstral analysis, adaptive threshold, AAC compression, compression robustness.

For citation: Popenya N. V., Romanenko D. M. An audio steganography method for AAC-compressed audio signals based on echo hiding and adaptive cepstral analysis. *Proceedings of BSTU, issue 3, Physics and Mathematics*. *Informatics*, 2025, no. 2 (296), pp. 110–119 (In Russian).
DOI: 10.52065/2520-6141-2025-296-14.

Введение. Обеспечение безопасности и конфиденциальности цифровой информации является одной из ключевых задач современной информатики [1]. Наряду с криптографией, скрывающей содержимое сообщения, активно развивается стеганография – область знаний, посвященная сокрытию самого факта существования тайного канала связи [2]. Основная цель стеганографии – встроить секретное сообщение в цифровой контейнер (например, изображение, аудио- или видеофайл) таким образом, чтобы наличие скрытых данных было невозможно обнаружить стандартными средствами или человеческим восприятием [3].

Видеофайлы, объединяющие визуальную и звуковую информацию, представляют особый интерес в качестве стеганографических контейнеров из-за их большого объема и широкого распространения. Аудиопоток видеофайла обладает значительной перцептивной избыточностью, которую можно использовать для встраивания дополнительной информации [4].

Однако широкое применение алгоритмов сжатия аудио с потерями, таких как AAC (Advanced Audio Coding) [5], создает серьезные препятствия для многих стеганографических методов. Простейшие подходы, например метод наименее значащего бита (LSB), модифицирующий младшие биты амплитудных отсчетов, оказываются крайне неустойчивыми, поскольку информация, внесенная таким способом, необратимо искажается или теряется в процессе квантования и удаления перцептивно незначимых компонент сигнала кодеком [4, 6]. Это существенно ограничивает практическую применимость LSB-методов в реальных сценариях передачи или хранения аудиоданных, особенно если данные подвергаются последующей обработке или сжатию.

Для повышения устойчивости стеганографических систем к сжатию разрабатываются более сложные методы, оперирующие в других областях представления сигнала или использующие его специфические свойства. Одним из таких методов является эхо-кодирование (echo hiding) [7]. Суть метода заключается во внесении в исходный сигнал его копии (эха), ослабленной по амплитуде и задержанной на определенное время. Скрываемая информация кодируется путем выбора одного из нескольких возможных значений временной задержки эха. Предполагается, что короткое эхо с малой амплитудой будет незаметно для слушателя благодаря эффектам маскировки слуховой системы [8], но при этом может быть обнаружено при извлечении с помощью автокорреляционного или кепстрального анализа [9]. Кепстральный анализ позволяет выявить периодичность, внесенную эхом, в виде пика в кепстре на соответствующей задержке. Эхо-кодирование потенциально более устойчиво к сжатию, чем LSB, так как оно модифицирует глобальные характеристики сигнала, а не отдельные биты отсчетов.

Тем не менее надежное детектирование пиков эха в кепстре после AAC-сжатия остается сложной задачей. Сжатие вносит искажения, которые могут ослаблять или маскировать пики эха, а также добавлять ложные пики, что приводит к ошибкам при извлечении данных.

Целью исследования является экспериментальный анализ и апробация метода аудиостеганографии видеоконтейнеров, основанного на эхокодировании, с акцентом на повышении надежности извлечения данных после AAC-сжатия. Для этого предлагается использовать детектор на основе кепстрального анализа с применением адаптивного порога шума, динамически подстраивающегося под характеристики каждого анализируемого аудиоблока. В работе представляются результаты оценки точности восстановления данных после сжатия кодеком AAC (256 кбит/с) и обсуждается применимость метода.

Основная часть. Предлагаемый метод аудиостеганографии предназначен для встраивания скрытой информации M в аудиокомпоненту (аудиопоток A_{in}) цифровых видеоконтейнеров V_{in}. Современные видеофайлы, такие как MP4, AVI, MKV и другие, представляют собой мультимедийные контейнеры, содержащие несколько потоков данных, включая видеоряд (последовательность изображений) и одну или несколько аудиодорожек. Именно такая аудиодорожка, извлеченная из видеофайла и представленная в виде последовательности цифровых отсчетов (сэмплов), служит контейнером для скрытия информации. Реализация метода включает несколько ключевых этапов, обеспечивающих как конфиденциальность и целостность передаваемых данных, так и их скрытность и устойчивость к последующей обработке, в частности к сжатию с потерями. Для обеспечения конфиденциальности и целостности исходное сообщение M предварительно шифруется с использованием алгоритма AES в режиме GCM (Galois/Counter Mode) [10, 11]. Этот режим генерирует как шифротекст, так и тег аутентификации, объединенные в токен C. Для повышения надежности к данным добавляется избыточность с применением специальных корректирующих кодов, что позволяет исправить определенное количество ошибок, возникающих при передаче или сжатии. Также формируется служебный заголовок H, содержащий метаданные, необходимые для корректного извлечения (как минимум, длину исходного шифротекста). Итоговый блок данных,

состоящий из заголовка H и шифротекста C, затем обрабатывается алгоритмом помехоустойчивого кодирования, в результате чего к нему добавляются контрольные символы (или символы избыточности). Полученный блок данных преобразуется в битовый поток E для последующего встраивания.

Ключевой особенностью предлагаемого метода является обеспечение повышенной устойчивости к сжатию с потерями при одновременном сохранении скрытности встраиваемых данных. Устойчивость достигается за счет использования эхо-кодирования, которое модифицирует кепстральные характеристики сигнала [7]. Скрытность обеспечивается путем встраивания данных преимущественно в сегменты аудиосигнала с низкой энергией («тихие» участки), где внесенные изменения менее заметны для человеческого слуха [8]. Однако, как показали предварительные эксперименты, AAC-сжатие может изменять характеристики этих «тихих» сегментов, что приводит к проблеме синхронизации при извлечении, если полагаться на повторный анализ тишины. Для решения этой проблемы предлагается использовать синхромаркеры – специальные сигналы, встраиваемые перед блоками данных и устойчивые к сжатию.

Алгоритм встраивания информации можно представить в виде последовательности из нескольких действий.

1. Осуществляется анализ исходного монофонического аудиосигнала $x(t)$ (полученного из A_{in}) для поиска участков, подходящих для скрытия данных. Используется метод, основанный на оценке кратковременной энергии сигнала. Сигнал разбивается на перекрывающиеся временные фреймы длительностью T_{frame} (например, 30 мс) с шагом T_{hop} (T_{frame} / 2). Для каждого i-го фрейма $x_i(t)$ вычисляется его спектр с помощью быстрого преобразования Фурье (БПФ) с применением оконной функции (Ханна $w(t)$) для сглаживания краев: $X_i(f) = FFT\{x_i(t)w(t)\}$. Энергия фрейма P_i оценивается как сумма квадратов амплитуд спектральных компонент: $P_i = \Sigma|X_i(f)|^2$.

2. Фреймы, энергия которых P_i оказывается ниже заданного порога Th_{energy}, помечаются как «тихие». Этот порог является важным параметром, влияющим на количество найденных участков и их характеристики. Последовательности смежных тихих фреймов объединяются в непрерывные «тихие» сегменты. Для дальнейшей обработки отбираются только те сегменты, чья длительность превышает минимально допустимые пороги (Dur_{min}, Dur_{embed}), чтобы обеспечить достаточное пространство для встраивания блока данных или синхромаркера и избежать артефактов на слишком коротких участках. Результатом этого этапа является список временных интервалов $\{(s_1, e_1), (s_2, e_2), ...\}$, соответствующих границам (начало s, конец e в сэмплах) подходящих тихих сегментов. На рис. 1 приведен пример результатов анализа тишины для фрагмента одного из тестовых аудиосигналов. Предложенный алгоритм анализа успешно идентифицирует участки сигнала с низкой энергией. Логарифмический масштаб по оси энергии на нижнем графике позволяет наглядно оценить значительное падение энергии в этих сегментах по сравнению с участками активной речи.

3. Для непосредственного скрытия битов подготовленного потока данных E в выбранные «тихие» сегменты используется метод эхо-кодирования. Принцип данного метода заключается во внесении в исходный аудиосигнал искусственного эха – его копии, ослабленной по амплитуде и смещенной во времени. Бинарная информация («0» или «1») кодируется путем выбора одной из двух предопределенных временных задержек эха.

4. Пусть для кодирования бита «0» выбрана задержка $d0$, а для бита «1» – задержка $d1$ (где $d0 \neq d1$, и обе задержки достаточно малы, например, в диапазоне 15–75 мс, чтобы эхо было перцептивно замаскировано, но достаточно велико для разрешения кепстром). Выбранный для встраивания «тихий» сегмент аудиосигнала $x(t)$ разбивается на блоки фиксированной длины T_{seg} (например, 100 мс), соответствующей длине, используемой при кепстральном анализе. Для каждого j-го блока $x_j(t)$, в который необходимо встроить бит b_j, формируется модифицированный блок $x'_j(t)$ по формуле

$$x'_j(t) = x_j(t) + \alpha \times x_j(t - d_{bj}),$$

где d_{bj} равно d_0, если $b_j = 0$, и $d1$, если $b_j = 1$, параметр α – это коэффициент ослабления эха (например, 0,6–0,8), который определяет баланс между устойчивостью (более сильное эхо легче детектировать) и незаметностью (слабое эхо менее слышимо). Важно отметить, что при добавлении эха амплитуда результирующего сигнала может превысить допустимый диапазон (например, [–1, 1] для нормализованных данных), поэтому после суммирования необходимо применять клиппирование (ограничение значений). Модифицированные блоки $x'_j(t)$ затем заменяют исходные блоки в общем аудиопотоке.

Далее рассмотрим процесс извлечения скрытых бит из принятого (и потенциально искаженного AAC-сжатием) аудиосигнала $x''(t)$, который выполняется с помощью кепстрального анализа. Наличие эха с задержкой d в сигнале $x(t)$ приводит к появлению пика в кепстре $c(\tau)$ этого сигнала при «квефренсии» τ, равной задержке d.

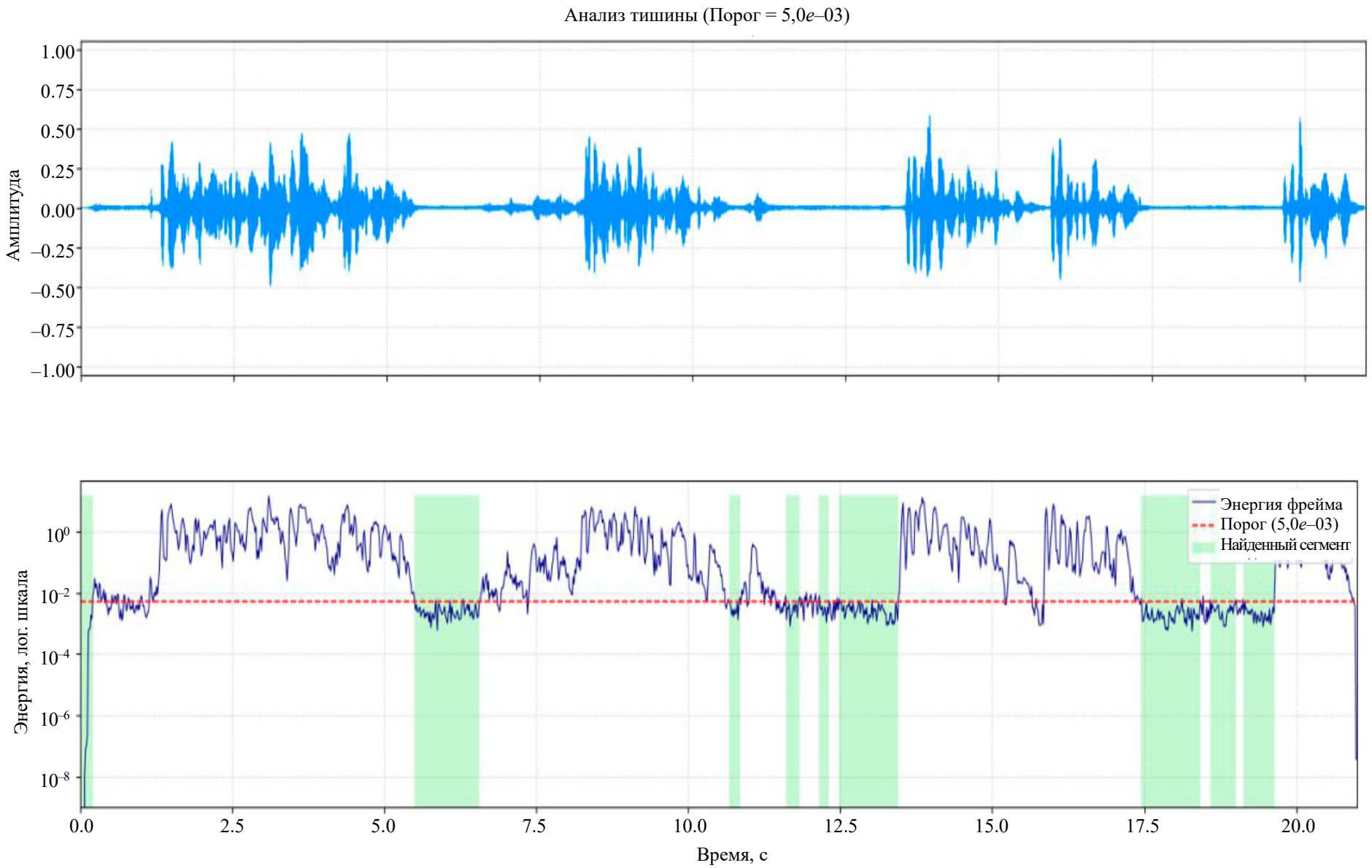


Рис. 1. Пример результатов анализа тишины для фрагмента аудиосигна

Для каждого блока $x''_j(t)$ процедура извлечения информации включает ряд последовательных действий.

1. Вычисление кепстра $c_j(\tau)$ блока $x''_j(t)$.

2. Вычисление адаптивного порога шума NF_j. Для этого анализируется сам кепстр $c_j(\tau)$ в областях, где заведомо нет ожидаемых пиков эха (т. е., исключая окрестности *d0*, *d1* и область низких квефренсий). По значениям кепстра в этих «шумовых» областях вычисляется статистическая оценка, например, $NF_j = median(c_{noise}) + k_{std} \times \times std(c_{noise})$. Это позволяет порогу подстраиваться под текущий уровень шума кепстра конкретного блока.

3. Поиск пиков. В кепстре $c_j(\tau)$ ищутся локальные максимумы *Peak0* и *Peak1* в небольших окрестностях ожидаемых квефренсий *d0* и *d1* соответственно.

4. Проверка валидности пиков. Найденные пики сравниваются с адаптивным порогом шума: $is_{valid0} = Peak0 > NF_j$, $is_{valid1} = Peak1 > NF_j$.

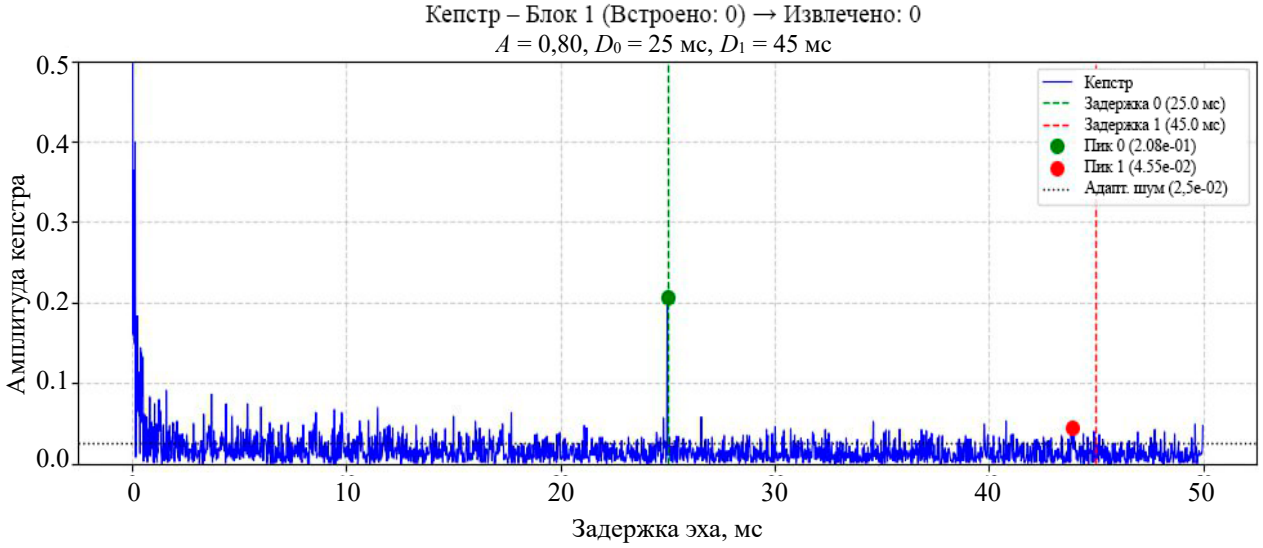
5. Принятие решения. Используется модифицированный детектор:

– если валиден только *Peak1*, декодируется бит «1»;

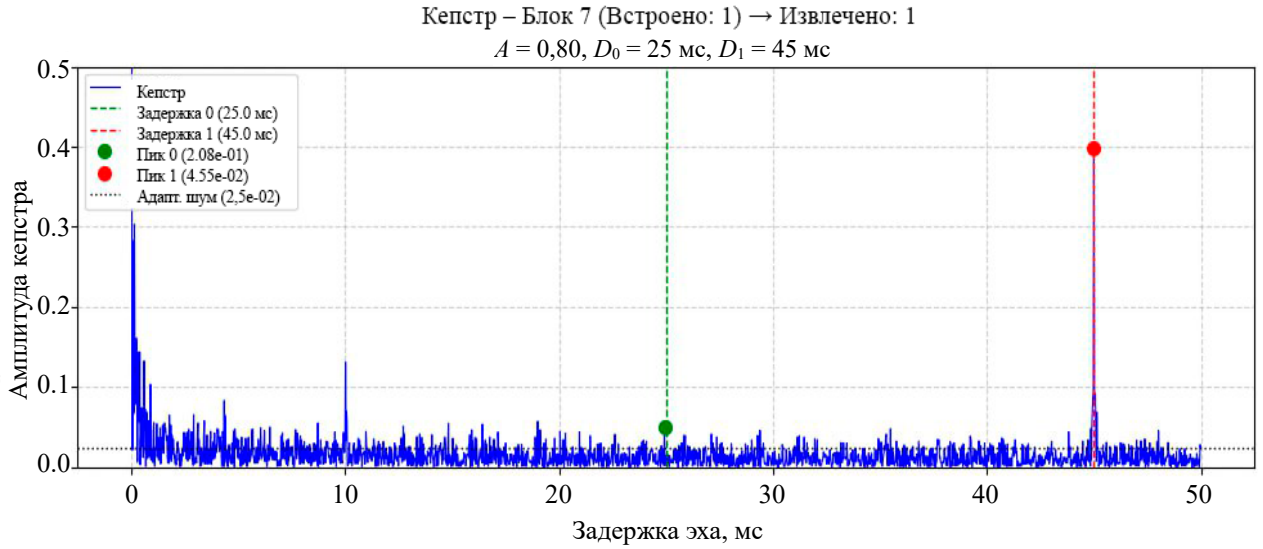
– если валиден только *Peak0*, декодируется бит «0»;

– если оба пика валидны, бит определяется по тому, какой пик выше: «1», если *Peak1* > *Peak0*, иначе «0» (рис. 2);

– если оба пика невалидны (ниже адаптивного порога), декодируется бит «0».



а



б

Рис. 2. Иллюстрация работы адаптивного кепстрального детектора:
а – декодирование бита «0»; *б* – декодирование бита «1»

Вертикальные линии – ожидаемые квефренсии для *d0* и *d1*. Маркеры – найденные пики *Peak0* и *Peak1* (закрашенные – валидные, т. е. выше адаптивного порога). Горизонтальная пунктирная линия – адаптивный порог шума NF_j, вычисленный для данного блока.

Подход с адаптивным порогом, позволяющий детектору подстраиваться под локальные характеристики кепстра. Такая адаптация обеспечивает лучшую надежность по сравнению с использованием единого фиксированного порога для всего сигнала, особенно при наличии вариаций уровня шума кепстра между различными аудиоблоками, что характерно для сигналов после сжатия с потерями.

Тем не менее применение AAC-кодека модифицирует статистические и спектральные свойства аудиосигнала, в результате чего алгоритм поиска «тихих» сегментов, использованный при встраивании, примененный повторно на этапе извлечения, идентифицирует иной набор временных интервалов. Это делает невозможной точную синхронизацию при извлечении данных. Для обеспечения корректной синхронизации при извлечении предлагается использовать синхромаркеры. Перед каждым блоком аудио $x'_j(t)$, в который встроены данные методом эха, вставляется короткий, уникальный и устойчивый к искажениям маркерный сигнал $m(t)$.

Требование к маркеру $m(t)$ состоит в том, что он должен легко детектироваться в сигнале $x''(t)$ даже после AAC-сжатия и иметь низкую вероятность ложного срабатывания на участках обычного аудио. В качестве маркера используются сигналы с хорошими автокорреляционными свойствами, например шумоподобные, сгенерированные на основе псевдослучайных последовательностей (*M*-последовательности, коды Баркера, Голда).

Использование синхромаркеров позволяет отказаться от анализа тишины при извлечении и точно локализовать блоки данных, но требует разработки надежного маркера (детектора) и вносит дополнительную избыточность, уменьшая полезную емкость стегоканала.

Для оценки работоспособности и устойчивости предложенного метода эхо-кодирования с кепстральным анализом и адаптивным детектором проведена серия вычислительных экспериментов. Тестирование выполнялось на реальном аудиовизуальном контенте – видеофайлах формата MP4, содержащем видеодорожку (H.264) и стереоаудиодорожку с частотой дискретизации 44,1 кГц (в экспериментах использовался один канал).

В качестве контейнеров использовался набор из видеофайлов формата MP4 (кодек видео H.264, кодек аудио исходный – AAC, частота дискретизации аудио 44,1 кГц, стерео). Длительность тестовых фрагментов варьировалась от десятков секунд до нескольких минут. Файлы были подобраны таким образом, чтобы представить различные типы аудиоконтента:

– содержащие записи голоса с различным качеством записи и уровнем фонового шума;

– включающие фрагменты различных жанров музыки, характеризующиеся разной динамикой, спектральной насыщенностью и наличием (отсутствием) пауз;

– содержащие записи окружающей среды с преобладанием стационарных или квазистационарных шумов (шум улицы, гул аппаратуры, природные звуки) при отсутствии доминирующего полезного сигнала;

– объединяющие речь, музыку и фоновые шумы.

Программная реализация алгоритмов выполнена на языке Python с использованием библиотек moviepy для работы с медиафайлами, numpy и scipy для цифровой обработки сигналов (БПФ, ОБПФ, кепстр), soundfile для промежуточного сохранения аудио.

На первом этапе проверялась корректность работы алгоритмов встраивания и извлечения в идеальных условиях, без влияния сжатия с потерями. Была сгенерирована псевдослучайная битовая последовательность известной длины ($N = 50$ бит). Исходный аудиосигнал был проанализирован для поиска «тихих» сегментов с использованием порога энергии $Th_{\text{energy}} = 0{,}005$ и минимальной длительности сегмента для обработки $Dur_{\text{embed}} = 150$ мс (при длине фрейма 30 мс и минимальной длине тишины 75 мс). Найденные сегменты были разбиты на блоки фиксированной длины $T_{\text{seg}} = 100$ мс. Для эхо-кодирования были выбраны следующие параметры: коэффициент ослабления эха $\alpha = 0{,}8$; задержка для бита «0» *d0* = 25 мс, задержка для бита «1» *d1* = 45 мс.

Тестовые биты были встроены в соответствующие блоки копии исходного аудиосигнала. Затем из этого же модифицированного массива в памяти были извлечены биты. Сравнение извлеченной последовательности с исходной показало точность 100%.

Далее модифицированный аудиосигнал был сохранен во временный файл формата WAV (без потерь) и немедленно загружен обратно. Из загруженных данных снова были извлечены биты с использованием тех же параметров. Результат сравнения показал идентичную точность 100%. Это подтверждает, что операции сохранения (загрузки) в формате WAV и используемая библиотека soundfile не вносят существенных искажений, влияющих на кепстральный детектор эха. Высокая точность в этих тестах свидетельствует о корректной реализации базовых алгоритмов эхо-кодирования и адаптивного извлечения.

Основной целью экспериментов является оценка устойчивости метода к сжатию с потерями. Для этого был проведен тест, имитирующий реальный сценарий сохранения видеоконтейнера.

В рамках теста сгенерирована новая последовательность из 100 бит. Эти биты были встроены методом эха с теми же параметрами (*alpha* = 0,8; *d0* = 25; *d1* = 45; *len* = 100) в предварительно идентифицированные «тихие» сегменты моноверсии исходного аудиосигнала. Модифицированный аудиосигнал был затем мультиплексирован с исходным видеопотоком (перекодированным с libx264) и сохранен в файл MP4 с использованием аудиокодека AAC с битрейтом 256 кбит/с.

После сохранения стего-контейнер был снова загружен, аудиодорожка (уже прошедшая AAC-сжатие и декодирование) извлечена. Из этого аудиосигнала были извлечены биты эха. Анализировались участки, соответствующие исходным позициям «тихих» сегментов, использованных для встраивания, путем их разбиения на блоки и применения адаптивного кепстрального детектора. Экспериментальная оценка устойчивости метода проводилась на выборке видеофайлов, содержащих аудиоконтент различного типа: речь, музыка, фоновые шумы. Для понимания влияния характеристик контейнера на точность восстановления данных для каждого тестового файла были вычислены акустические признаки в пределах найденных «тихих» сегментов. Результаты оценки устойчивости предложенного метода к AAC-сжатию для различных типов аудиоконтента сведены в таблицу.

Результаты экспериментов по восстановлению данных после AAC-сжатия для различных типов аудиоконтента

Тип файла	RMS, норм.	Центр, Гц	Шир., Гц	Тиш., %	Точн. AAC, %
Смешанный звук	0,0015	6014	5578	45,8	94
Речь тихая	0,0013	7171	5736	62,1	98
Речь громкая	0,0025	6850	5612	18,5	98
Музыка тихая	0,0017	4955	4793	71,2	100
Музыка громкая	0,0031	3211	4432	12,6	98
Фоновый шум громкий	0,0040	3513	4809	11,8	98
Фоновый шум тихий	0,0014	2855	2607	15,3	100
Фоновый шум смешанный	0,0031	3660	4274	22,6	95

В таблице в качестве характеристик тихих сегментов были выбраны:

1) RMS – среднеквадратичное значение амплитуды, усредненное по всем найденным «тихим» участкам, отражающее средний уровень остаточного сигнала или шума;

2) спектральный центроид – среднее значение спектрального центроида по «тихим» сегментам, показывающее преобладающие частоты;

3) спектральная ширина – средняя спектральная ширина, характеризующая разброс энергии по спектру в «тихих» зонах;

4) тишина – процентное соотношение общей длительности всех найденных «тихих» сегментов к общей длительности аудиосигнала;

5) точность AAC – показатель итоговой точности восстановления бит данных, встроенных в блоки модифицированного аудиосигнала перед его AAC-сжатием, и последующего извлечения из сжатого-разжатого сигнала.

Анализ результатов на различных аудиоконтейнерах показал зависимость точности метода от характеристик сигнала. На файлах с относительно низким уровнем фонового шума в «тихих» сегментах (RMS ≈ 0,0013–0,0017) и простым спектральным составом (низкие значения спектральной ширины) была достигнута высокая точность восстановления данных после AAC-сжатия, составляющая 98–100%. Некоторое снижение точности до 94–95% наблюдалось для файлов со смешанным аудиоконтентом и для записи с выраженными реверберационными характеристиками (фоновый смешанный шум), несмотря на наличие достаточного количества «тихих» сегментов. Это свидетельствует о том, что не только уровень фонового шума, но и его спектральная сложность (например, более высокая спектральная ширина в смешанном звуке) или наличие специфических акустических эффектов (реверберация в смешанном фоновом шуме) в сегментах, выбранных для встраивания (даже если они формально «тихие»), влияют на стабильность кепстрального представления после AAC-сжатия и, как следствие, на надежность детектирования пиков эха. Более высокий или нестационарный уровень шума, а также наличие реверберации могут затруднять работу адаптивного порога или маскировать слабые пики эха, что приводит к увеличению количества ошибок декодирования. Тем не менее стабильно высокая точность (>94%) на разнообразных типах контента подтверждает робастность подхода с адаптивной детекцией по сравнению с методами, использующими фиксированные пороги.

Исследование характера ошибок показало, что в большинстве случаев они являются одиночными независимыми битовыми искажениями. Лишь в одном эксперименте был зафиксирован пакет ошибок длиной 2 бита. Преобладание одиночных ошибок свидетельствует о том, что AAC-сжатие вносит скорее локальные искажения в кепстр, влияющие на решение детектора для отдельных блоков, а не протяженные помехи. Наличие остаточных ошибок предполагает необходимость в дальнейшем использования кодов коррекции ошибок для гарантированного восстановления данных.

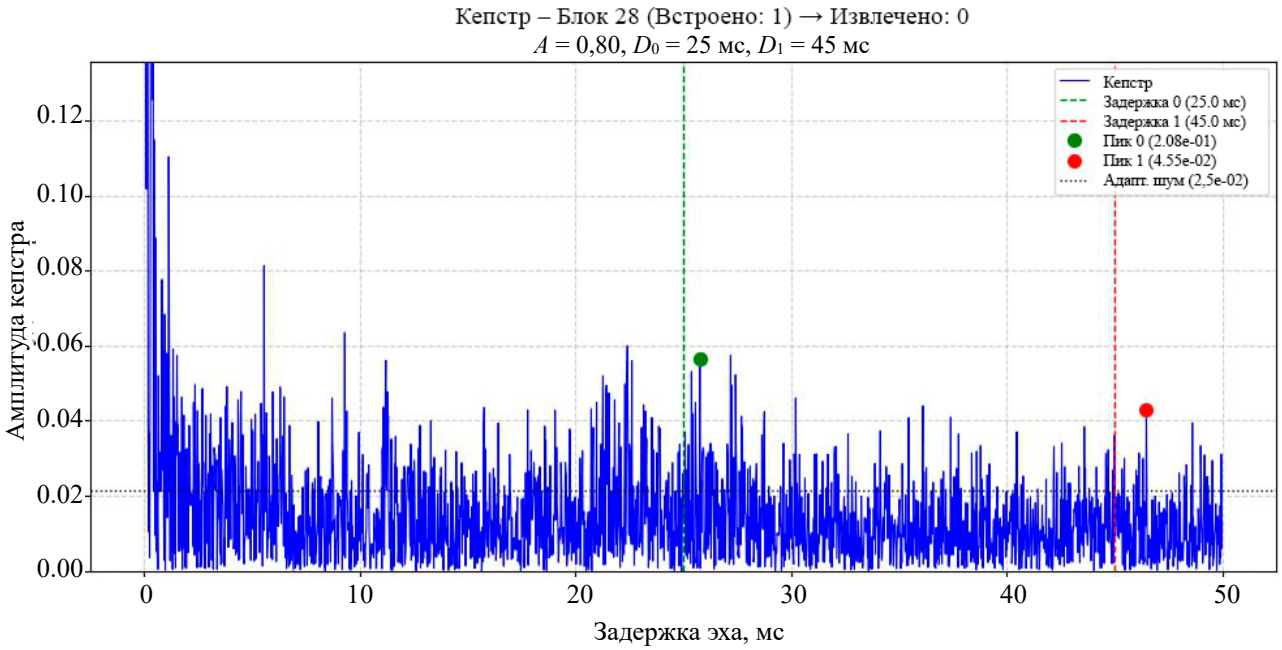


Рис. 3. Иллюстрация неверного декодирования бита

Полученные экспериментальные данные свидетельствуют о принципиальной работоспособности предложенного метода аудиостеганографии на основе эхо-кодирования. Тесты в идеальных условиях показали высокую точность восстановления данных, что подтверждает корректность реализации алгоритмов встраивания эха и его извлечения с помощью кепстрального анализа и адаптивного порога шума.

Ключевым результатом проведенного исследования является экспериментально подтвержденная высокая устойчивость предложенного метода эхо-кодирования к сжатию аудиокодеком AAC с битрейтом 256 кбит/с. Контрольные эксперименты по LSB-встраиванию, проведенные на аналогичных тестовых файлах и с тем же уровнем AAC-сжатия показали точность восстановления скрытых данных на уровне около 50%, что свидетельствует о существенной деградации скрытого. Высокая эффективность предложенного эхо-метода, превосходящая LSB в условиях сжатия, достигается благодаря использованию кепстрального анализа для обнаружения эха и применения адаптивного порога шума. Адаптивный порог, вычисляемый на основе статистик кепстра вне зон ожидаемых пиков, позволяет надежно отделять пики, соответствующие эху, от фонового шума кепстра, уровень которого повышается после сжатия.

Тем не менее наличие ошибок после AAC-сжатия указывает на существующие ограничения метода. Анализ графиков кепстра для ошибочных случаев показывает, что основной причиной ошибок является значительное ослабление или искажение пика эха в кепстре из-за процесса квантования и удаления информации кодеком AAC. В некоторых блоках пик эха может опускаться ниже адаптивно вычисленного порога шума, что приводит к неверному декодированию бита (рис. 3).

Выбор параметров метода определяется необходимостью найти компромисс между устойчивостью и скрытностью. Представленные результаты высокой устойчивости были получены при встраивании данных в «тихие» блоки аудиосигнала. Хотя эхо-кодирование само по себе считается относительно незаметным [7], добавление эха в «громкие», «нетихие» участки сигнала может повышать вероятность его обнаружения как специализированными алгоритмами стегоанализа, так и, возможно, при внимательном прослушивании.

Встраивание только в «тихие» сегменты повысит скрытность, но, как было показано, требует решения проблемы синхронизации при извлечении после сжатия. Разработка надежных синхромаркеров или методов встраивания информации о позициях является важным направлением для дальнейшего улучшения практической применимости метода.

Заключение. В статье предложен и исследован метод аудиостеганографии на основе эхо-кодирования, предназначенный для использования в видеоконтейнерах и обладающий повышенной устойчивостью к AAC-сжатию. Метод использует две временные задержки для кодирования бит и кепстральный анализ с адаптивным порогом шума для их извлечения.

Экспериментально доказана работоспособность метода и его устойчивость к сжатию аудиокодеком AAC с битрейтом 256 кбит/с, при котором достигнута точность восстановления данных в «тихие» блоки аудиосигнала до 95–99% (в зависимости от аудиоконтента). Показана высокая

эффективность применения адаптивного порога шума для надежного детектирования пиков эха в кепстре после сжатия.

Полученные результаты свидетельствуют о перспективности использования данного метода для задач скрытой передачи данных или встраивания цифровых водяных знаков в аудиопотоки, подверженные сжатию с потерями. Для обеспечения достоверного восстановления информации необходимо применение кодов коррекции ошибок. Дальнейшие исследования могут быть направлены на разработку методов синхронизации для повышения скрытности путем встраивания только в «тихие» сегменты, на дальнейшее улучшение робастности эхо-детектора, а также на количественную оценку скрытности предложенного метода и его сравнение с другими устойчивыми алгоритмами аудиостеганографии.

Список литературы

1. Урбанович П. П. Защита информации методами криптографии, стеганографии и обфускации. Минск: БГТУ, 2016. 116 с.

2. Шутько Н. П. Защита авторских прав на электронные текстовые документы методами стеганографии // Труды БГТУ. Сер. 3, Физико-математические науки и информатика. 2013. № 6 (162). С. 131–134.

3. Fridrich J. Steganography in Digital Media: Principles, Algorithms, and Applications. Cambridge: Cambridge University Press, 2010. 438 p.

4. Бранденбург К. Кодирование звука с высоким разрешением // Звукорежиссер. 2001. № 7. С. 99–105.

5. Westfeld A., Pfitzmann A. Attacks on Steganographic Systems // Information Hiding: Third International Workshop, IH'99, Dresden, Germany, September 29 – October 1, 1999. Proceedings / ed. A. Pfitzmann. Berlin; Heidelberg: Springer, 2000. P. 61–76 (Lecture Notes in Computer Science; vol. 1768).

6. Techniques for data hiding / W. Bender [et al.] // IBM Systems Journal. 1996. Vol. 35, no. 3&4. P. 313–336.

7. Алдошина И. А. Основы психоакустики. М.: [б. и.], 2000. 248 с.

8. Оппенгейм А. В., Шафер Р. В. Цифровая обработка сигналов; пер. с англ.; под ред. С. Я. Шаца. М.: Связь, 1979. 416 с.

9. National Institute of Standards and Technology. Advanced Encryption Standard (AES). Gaithersburg, MD: U.S. Department of Commerce, 2001. 47 p. (Federal Information Processing Standards Publication (FIPS PUB); 197). URL: https://nvlpubs.nist.gov/nistpubs/FIPS/NIST.FIPS.197.pdf (дата обращения: 23.04.2025).

10. Dworkin M. Recommendation for Block Cipher Modes of Operation: Galois/Counter Mode (GCM) and GMAC. Gaithersburg, MD: National Institute of Standards and Technology, 2007. 39 p. (NIST Special Publication; 800-38D). URL: https://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-38d.pdf (дата обращения: 23.04.2025).

11. Морелос-Сарагоса Р. Искусство помехоустойчивого кодирования. Методы, алгоритмы, применение. М.: Техносфера, 2006. 320 с.

References

1. Urbanovich P. P. *Zashchita informatsii metodami kriptografii, steganografii i obfuskatsii* [Information security by methods of cryptography, steganography and obfuscation]. Minsk, BGTU Publ., 2016. 116 p. (In Russian).

2. Shutko N. P. Copyright protection of electronic text documents by steganography methods. *Trudy BGTU* [Proceedings of BSTU], issue 3, Physics and mathematics. Informatics, 2013, no. 6 (162), pp. 131–134. (In Russian).

3. Fridrich J. Steganography in Digital Media: Principles, Algorithms, and Applications. Cambridge, Cambridge University Press Publ., 2010. 438 p.

4. Brandenburg K. High-resolution audio coding. *Zvukorezhisser* [Sound Engineer], 2001, no. 7, pp. 99–105 (In Russian).

5. Westfeld A., Pfitzmann A. Attacks on Steganographic Systems. *Information Hiding: Third International Workshop, IH'99, Dresden, Germany, September 29 – October 1, 1999. Proceedings*. Ed. by A. Pfitzmann. Berlin; Heidelberg, Springer Publ., 2000, pp. 61–76 (Lecture Notes in Computer Science; vol. 1768).

6. Bender W., Gruhl D., Morimoto N., Lu A. Techniques for data hiding. *IBM Systems Journal*, 1996, vol. 35, no. 3&4, pp. 313–336.

7. Aldoshina I. A. *Osnovy psikhoakustiki* [Fundamentals of psychoancoustics]. Moscow, 2000. 248 p. (In Russian).

8. Oppenheim A. V., Schafer R. W. *Tsifrovaya obrabotka signalov* [Digital Signal Processing]. Ed. by S. Ya. Shats. Moscow, Svyaz' Publ., 1979. 416 p. (In Russian).

9. National Institute of Standards and Technology (NIST). FIPS PUB 197: Advanced Encryption Standard (AES). November 26, 2001. 47 p. Available at: https://nvlpubs.nist.gov/nistpubs/FIPS/NIST.FIPS.197.pdf (accessed 23.04.2025).

10. Dworkin M. Recommendation for Block Cipher Modes of Operation: Galois/Counter Mode (GCM) and GMAC. NIST Special Publication 800-38D. November, 2007. 39 p. Available at: https://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-38d.pdf (accessed 23.04.2025).

11. Morelos-Zaragoza R. *Iskusstvo pomekhoustoychivogo kodirovaniya* [The art of error correcting coding. Methods, algorithms, application]. Moscow, Tekhnosfera Publ., 2006. 320 p. (In Russian).

Информация об авторах

Попеня Наталья Владимировна – аспирант кафедры информатики и веб-дизайна. Белорусский государственный технологический университет (ул. Свердлова, 13а, 220006, г. Минск, Республика Беларусь). E-mail: popenya@belstu.by

Романенко Дмитрий Михайлович – кандидат технических наук, доцент, заведующий кафедрой информатики и веб-дизайна. Белорусский государственный технологический университет (ул. Свердлова, 13а, 220006, г. Минск, Республика Беларусь). E-mail: rdm@belstu.by

Information about the authors

Popenya Natalya Vladimirovna – PhD student, the Department of Computer Science and Web-design. Belarusian State Technological University (13a Sverdlova str., 220006, Minsk, Republic of Belarus). E-mail: popenya@belstu.by

Romanenko Dmitri Mikhailovich – PhD (Engineering), Associate Professor, Head of the Department of Computer Science and Web-design. Belarusian State Technological University (13a Sverdlova str., 220006, Minsk, Republic of Belarus). E-mail: rdm@belstu.by

Поступила 12.05.2025