

УДК 004.4'277+78.06

## ИССЛЕДОВАНИЕ ВЛИЯНИЯ АУДИОКОДЕКОВ НА КАЧЕСТВО ЗВУЧАНИЯ СТРИМИНГОВЫХ ПЛАТФОРМ

А.Д. ВАСЮТИЧ<sup>1,2</sup>, Н.С. ДАВЫДОВА<sup>1</sup>

1 – Белорусский государственный университет информатики и радиоэлектроники, Республика Беларусь,

2 – ОАО «Пеленг», Республика Беларусь

*Поступила в редакцию 3 апреля 2025*

**Аннотация.** В работе представлены оптимальные решения по выбору аудиокодексов для стриминговых платформ, обеспечивающие баланс между качеством звучания и размером файла. Экспериментально подтверждена зависимость характеристик аудиосигнала от типа используемого кодека.

**Ключевые слова:** кодек аудио без потерь, кодек аудио с потерями, частота дискретизации, битовая разрядность, стриминговые сервисы, АЧХ.

### Введение

Рост популярности медиаконтента в современном мире привел к развитию стриминговых платформ, которые позволяют пользователям получать доступ к разнообразным медиаресурсам посредством мультисервисных сетей. Отличительными особенностями стриминговых платформ являются потоковое вещание, интерактивность и адаптивность, что делает их удобными и востребованными сетевыми сервисами [1].

Качество и производительность потокового медиа зависит от параметров мультисервисной сети и, в первую очередь, от средней пропускной способности. Для уменьшения последствий задержек и изменений пропускной способности сети применяют различные методы, такие как буферизация данных на клиентской стороне, предварительная загрузка данных и адаптивное изменение качества медиа в зависимости от пропускных возможностей канала [2].

Наиболее эффективной технологией является DASH. Dynamic Adaptive Streaming over HTTP – технология адаптивной потоковой передачи данных по HTTP, которая позволяет клиентам с разным качеством сетевого соединения потреблять медиаконтент с разной шириной потока. То есть DASH динамически подстраивает качество медиа в реальном времени в зависимости от условий сети и возможностей устройства пользователя. Это возможно благодаря тому, что все медиафайлы на стриминговой платформе имеют несколько версий с разной шириной потока (битрейта) и, соответственно, с разным качеством [3].

Получение нескольких версий одного и того же медиафайла с несколькими уровнями качества возможно с помощью алгоритмов сжатия посредством аудио- и видеокодексов. При этом цифровые аудиоданные имеют более низкие требования к пропускной способности сети, чем видеоданные. И несмотря на то, что по сравнению с видеоданными звуковые файлы имеют значительно меньший размер, пользователи в целом более чувствительны именно к аудиопомехам.

Целью представленной работы является исследование влияния аудиокодексов на качество звучания стриминговых платформ.

### Методология

Кодеки аудио – это алгоритмы сжатия аудиоданных. Их можно разделить на два основных типа: с потерями (loosy) и без потерь (lossless). Форматы с потерями вырезают из исходного трека

звуковые данные, которые по мнению алгоритмов человеческого уха не заметит. При этом размер итогового файла получается относительно небольшим, однако теряется некоторая частотная информация, особенно на высоких частотах, что в свою очередь влияет на детальность звучания. Среди наиболее часто используемых кодеков аудио с потерями можно выделить такие, как MP3, AAC, OGG [4]. В настоящее время одним из самых популярных среди них является MP3, однако его постепенно заменяет AAC. В случае, если качество звучания стоит на первом месте, выбор отдается форматам без потерь. Они идеальны для прослушивания музыки на высококачественной аудиотехнике и хорошо подходят для хранения аудиофайлов студийного качества. В качестве примера можно привести такие кодеки аудио без потерь, как WAV, AIFF, FLAC. При экспорте музыкального трека из проекта в кодек аудио без потерь, например, FLAC, можно также выбирать и степень сжатия. Она будет напрямую влиять на качество звучания полученного файла. Чем меньше степень сжатия, тем более качественное звучание будет получено. На рис. 1 и рис. 2 представлены графики АЧХ одного и того же музыкального трека с использованием разных аудиокодеков.

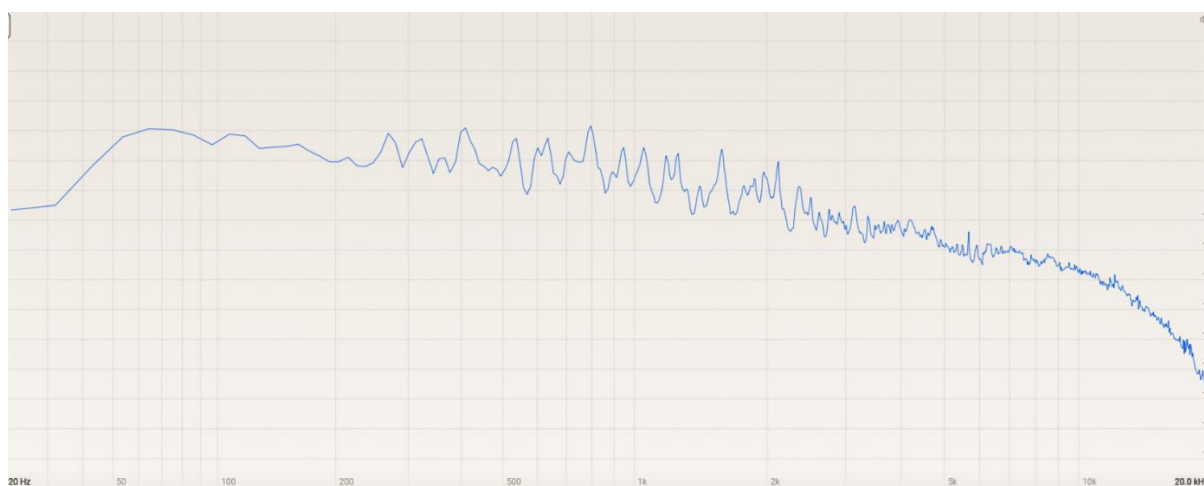


Рис. 2. График АЧХ музыкального трека с использованием кодека аудио без потерь (lossless) WAV  
44,1 кГц, 24 бит

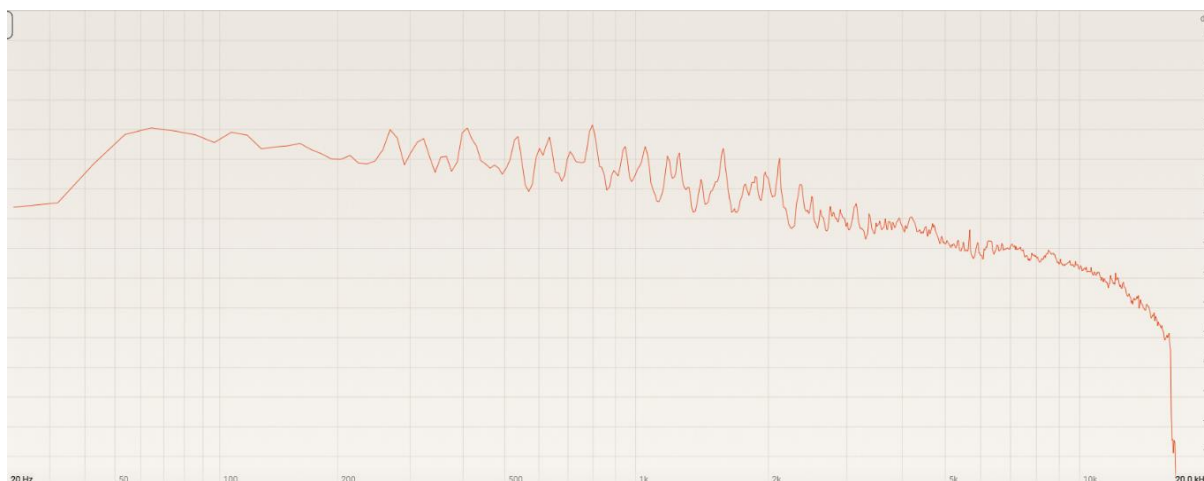


Рис. 2. График АЧХ музыкального трека с использованием кодека аудио с потерями (loosy) MP3  
44,1 кГц, 128 kbps

Из полученных графиков АЧХ видно, что при использовании кодека аудио без потерь частотный диапазон (20 Гц – 20 кГц), в том числе и на верхних частотах, сохраняется. Кодек аудио с потерями приводит к обрезанию частот, находящихся выше 16 кГц и вносит некоторые искажения в исходный сигнал. На рис. 3 представлен график АЧХ, показывающий различие в спектрах между кодеком аудио без потерь и с потерями (серым цветом). Наиболее явные отличия находятся в области высоких частот (около 16 – 20 кГц).

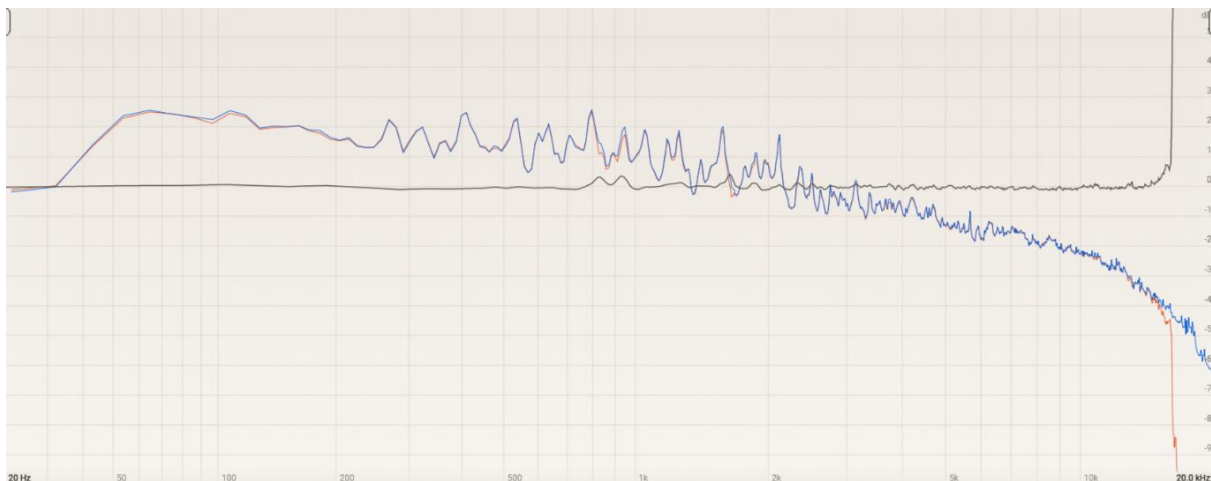


Рис. 3. Разность спектров WAV и MP3

### Факторы, влияющие на качество аудио

Частота дискретизации звукового сигнала определяет количество выборок данных в единицу времени, взятых из непрерывного сигнала для последующей оцифровки и хранения в памяти [5]. Человеческое ухо способно слышать звуки в частотном диапазоне примерно от 20 Гц до 20 кГц. Согласно теореме Котельникова-Найквиста, для точного воспроизведения звука частота дискретизации должна быть как минимум в два раза больше максимальной частоты в сигнале. Учитывая диапазон частот, которые может услышать человек, и данную теорему, получаем, что минимальная частота дискретизации должна быть 40 кГц. В настоящее время в студийных записях используются следующие частоты дискретизации: 44,1 кГц; 48 кГц; 88 кГц; 96 кГц; 192 кГц. Поэтому минимальной частотой дискретизации для полной передачи качества звучания трека является 44,1 кГц. Стандарт CD использует данную частоту дискретизации [6]. Более высокие значения частоты дискретизации в основном используются в профессиональных условиях при студийной записи, сведении и мастеринге. Это связано с тем, что они улучшают качество передачи быстрых переходных процессов в треке (транзиентной составляющей), таких как атака барабанов, щелчок медиатора по гитарной струне, нажатие на клавишу пианино. Также улучшается точность в слышимом диапазоне за счет уменьшения алиасинга (искажений).

Битовая разрядность (bit depth) влияет на динамический диапазон музыкального трека – разность между самым тихим и самым громким звуком в треке [5]. Стандарт CD использует битовую разрядность, равную 16 битам, что обеспечивает динамический диапазон равный 96 дБ [5]. Данного значения динамического диапазона достаточно для большинства аудиозаписей. Однако в студийных условиях, особенно при сведении, чаще всего используется значение, равное 24 или даже 32 бит. Данные значения снижают шум квантования аудиосигнала и позволяют работать с более тонкими нюансами, например, с тихими пассажами в классической музыке.

### Инструментарий и сбор данных

Выбор аудиокодеков для исследования обоснован целевой аудиторией стриминговых платформ. Стриминговыми платформами могут пользоваться как обычные пользователи с низкогокачественными устройствами воспроизведения аудио, так и продвинутые пользователи с высококачественной звуковой аппаратурой и развитым профессиональным слухом. Соответственно, от данного выбора будут зависеть и используемые в сервисе кодеки аудио. Стоит сразу отметить, что в большинстве случаев нет никакого смысла использовать высококачественные кодеки аудио без потерь, если исходный трек был экспортирован в формат с потерями, плохо сведен и нет возможности переэкспортировать его из оригинального проекта в формат без потерь. Также бессмысленно перекодировать аудиофайл с использованием кодека с частотой дискретизации или битовой разрядностью большей, чем в нем сейчас используется, так как невозможно добавить в него то, чего уже нет. Соответственно, при такой

переконвертации мы получим такое же качество звучания, которое и было, но вес файла увеличится, что будет означать нежелательное использование памяти.

Если пользователи сервиса слушают треки в наушниках среднего ценового класса, через встроенные динамики смартфона или ноутбука, через Bluetooth-колонку, то нет смысла использовать кодеки аудио без потерь, так как разница будет неощутима, однако сам звуковой файл будет занимать намного больше места.

Если же пользователи будут использовать высококачественные наушники, динамики, качественные цифро-аналоговые преобразователи и обладать натренированным профессиональным слухом, то форматы без потерь раскроют весь потенциал аудиозаписей. С высокими значениями частоты дискретизации и битовой разрядности, хорошо сведенными исходниками профессионально созданный трек в формате без потерь сохранит даже такие нюансы, как дыхание вокалиста и реверберацию комнаты, в которой он записывался.

Для исследования использовались профессиональный аудиоинтерфейс с максимальной частотой дискретизации 96 кГц и битовой разрядностью 24 бит, профессиональные студийные наушники с сопротивлением 250 Ом, чувствительностью 102 дБ/мВт и частотным диапазоном 5 Гц – 40 кГц, проект в цифровой рабочей станции с профессионально сведенным треком.

Целью работы было определение оптимального кодека аудио с минимальным размером файла, но приемлемым качеством звучания. В ходе эксперимента слушателем с развитым профессиональным слухом оценивалось качество звучания трека с использованием разных аудиокодеков. Для оценки использовались как количественные, так и качественные критерии оценки качества звука.

Количественные критерии:

1. Сохранность АЧХ оценивалась через сравнение спектров исходного и сжатого аудио с помощью программного спектрального анализатора. Из спектра исходного файла вычитался спектр сжатого аудио. Результат вычитания – разностный спектр, который визуализирует отклонения, вызванные сжатием. На графике разностного спектра фиксировалось максимальное значение отклонения ( $\Delta$ , дБ) в пределах всего частотного диапазона (20 Гц – 20 кГц). Например, если на частоте 10 кГц разница между исходным и сжатым сигналом составила 2,5 дБ, а на 15 кГц 4 дБ, максимальной дельтой было 4 дБ. Чем меньше максимальная дельта, тем выше сохранность АЧХ. На основе этого значения применялась 5-балльная шкала, которая представлена в табл. 1.

Табл. 1. Шкала оценки сохранности АЧХ

| Оценка | Максимальная дельта (разность спектров) | Интерпретация                                                         |
|--------|-----------------------------------------|-----------------------------------------------------------------------|
| 5      | $\leq 0.5$ дБ                           | Идеальная сохранность АЧХ. Отличия от оригинала отсутствуют.          |
| 4      | 0.6 – 1 дБ                              | Минимальные искажения. Потери заметны только при детальном анализе.   |
| 3      | 1.1 – 3 дБ                              | Умеренные искажения. Приемлемо для непрофессионального использования. |
| 2      | 3.1 – 6 дБ                              | Значительные потери. Ухудшение качества, особенно в высоких частотах. |
| 1      | $> 6$ дБ                                | Неприемлемые искажения. Частотный диапазон обрезан или искажен.       |

2. Размер файла соответствовал размеру полученного выходного файла при экспорте музыкального трека из проекта цифровой рабочей станции.

Качественные критерии:

1. Проведено слепое тестирование с использованием цифровой рабочей станции, в ходе которого оценивался трек с разными аудиокодеками. Оценка слушателя по 10-балльной шкале:

- 1) 9, 10 – неотличимо от оригинала (артефакты отсутствуют).
- 2) 6–8 – приемлемо (минимальные искажения).
- 3)  $\leq 5$  – неприемлемо (заметные артефакты, шумы).

## Результаты и их обсуждение

Результаты исследования представлены в табл. 2 и 3. Условием работы был поиск кодека аудио с приемлемым качеством звучания, но минимальным размером, поэтому использовалась частота дискретизации 44,1 кГц. Итоговые результаты были разделены на две таблицы. В табл. 2 представлены результаты эксперимента с использованием кодеков аудио без потерь.

Табл. 2. Сравнение кодеков аудио без потерь

| Кодек | Частота дискретизации | Битовая разрядность | Размер файла | Оценка сохранности АЧХ | Оценка звучания слушателем |
|-------|-----------------------|---------------------|--------------|------------------------|----------------------------|
| WAV   | 44,1 кГц              | 64 бит Float        | 127 МБ       | 5                      | 10                         |
| WAV   | 44,1 кГц              | 32 бит Float        | 63,5 МБ      | 5                      | 10                         |
| WAV   | 44,1 кГц              | 24 бит              | 47,6 МБ      | 5                      | 10                         |
| WAV   | 44,1 кГц              | 16 бит              | 31,8 МБ      | 4                      | 9                          |
| WAV   | 44,1 кГц              | 8 бит               | 15,9 МБ      | 2                      | 6                          |
| AIFF  | 44,1 кГц              | 32 бит Float        | 63,5 МБ      | 5                      | 10                         |
| AIFF  | 44,1 кГц              | 24 бит              | 47,6 МБ      | 5                      | 10                         |
| AIFF  | 44,1 кГц              | 16 бит              | 31,8 МБ      | 4                      | 9                          |
| AIFF  | 44,1 кГц              | 8 бит               | 15,9 МБ      | 2                      | 6                          |
| FLAC  | 44,1 кГц              | 32 бит              | 50,3 МБ      | 5                      | 10                         |
| FLAC  | 44,1 кГц              | 24 бит              | 34,5 МБ      | 4                      | 10                         |
| FLAC  | 44,1 кГц              | 16 бит              | 18,9 МБ      | 4                      | 8                          |
| ALAC  | 44,1 кГц              | 24 бит              | 35,2 МБ      | 4                      | 10                         |
| ALAC  | 44,1 кГц              | 16 бит              | 19,3 МБ      | 4                      | 9                          |

В табл. 3 представлены результаты эксперимента с использованием кодеков аудио с потерями.

Табл. 3. Сравнение кодеков аудио с потерями

| Кодек | Частота дискретизации | Битовая разрядность | Размер файла | Оценка сохранности АЧХ | Оценка звучания слушателем |
|-------|-----------------------|---------------------|--------------|------------------------|----------------------------|
| MP3   | 44,1 кГц              | 320 kbps            | 7,2 МБ       | 4                      | 7                          |
| MP3   | 44,1 кГц              | 256 kbps            | 5,8 МБ       | 1                      | 6                          |
| MP3   | 44,1 кГц              | 192 kbps            | 4,3 МБ       | 1                      | 5                          |
| MP3   | 44,1 кГц              | 128 kbps            | 2,9 МБ       | 1                      | 3                          |
| MP3   | 44,1 кГц              | 64 kbps             | 1,4 МБ       | 1                      | 1                          |
| AAC   | 44,1 кГц              | 256 kbps            | 5,7 МБ       | 4                      | 7                          |
| AAC   | 44,1 кГц              | 224 kbps            | 5 МБ         | 4                      | 6                          |
| AAC   | 44,1 кГц              | 192 kbps            | 4,3 МБ       | 1                      | 5                          |
| AAC   | 44,1 кГц              | 160 kbps            | 3,6 МБ       | 1                      | 3                          |
| AAC   | 44,1 кГц              | 128 kbps            | 2,9 МБ       | 1                      | 3                          |
| AAC   | 44,1 кГц              | 96 kbps             | 2,2 МБ       | 1                      | 2                          |
| OGG   | 44,1 кГц              | Variable            | 5,4 МБ       | 4                      | 5                          |

Анализ полученных данных показывает, что среди кодеков аудио без потерь оптимальным вариантом с приемлемым качеством, но минимальным размером файла, будет кодек FLAC или ALAC с частотой дискретизации 44,1 кГц и битовой разрядностью 16 бит. Наилучшим по качеству вариантом будет использование кодека WAV или AIFF. Однако, как уже было упомянуто ранее, чем больше частота дискретизации и битовая разрядность, меньше степень сжатия, тем более качественное звучание будет получено на профессиональных устройствах воспроизведения. При этом следует отметить, что треки должны быть изначально отличного качества и экспортированы из проекта в формат без потерь. Если же сервис ориентирован на обычных пользователей, не использующих высококачественные устройства воспроизведения, то оптимальным вариантом будет кодек MP3 или AAC с частотой дискретизации 44,1 кГц и битрейтом 256 или 320 кбит/с. В современных сервисах можно использовать и оба варианта

кодеков. В таком случае для продвинутых пользователей можно предложить формат без потерь, а для всех остальных – формат с потерями. Данный вариант разделения используется в некоторых стриминговых сервисах: платная подписка позволяет слушать музыкальные треки в формате без потерь.

### Заключение

Результаты проведенных исследований показали, что для аудио без потерь оптимальным выбором являются кодеки FLAC и ALAC (16 бит; 44,1 кГц), обеспечивающие малый размер файла (18,9 МБ; 19,3 МБ) при оценке сохранности АЧХ в 4 балла и оценке качества звучания в 8, 9 баллов. Для пользователей с высокими требованиями к качеству звучания рекомендованы WAV/AIFF (24, 32 бит), несмотря на увеличенный размер (47,6 МБ; 63,5 МБ). Среди кодеков с потерями лучшие результаты продемонстрировали MP3 (320 kbps) и AAC (256 kbps) с оценкой сохранности АЧХ в 4 балла и размером 7,2 МБ; 5,7 МБ.

Полученные данные могут быть использованы стриминговыми сервисами для гибкой настройки аудиопотоков: предоставление lossless-форматов (FLAC/ALAC) пользователям с профессиональным оборудованием и lossy-форматов (MP3/AAC) – основной аудитории. Это позволит сократить затраты на хранение и передачу данных без значимого ухудшения качества звучания в целевых сегментах. Внедрение многоуровневой системы (например, платной подписки для lossless форматов) повысит удовлетворенность пользователей и оптимизирует ресурсы платформы.

## STUDY OF THE IMPACT OF AUDIO CODECS ON THE SOUND QUALITY OF STREAMING PLATFORMS

A.D. VASIUTICH, N.S. DAVYDOVA

**Abstract.** The paper presents optimal solutions for choosing audio codecs for streaming platforms, providing a balance between sound quality and file size. The dependence of audio signal characteristics on the type of codec used is experimentally confirmed.

**Keywords:** lossless audio codec, lossy audio codec, sampling rate, bit rate, streaming services, AFC.

### Список литературы

1. Arditi D. Streaming culture: Subscription platforms and the unending consumption of culture. – Emerald Publishing Limited, 2021.
2. Куроуз Д., Росс К. Компьютерные сети. Нисходящий подход. – Litres, 2022.
3. Michalos M. G., Kessanidis S. P., Nalmpantis S. L. Dynamic Adaptive Streaming over HTTP // Journal of Engineering Science & Technology Review. – 2012. – Т. 5. – №. 2.
4. Bosi M., Goldberg R.E. Introduction to Digital Audio Coding and Standards. Springer, 2003.
5. Райков В. Цифровая обработка сигналов: теория и практика. М.: Энергия, 2005.
6. Watkinson J. The Art of Digital Audio. Focal Press, 2013.