

**Ковтуненко Аркадий Алексеевич**, аспирант, Петербургский  
государственный университет путей сообщения Императора Александра I, г.

Санкт-Петербург

## **ВНЕДРЕНИЕ МАРКЕРОВ В АУДИОСИГНАЛЫ С АДАПТИВНОЙ СИЛОЙ: ПРЕИМУЩЕСТВА И ПОДХОДЫ**

**Аннотация.** В большинстве как традиционных, так и современных методов внедрения цифровых водяных знаков в аудиосигналы используется статическая сила внедрения сигнала. Этот подход достаточно ограничен и не использует ресурсы сигнала максимально эффективно, теряя в устойчивости к искажениям или незаметности водяного знака. Для решения этих проблем разрабатываются подходы с динамической силой внедрения сигнала. Данная статья представляет из себя обзор на основные современные подходы к адаптации силы внедрения водяных знаков в аудиосигналы с целью повышения осведомлённости в этой области. Кратко описываются сильные и слабые стороны методов, области их применения и перспективы развития в будущем.

**Abstract.** In most traditional and modern methods of embedding digital watermarks into audio signals, a static embedding strength is used. This approach is inherently limited and fails to utilize the signal's capacity efficiently, resulting in reduced robustness to distortions or decreased imperceptibility of the watermark. To address these shortcomings, methods with dynamically adjustable embedding strength are being developed. This article provides an overview of current approaches to adapting watermark embedding strength in audio signals, aiming to improve awareness in this field. The paper briefly outlines the strengths and weaknesses of different methods, their areas of application, and future development prospects.

**Ключевые слова:** аудио, стеганография, водяные знаки, сила внедрения, эвристические методы, психоакустические модели, глубокое обучение.

**Keywords:** audio, steganography, watermarking, embedding strength, rule-based methods, psychoacoustic models, deep learning.

## **Введение**

В условиях стремительного роста объёмов цифрового аудиоконтента и развития технологий фальсификации и генерации аудио [1] возрастает необходимость в эффективных средствах защиты. Одним из самых широко применяемых методов в этой области является использование цифровых водяных знаков. Они представляют из себя незаметно внедрённую в цифровой объект информацию об этом объекте [2], служащую для идентификации, отслеживания использования или технических задач.

Однако традиционные методы внедрения водяных знаков в аудиосигналы используют статическую силу внедрения сигнала. Такой подход, пытаясь балансировать между устойчивостью и незаметностью водяного знака, не учитывает акустические особенности различных фрагментов сигнала. Для исправления этого недостатка были созданы и по сей день развиваются подходы динамической настройки параметров внедрения водяного знака в сигнал.

Целью данной статьи является обзор существующих подходов к внедрению цифровых водяных знаков в аудиосигналы с динамической регулировкой силы внедрения, обозначение их преимуществ, недостатков и перспектив развития.

### **Преимущества динамической силы внедрения**

Сила внедрения – это коэффициент, определяющий степень воздействия внедряемого цифрового водяного знака на аудиосигнал. Сила внедрения напрямую влияет на устойчивость водяного знака к искажениям и его незаметность в оригинальном аудиосигнале. Таким образом, высокая сила внедрения повышает устойчивость водяного знака к атакам и искажениям, но делает его более заметным на слух в оригинальном сигнале, а низкая сила внедрения делает водяной знак практически неслышимым, но легко разрушаемым при обработке и искажении аудиосигнала.

Во многих методах сила внедрения является константой, одинаковой для всего аудиофайла. Такой подход может не обеспечивать достаточную устойчивость при слишком слабом внедрении и, напротив, вызывать слышимые искажения, если сила внедрения превышает пороги слышимости. Во многом это происходит из-за того, что выбор фиксированного значения не учитывает особенности восприятия звука человеком.

Динамическая же сила внедрения регулируется для каждого сегмента аудиосигнала индивидуально в зависимости от его характеристик. Это оптимизирует внедрение водяного знака на уровне отдельного аудио отсчёта, обеспечивая большую незаметность на «тихих» участках и большую устойчивость на «шумных» участках. Таким образом, данный подход подбирает оптимальное соотношение незаметности и устойчивости цифрового водяного знака [3], что особенно полезно при потоковой передаче и массовой обработке аудио в реальном времени.

### **Подходы адаптации силы внедрения**

#### **Эвристические методы.**

Эвристические (rule-based) методы являются одними из самых простейших и основаны на заранее заданных правилах и пороговых значениях. Принцип этих методов заключается в том, что водяной знак внедряется в конкретный аудио отсчёт с определённой силой, если акустические или статистические признаки на этом аудио отсчёте соответствуют определённым условиям [4]. Например, при высокой дисперсии у сигнала более сложная структура, поэтому можно внедрять водяной знак с большей силой.

Таким образом, эвристические методы просты в реализации и имеют низкие вычислительные затраты, а также подходят для работы в реальном времени. Однако этот подход очень неустойчив к атакам, а из-за того, что анализ файла фрагментарный, методы не учитывают глобальный контекст аудиосигнала и трудно масштабируются на большие наборы данных.

## **Психоакустические модели.**

Психоакустические модели – это математические или алгоритмические модели слуховой системы человека. С их помощью можно определять силу внедрения водяного знака так, чтобы он не воспринимался человеческим ухом на фоне оригинального аудиосигнала. Пример такой модели приведён в [5].

Использование психоакустических моделей для динамического внедрения водяного знака гарантирует незаметность, но при этом требует множества сложных вычислений.

## **Статистические методы**

Статистические подходы определяются тем, что параметры внедрения водяного знака в них регулируются на основе исключительно статистических свойств аудиосигнала. В отличие от эвристических методов, где решения принимаются по фиксированным правилам, в статистических за основу берутся измерения распределений, отклонений и т.п. [6]. Например, при эвристическом подходе водяной знак внедряется с фиксированной силой для случая высокой дисперсии сигнала, если она превышает определённый порог; а при статистическом подходе сила внедрения масштабируется пропорционально значению дисперсии [7].

Ввиду своих свойств статистические методы универсальны, требуют меньше ручной настройки и небольших вычислительных мощностей, но, из-за опоры исключительно на статистические свойства аудиосигнала, игнорируют перцептивные свойства слуха и могут вызывать слышимые искажения.

## **Глубокое обучение**

Методы глубокого обучения используют нейронные сети для определения силы и места внедрения водяного знака и непосредственно для его внедрения и извлечения. И хоть существуют разные виды нейронных сетей, суть их работы одна и та же: обучить на базе примеров сеть определять в какие сегменты аудиосигнала и с какой силой нужно внедрять водяной знак максимально

энергоэффективно, чтобы он не искажал оригинальный сигнал и был устойчив к атакам и искажениям [8]. Особенность такого подхода заключается в обучении на примерах, а значит не требуется ручная настройка маскирующих порогов или статистических вычислений.

Это один из самых прогрессивных подходов в данное время и крайне адаптивный, однако он требует высоких вычислительных затрат как на обучение, так и на эксплуатацию модели, а также не обосновывает свои решения в отличие от более простых по структуре методов.

### **Оценка подходов в современных системах**

В наше время чисто эвристические методы не популярны для сложных задач, поскольку достаточно примитивны, однако благодаря своей ресурсоёмкости они хороши для простых задач в условиях ограниченных вычислительных ресурсов, как, например, в микроконтроллерах или для экспертизы аудиосигналов, поскольку легко интерпретируемы и подчиняются понятным правилам.

Психоакустические методы остаются полезными в сферах, где требуется сохранить качество звука аудио, например, в музыкальной сфере для защиты авторских прав. Статистические методы одни из самых актуальных, поскольку являются золотой серединой между устойчивостью и незаметностью внедрённого водяного знака, они предсказуемы и адаптивны.

Однако за последние несколько лет, с развитием нейронных сетей, развиваются методы глубокого обучения. Они уже в некоторых случаях превосходят традиционные подходы по точности и устойчивости, но из-за высоких вычислительных требований и отсутствия стандартизации ещё не стали самыми популярными. Особенно перспективны разработки гибридных методов, объединяющие нейронные сети с психоакустическими или статистическими подходами на этапе обучения модели [9].

### **Заключение**

В результате краткого обзора на подходы к динамическому изменению силы внедрения цифрового водяного знака в аудиосигнал, их преимущества и недостатки были сделаны выводы об актуальности этих подходов. В частности, наиболее перспективным направлением развития являются методы глубокого обучения на основе нейронных сетей, в то время как самыми распространёнными всё ещё являются статистические методы. Однако именно за объединением классических подходов, вроде психоакустических моделей, с методами глубокого обучения находится будущее этой сферы.

### Литература

1. ElevenLabs представила третью версию модели для генерации речи — с поддержкой эмоций и 70 языков // ROZETKED [Электронный ресурс] / URL: <https://rozetked.me/news/39669-elevenlabs-predstavila-tret-yu-versiyu-modeli-dlya-generacii-rechi-s-podderzhkoy-emociy-i-70-yazykov> (дата обращения: 07.06.2025).
2. I. Cox, M. L. Miller and J. A. Bloom, Digital Watermarking, Morgan Kaufmann Publishers, San Francisco, 2002.
3. Bagheri, Mahnoosh & Mohrekesh, Majid & Karimi, Nader & Samavi, Shadrokh. (2020). Adaptive Control of Embedding Strength in Image Watermarking using Neural Networks.
4. Hartung, Frank & Kutter, Martin. (1999). Multimedia watermarking techniques. Proceedings of the IEEE. 87. 1079 - 1107. 10.1109/5.771066.
5. ГОСТ Р 54712-2011 «Звуковое вещание цифровое. Кодирование сигналов звукового вещания с сокращением избыточности для передачи по цифровым каналам связи. MPEG-2 часть III (MPEG-2 audio)» – Москва: СТАНДАРТИНФОРМ. – 2014.
6. Chen, Shou & Huang, Huang-Nan & Tu, S.. (2011). Quantization Audio Watermarking with Optimal Scaling on Wavelet. Coefficients. Computing Research Repository - CORR.
7. Zhang, Jinqun & Han, Bin. (2017). Robust Audio Watermarking Algorithm Based on Moving Average and DCT. 10.48550/arXiv.1704.02755.

8. Purwins, Hendrik & Li, Bo & Virtanen, Tuomas & Schlüter, Jan & Chang, Shuo-yiin & Sainath, Tara. (2019). Deep Learning for Audio Signal Processing. IEEE Journal on Selected Topics in Signal Processing. 13. 10.1109/JSTSP.2019.2908700.
9. Singh, Mayank & Takahashi, Naoya & Liao, Weihsiang & Mitsufuji, Yuki. (2024). SilentCipher: Deep Audio Watermarking. 2235-2239. 10.21437/Interspeech.2024-174.