

*Санкт-Петербургский государственный электротехнический университет «ЛЭТИ»
им. В. И. Ульянова (Ленина), г. Санкт-Петербург, Россия*

Аннотация. Обсуждается проблема создания большой системы трансляции аудиопотоков. Рассматривается метод микширования аудио данных на стороне web-сервера, а также способы оптимизации аудиопотоков на стороне клиента. Предложен алгоритм микширования аудиопотоков.

Ключевые слова: дистанционное обучение; технология WebRTC; микширование аудиопотоков; peer-to-peer соединение

WebRTC (Web Real-Time Communications) – это технология, которая позволяет Web-приложениям и сайтам захватывать и выборочно передавать аудио и/или видео медиа-потоки, а также обмениваться произвольными данными между браузерами, без обязательного использования посредников. Технология WebRTC обеспечивает потоковую трансляцию данных одноранговых соединений используя соответствующий программный интерфейс приложения API (Application Programming Interface). Однако технология ограничена низким количеством узлов, поскольку не имеет в базовой реализации способа микширования данных. В случае peer-to-peer соединения максимальное число одноранговых пользователей не сможет превысить пяти, поскольку обработка данных видео и аудио потребляет большое количество системных ресурсов [1].

Общая схема технологии WebRTC приведена на рис. 1. Для потоковой передачи мультимедиа по сети WebRTC используется протокол RTP/SRTP (Secure Real-time Transport Protocol – протокол управления передачей в реальном времени и протокол интерактивного установления соединения ICE (Interactive Connectivity Establishment). Инфраструктура ICE в WebRTC поддерживает утилиты обхода сеанса, выделенную потоковую передачу мультимедиа в сети, преобразование сетевых адресов (NAT), утилиты обхода сеанса для NAT (STUN) и обход с использованием реле вокруг NAT (TURN). Когда установить прямое одноранговое соединение затруднительно из-за сетевых ограничений, вместо NAT(STUN) применяются альтернативы NAT (TURN) [2].

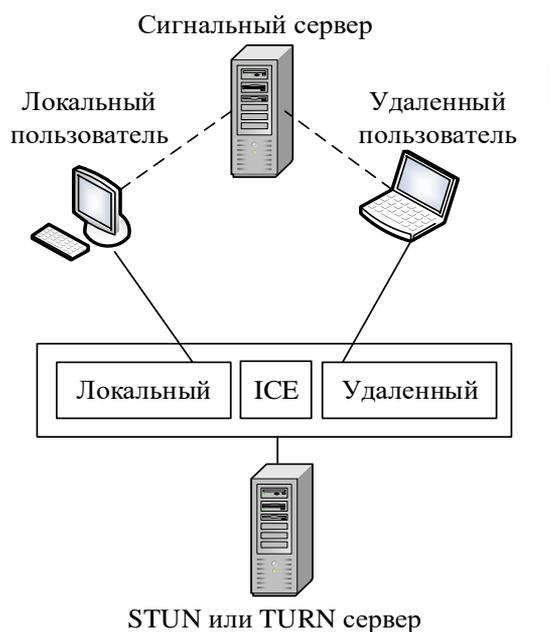


Рис. 1. Peer-to-peer соединение для двух пользователей

Предложен метод микширования аудио-потоков от нескольких одноранговых узлов с реализацией трансляции на большую аудиторию.

Методы реализации микширования аудио и видео можно разделить на две категории: методы на стороне клиента и методы на стороне сервера.

Проблема потоковой передачи на стороне клиента заключается в том, что требования к производительности клиента высоки, особенно при наличии нескольких источников ввода. Микширование на стороне сервера может обрабатывать множество распределенных полученных потоков, а затем передавать смешанный поток аудитории. Сторона сервера может централизовать процесс микширования и сократить ресурсы обработки на стороне клиента, а также может обеспечить унифицированный поток обработки данных и спецификации.

Микширование звука накладывает формы сигналов нескольких источников звука в соответствии с определенным алгоритмом и выводит их как один источник звука. Обычно звук разных форматов кодирования преобразуется в волны PCM, а затем смешивается. Целью микширования является получение наилучших результатов при многорожечной записи путем регулировки уровней, панорамирования и временных звуковых эффектов (хорус, реверберация, задержка) [3]. Также важно уменьшить шум и эхо во время микширования звука на онлайн-встречах.

Для решения данной проблемы предлагается изменить архитектуру системы. Основную роль в этом сыграет web-сервер, который будет являться единственным узлом для каждого пользователя. В момент инициализации приложения каждый пользователь устанавливает соединение с сервером, после обмена информацией о аудиорожках, сервер начинает передавать один микшированный аудиопоток данных, обрабатываемый в режиме реального времени.

Изменяя архитектуру таким образом, удастся уменьшить нагрузку на клиентов за счет увеличения нагрузки на web-сервер. Далее необходимо предложить способ оптимизации аудиорожек на сервере. Вместо того, чтобы сервер передавал (N-1) аудиорожек пользователю, где N – количество участников конференции, сервер может микшировать их на своей стороне и передавать одну суммарную аудиорожку.

После того, как web-сервер принял N аудиорожек от N пользователей, ему необходимо для каждого k-го пользователя передать результирующую дорожку. Данный процесс был разбивается на 4 этапа:

1. Суммирование всех полученных сэмплов на web-сервере согласно выражению:

$$\vec{M}_t = \sum_{k=0}^n \vec{U}_{k,t},$$

где $\vec{U}_{k,t}$ – вектор семпла, полученного от k-го пользователя за период времени t;

\vec{M}_t – суммарный вектор микширования – основной результат web-сервера.

2. Исключение или эхокомпенсация: происходит исключение k-го семпла из суммарного вектора. Таким образом k-му пользователю будет проходить аудиопоток без собственного аудио потока. Это можно выразить упрощённой формулой:

$$\vec{M}_{k,t} = \vec{M}_t - \vec{U}_{k,t},$$

где $\vec{M}_{k,t}$ – итоговой вектор микширования для k-го устройства.

3. Сжатие: для корректности итогового вектора $\vec{M}_{k,t}$ необходимо, чтобы его размерность совпадал с размерностью входящих векторов. Однако на 1 и 2 этапе исходный вектор может быть переполнен, что в конечном результате приведёт к дополнительным шумам. Сжатие происходит по формуле:

$$\vec{M}_{k,t} = \vec{M}_{k,t} * L_{k,t}$$

где $L_{k,t}$ – коэффициент сжатия для k-го семпла, который высчитывается автоматически, исходя из максимального сжатого семпла.

4. Генерация web-сервером для каждого пользователя по одной аудиорожке.

Таким образом, используя технологию WebRTC и микширование аудиопотоков на web-сервер решается проблема большого количества peer-to-peer соединений. Технология найдет применение в организации дистанционного образования [4], [5].

Список литературы:

1. Шитько А.М., Пацей Н.В. Использование протокола peer-to-peer для защищенного обмена данными // Труды БГТУ. 2015. № 6. С. 162–165.
2. Медведева Е.Г., Гайдамака Ю.В. К анализу параметров качества передачи мультиканального потокового трафика в одноранговой сети // Современные информационные технологии и ИТ-образование. 2015. Т. 11. № 2. С. 192-198.
3. Шестаков А.Л., Свиридюк Г.А., Худяков Ю.В. Динамические измерения в пространстве «шумов» // Вестник ЮУрГУ. 2013. Т. 13, №. 2. С. 4–11.
4. Tatarnikova T.M., Palkin I.I. Simulation as a high technology that contributes to the learning process at the university // CEUR Workshop Proceedings. 4. Сеп. "DLT 2019 – Selected Papers of the 4th All-Russian Scientific and Practical Conference with International Participation "Distance Learning Technologies"". 2021. С. 533-538.
5. Tatarnikova T.M., Miklush V.A., Kraeva E.V. Information technology for distance education // CEUR Workshop Proceedings. 4. Сеп. "DLT 2019 – Selected Papers of the 4th All-Russian Scientific and Practical Conference with International Participation "Distance Learning Technologies"". 2021. С. 539–545.

E. D. Arkhipsev, N. S. Mokretsov, T. M. Tatarnikova
Mixing audio streams using WebRTC technology in distance education

Saint Petersburg Electrotechnical University, Russia

***Abstract.** The problem of creating a large system for broadcasting audio streams is discussed. A method for mixing audio data on the web server side is considered, as well as ways to optimize audio streams on the client side. An algorithm for mixing audio streams is proposed.*

Keywords: distance learning; WebRTC technology; mixing of audio streams; peer-to-peer connection