

On the way to

Loudness nirvana

— audio levelling with EBU R 128

Внимание!

Данный перевод НЕ претендует на аутентичность и может содержать отдельные неточности. Оригинал документа на сайте <https://tech.ebu.ch>

На пути к нирване

Громкости

— регулирование уровня звука с EBU R 128

Florian Camerer
ORF

В статье описана одна из самых фундаментальных перемен в истории звука в вещании: смена парадигмы установки уровня с *нормализации пиков на нормализацию громкости*. Эта переменна жизненно необходима из-за проблемы, ставшей главным источником раздражения телезрителей и радиослушателей во всем мире – скачки уровней звука в перерывах внутри программ, между программами и между каналами. Нормализация громкости – решение по устранению этой проблемы.

Нормализация громкости – настоящая революция в регулировке уровня!

Рекомендация EBU R 128 [1] определяет прогнозируемый и строго определенный метод измерения уровня громкости¹ для новостей, спорта, рекламы, драмы, музыки, анонсов, фильмов и т.д. во всей цепи вещания и таким образом помогает профессионалам в создании робастных спецификаций для загрузки, производства, воспроизведения и распространения на множестве платформ. EBU создал четыре дополнительных документа, помогающих аудио индустрии в работе с Rec. R 128. Она полностью основана на открытых стандартах и нацелена на гармонизацию метода международного производства и измерения звука. Кроме **громкости программ**, R 128 вводит еще два аудио дескриптора, **диапазон громкости** и **максимальный уровень реальных пиковых значений**. Все три предназначены работать вместе, образуя набор необходимых дескрипторов, характеризующих аудио сигнал.

Измерение громкости и нормализация громкости – признаки *настоящей революции в регулировке уровня* (см. Рис. 1). Кроме того, эта парадигма регулировки громкости влияет на **все этапы** вещательного аудиосигнала, от сбора до распределения и передачи. Таким образом, конечная цель – гармонизация уровней громкости звука внутри вещательных каналов, а также между каналами для получения **одинакового универсального уровня громкости** в интересах слушателей. Поясним: уровень громкости может (и должен!) меняться в соответствии с художественными и техническими задачами внутри программы. Метод нормализации громкости использует **среднюю громкость** программы, чтобы сделать уровень согласованным между программами и между каналами.

¹ Громкость относится к воспринимаемой мощности звукового фрагмента (музыки, речи, звуковых эффектов и т.д.). Громкость зависит от уровня, частоты, контента и длительности звука, среди всего прочего.

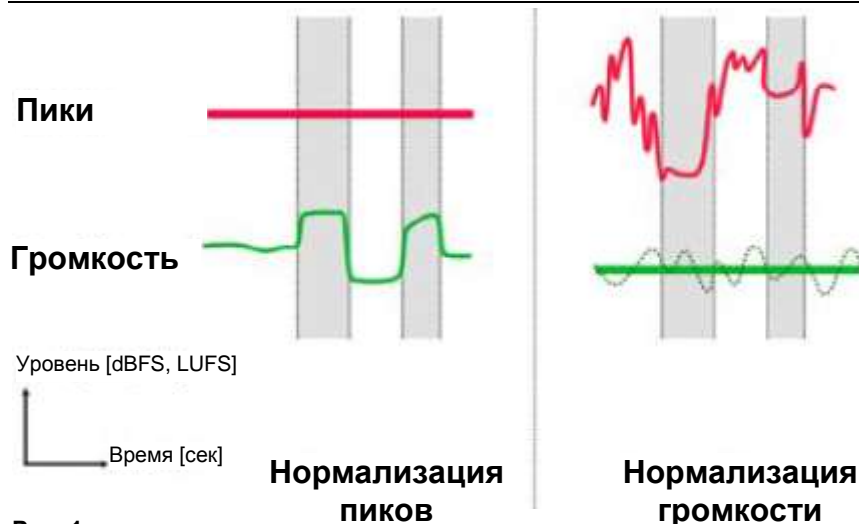


Рис. 1

Нормализация пикового уровня против нормализации уровня громкости программ

Опыт нескольких членов EBU показал, что работа с парадигмой громкости освобождает и удовлетворяет. Исчезает борьба «кто громче», сведение может быть более динамичным, меньше артефактов динамической компрессии, например, «качания», и, таким образом, наблюдается **общее повышение качества звука!** Производители программ, предпочитавшие раньше динамическое сведение, теперь избавлены от потенциальных компромиссов, т.к. их программа уже не звучит тише, чем более сжатые. С нормализацией громкости этого компромисса нет. Нирвана ближе чем всегда!

Источник проблемы – “война громкости”

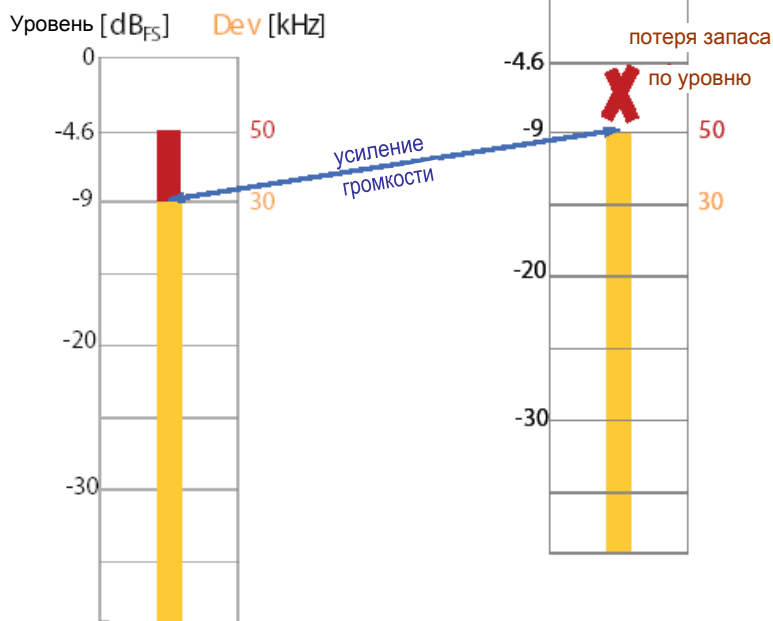


Рис. 2

Отношение между максимально допустимым уровнем (с QPPM) и отклонением FM и нарушение нормы: усиление громкости, но потеря запаса по уровню

Однако началось коммерческое давление, и чтобы не отставать от конкурентов, приходилось быть громче их. Современное измерение пиков и мощная обработка динамического диапазона позволили организациями переустановить PML, сравнившись с максимальным отклонением (50 kHz) FM несущей (см. Рис. 2). Все импульсные помехи должны были обрубаться на PML во избежание искажений, но те, кто это реализовал, считали это допустимым компромиссом.

Измерение звука в вещании сегодня обычно основано на квазипиковых вольтметрах (чаще QPPM, известных просто как PPM). «Квази» он называется из-за ограниченного времени реакции 10 мс (хотя бывает и 5 мс). На практике это означает, что пики сигнала короче этого времени реакции если и будут отображаться, то некорректно (например, импульсные помехи, созданные брэнчанием клавиш). Для обеспечения запаса по уровню для этих помех, которые не видны на приборе, но, тем не менее, должны быть и способствовать «открытости» аудио сигнала, согласованный **максимально допустимый уровень** (PML) был установлен на **-9 dBFS**.

Это значение основано на знакомом – и сохранившемся еще во многих местах – методе передачи звука слушателям в FM несущей. Максимальное отклонение несущей для ТВ во многих странах стандартизировано на 50 kHz и PML с отклонением 30 kHz (что равняется -9 dBFS), что таким образом позволяет **запас по уровню** 20 kHz, или 4.4 dB.

Когда слушатель переключался между одной из этих станций и станцией, не вступившей в «войну», или слышал громкую рекламу в сбалансированной программе с более широким динамическим диапазоном, уровень звука прыгал и приходилось хвататься за пульт и регулировать звук на более приемлемый уровень. В случае громкой рекламы звук потом надо было регулировать обратно при возвращении к к основной программе. Не удивительно, что вещатели получали так много жалоб. Другие решали проблему, заглушая звук во время рекламы, что как минимум сильно обедняло сообщение.

Появляется стандарт – и EBU его развивает

Международный союз электросвязи (ITU) признал проблему и выработал **ITU-R BS.1770** [2]. Цель этого стандарта – создание согласованного алгоритма для измерения громкости и уровней реальных пиковых значений программ. Это робастный стандарт, отличающийся простой реализацией. Вкратце, он определяет кривую “К-взвешивания” (модифицированный фильтр верхних частот второго порядка), который формирует основу для совпадения субъективного впечатления с объективным измерением.

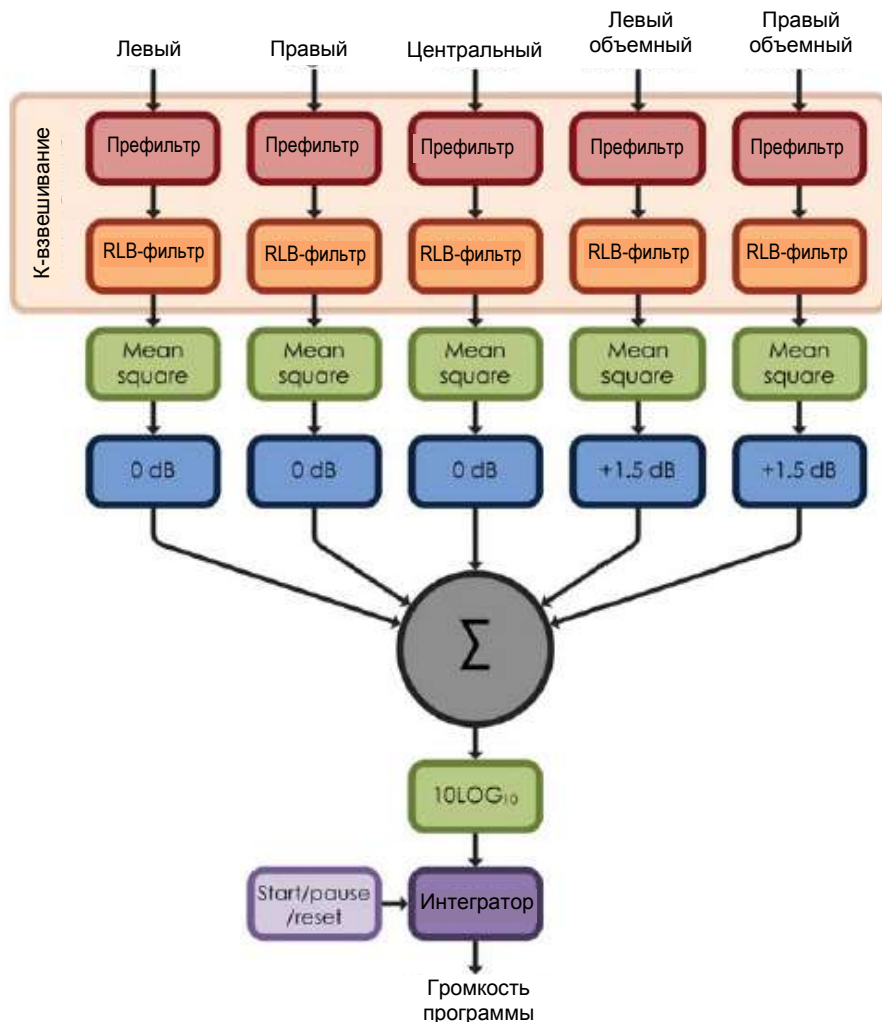


Рис. 3
Обработка канала и суммирование в ITU-R BS.1770

Эта кривая применяется ко всем каналам (кроме канала низкочастотных эффектов (LFE), который отбрасывается из этого измерения), затем вычисляется общая среднеквадратическая энергия (с коэффициентами дифференциального усиления для передних и объемных каналов, см. Рис. 3), и результат отображается как “LKFS” (громкость, K-взвешивание, относительно цифровой полной шкалы).

Для относительных измерений используются единицы громкости (LU), где 1 LU эквивалентно 1 dB.

Более детальное изучение алгоритма можно найти в ITUR BS.1770 [2], а также в EBU Tech Doc. 3343 – “Practical Guidelines” [3]).

ITU-R BS.1770 также определяет и рекомендует использовать измеритель реальных пиковых значений для измерения пиков. Такой измеритель работает с большей частотой дискретизации, чем аудио сигнал (обычно с передискретизацией 4x), чтобы улавливать пики между выборками, которые иначе могут превышать 0 dBFS и вызывать потом искажения в цепи.

R 128 и ITU-R BS.1770 – это основа. Детали содержатся в четырех технических документах EBU.

ITU-R BS.1770 обеспечивает основу для Рекомендации EBU R 128, которая расширяет стандарт ITU, фактически определяя специфический **целевой уровень** (см. ниже) для нормализации громкости, а также метод **стробирования** для улучшения совпадения громкости

программ, содержащих долгие периоды тишины или изолированные фрагменты активной речи. Разработки EBU нужны были для удовлетворения требований производителей программ, с особым вниманием к средствам для измерения общего сведения (а не просто одного компонента, например, диалога) и диапазона громкости программы. Для этого EBU определил три новых параметра:

- **Громкость программ**
- **Диапазон громкости**
- **Уровень реальных пиковых значений**

Громкость программ

Громкость программы описывает долговременную интегрированную громкость за все время программы. (В R 128 определение слова «программа» также используется для обозначения рекламы, заставок и других промежуточных материалов). Параметр состоит из одного числа (в LUFS²), с одним числом после десятичной точки), что означает «насколько громка программа в среднем». Это измеряется прибором, соответствующим ITU-R BS.1770 с добавлением функции стробирования. Стробирующий импульс служит для паузы в измерении громкости, когда сигнал падает ниже определенного порога. Без функции стробирования программы с более долгими периодами тишины или низкоуровневых фоновых звуков или шума получают слишком низкое значение для интегрированного уровня громкости. Потом при вещании такие программы будут слишком громкими.

После серии тестов был согласован стробирующий импульс –8 dB относительно измерения LUFS без стробирования с длиной блоков 400 мс. Тесты также подтвердили, наряду с другими результатами, выбор **целевого уровня**, по которому будет нормализован каждый аудио сигнал. Этот уровень:

–23 LUFS (–8 rel gate)

Отклонение **± 1 LU** допустимо для программ, где точная нормализация по целевому уровню –23 LUFS практически недостижима (например, прямые передачи или программы с крайне малым циклом). В случаях, когда уровни отдельных сигналов программы в большей степени **непредсказуемы** или когда программа состоит только из фоновых элементов (например, музыкальный фон прогноза погоды), допуск может быть слишком строгим. Поэтому в таких случаях ожидается, что интегрированный уровень громкости может выходить за пределы допуска, определенного R 128.

–23 LUFS – новый центр вселенной регулировки уровня.

Диапазон громкости

Loudness Range – общий дескриптор, помогающий решить необходимость динамической компрессии.

Еще один важный вопрос – диапазон громкости, необходимый для вмещения **всех** программ (при условии, что они не превышают допустимый диапазон громкости для домашнего прослушивания). Дескриптор **Loudness Range** (LRA) указывает (в LU) вариации измерения громкости программы. Он основан на статистическом распределении громкости внутри программы, исключая крайности. Таким образом, например, один выстрел не может повлиять на результат вычисления LRA. Рекомендация EBU R 128 не определяет максимально допустимый LRA, т.к. он зависит от таких факторов как окно допуска среднего слушателя станции, распределения жанров станции и т.д. R 128, однако, настоятельно **рекомендует использовать LRA** для определения необходимости динамической обработки аудио сигнала и для соответствия сигнала требованиям определенного канала или платформы передачи. Подробнее о LRA можно найти в EBU Tech Doc 3342 [4].

Он основан на статистическом распределении громкости внутри программы, исключая крайности. Таким образом, например, один выстрел не может повлиять на результат вычисления LRA. Рекомендация EBU R 128 не определяет максимально допустимый LRA, т.к. он зависит от таких факторов как окно допуска среднего слушателя станции, распределения жанров станции и т.д. R 128, однако, настоятельно **рекомендует использовать LRA** для определения необходимости динамической обработки аудио сигнала и для соответствия сигнала требованиям определенного канала или платформы передачи. Подробнее о LRA можно найти в EBU Tech Doc 3342 [4].

² **LUFS** показывает значение К-взвешенной громкости относительно цифровой полной шкалы. EBU рекомендует эту единицу для устранения расхождений между ITU-R BS.1770 и ITU-R BS.1771. Эта единица соответствует ISO 80000-8.

Первый опыт вещательных станций предполагает максимальное значение LRA около 20 LU для материала с большой динамикой, например, художественных фильмов или классической музыки. Большинству программ никогда не потребуется полное использование такого большого значения LRA или, на самом деле, они не смогут его достичь!

Уровень реальных пиковых значений

Уровень реальных пиковых значений аудио сигнала означает максимальное (положительное или отрицательное) значение формы сигнала в непрерывном временном интервале; в большинстве случаев это значение выше, чем показывает квазипиковый

вольтметр или даже пиковый вольтметр по выборкам, оба из которых пропускают реальные пиковые значения, находящиеся потенциально между выборками. Использование oversampling вольтметра, соответствующего BS.1770, позволяет обнаруживать эти пики.

Oversampling вольтметр все равно может чуть занижать показание реальных пиковых значений (в зависимости от частоты передискретизации), и **максимально допустимый уровень реальных пиковых значений** для производства будет:

-1 dBTP

Обратите внимание, что некоторые части цепи, например, аналоговые ретрансляторы и пользователи кодеров с низкой скоростью потока, требуют меньший уровень реальных пиковых значений. Полностью эта тема изложена в PLOUD Distribution Guidelines (EBU Tech Doc 3344 [5]).

Стратегии нормализации громкости



Рис. 4
Два принципиальных метода работы для получения однородной громкости в производстве и постпроизводстве

На Рис. 4 показано два подхода, в основном для **производства**. Первый больше подходит для первых этапов перехода и особенно полезен для тех, кто работает с прямым эфиром. Существующие измерители, лимитеры и практика сведения остаются, а сдвиг происходит на выходе пульта (после главных измерителей) для получения целевого уровня громкости **-23 LUFS**. Измеритель громкости расположен после сдвига уровня, позволяя инженерам понять точную величину сдвига (которая первоначально определяется наугад).

Использование измерителя громкости в прошедших программах одного жанра дает руководство по установке уровней. Первый опыт в NDR, ORF и RTBF показал, что для прямого сведения вполне возможно окно ± 1 LU, допустимое R 128.

Тем, кто работает с файлами, проще, т.к. всю программу можно быстро и легко нормализовать до **-23 LUFS** путем смещения уровня.

Организации или отделы, которые могут перейти на измерение громкости уже сейчас, тут же получают выгоду. Возможное увеличение динамического диапазона будет полезно для шума толпы в спортивных передачах – например, расширяя впечатление от игры для зрителей и слушателей.

EBU Tech Doc 3343 [3] дает практическое руководство по новому методу работы с уровнями звука.

Есть также **два способа** нормализации громкости для **потребителя**. Первый – фактическая **нормализация самого источника**, чтобы программы были одинаково громкими на этапе проектирования. Второй метод – **использование метаданных громкости**, которые описывают, насколько громка программа. В последнем фактические средние уровни громкости программ не нужно менять на нормализованное значение, и они по-прежнему могут быть очень разными. Для обладателей новейшего оборудования нормализация может производиться у потребителя с помощью индивидуальных значений метаданных громкости для выравнивания программ до одинакового уровня воспроизведения.

Сокращения

dBFS	dB относительно цифровой полной шкалы	LU	Loudness Unit Единица громкости
dBTP	dB True-Peak относительно цифровой полной шкалы, в измерении измерителем реальных пиковых значений	LUFS	K-взвешенная единица громкости относительно цифровой полной шкалы
FM	Frequency Modulation Частотная модуляция	PML	Permitted Maximum Level Максимально допустимый уровень
LKFS	Громкость, K-взвешивание, относительно полной шкалы	PPM	Peak Programme Meter Пиковый вольтметр
LRA	Loudness Range Диапазон громкости	QPPM	Quasi-Peak Programme Meter Квазипиковый вольтметр

В парадигме регулировки уровней EBU R 128 рекомендуется первое решение из-за следующих плюсов:

- простота;
- потенциальное повышение качества в источнике.

Нормализация громкости источника рекомендуется в производстве из-за простоты и потенциального повышения качества.

Второе решение не запрещается (см. также документ *Distribution Guidelines – EBU Tech Doc 3344* [5]), но единое число (**-23 LUFS**) имеет большую силу в распространении концепции регулировки

уровня громкости, т.к. легко в понимании и в руководстве к действию. Активная нормализация источника в некотором смысле «наказывает» слишком сжатые сигналы и таким образом автоматически стимулирует производственный персонал думать о других, более динамичных и творческих способах создания впечатления от своей программы. Другими словами, фактическое **техническое** изменение исходного уровня через активную нормализацию до **-23 LUFS** имеет прямое влияние на **художественный** процесс – и, притом положительное!

Вместе с тем необходимо подчеркнуть, что оба метода могут дополнять друг друга: их не следует считать оппонентами или черно-белым представлением одной проблемы. Оба подхода являются частью R 128 – но из-за вышеупомянутых преимуществ рекомендуется **нормализация источника**.

Работа с измерителями громкости

До настоящего времени сведение с QPPM и нормализация пиков производились относительно максимально допустимого уровня (обычно -9 dBFS) и ограничитель пиков на этом уровне обеспечивал «безопасный потолок», который может быть жестким по желанию – конечно, за счет менее привлекательного звука.

В отличие от этого, парадигма регулировки уровня громкости больше напоминает «движение в пространстве», как видно в схематическом представлении столбчатого измерителя на *Рис. 5*. На *Рис. 6* показано, как может выглядеть программный измеритель на базе «стрелки».

EBU намеренно не определил никаких графических деталей или деталей пользовательского интерфейса измерителя громкости, но определил расширения алгоритма, описанного в ITU-R BS.1770 и две шкалы:

- **Шкала EBU +9**, которая должна подходить для большинства программ;
- **Шкала EBU +18**, которая может использоваться для программ с широким LRA.

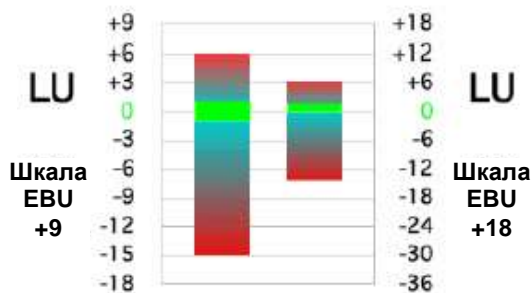


Рис. 5
Схематическое представление двух шкал громкости (здесь в LU) согласно EBU Tech 3341

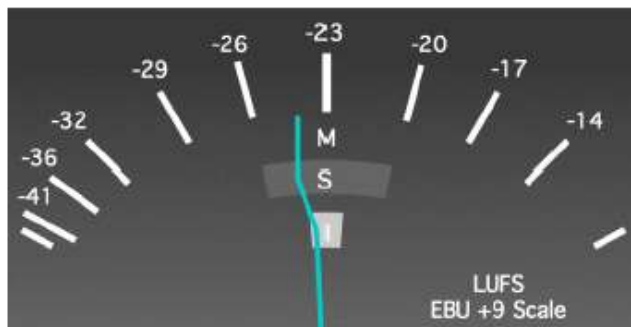


Рис. 6
Схематическое представление смоделированного измерителя громкости с «гибкой стрелкой»

Обе шкалы могут отображать либо относительный уровень громкости в **LU**, либо абсолютный уровень в **LUFS**. Производители прибора из группы PLOUD согласились реализовать набор параметров “EBUmode” для гарантии, что показания их приборов будут согласованы. Многие производители также приняли “EBUmode” или находятся на пути к этому.

Измеритель громкости “EBU Mode” согласно EBU Tech Doc 3341 [6] предлагает три шкалы времени:

- **Моментальная** громкость (сокращенно “**M**”) – временной интервал: **400 мс**
- **Кратковременная** громкость (сокращенно “**S**”) – временной интервал: **3 с**
- **Интегрированная** громкость (сокращенно “**I**”) – от “**start**” до “**stop**”

Интервалы **M** и **S**³ используются для мгновенной регулировки уровня и сведения аудио сигналов. Первичную установку уровня лучше всего делать с измерителем моментальной громкости, регулируя уровень ключевых или анкерных элементов (таких как речь, музыка или звуковые эффекты) около целевого уровня **-23 LUFS**. Советуем внимательно устанавливать начальные уровни, т.к. легче постепенно прибавлять интегрированный уровень громкости во время сведения, чем убавлять его. Обычно небольшое повышение в ходе программы также более естественно – изначально «защитная» стратегия оставляет оператору место для маневров в случае неожиданных или непредсказуемых сигналов и событий.

Когда уровни установлены, звукооператор может переключиться на **сведение только по слуху**. Наблюдение моментальной или кратковременной громкости и иногда взгляд на значение интегрированной громкости должны давать подтверждение, что сведение находится в пределах допуска вокруг целевого уровня. С числовым показанием I-значения, с точностью до десятичной точки, или с графическим отображением аналогичного разрешения можно предвидеть тенденции и принимать соответствующие меры.

В общем ... улучшение измерений звука путем замены PPM на измеритель громкости – шаг к лучшему измерительному прибору – человеческому слуху.

Громкость в цепи распространения

EBU Tech Doc 3344 [5] определяет соответствующие установки и обработку звука после выхода из вещательного центра и принимает парадигму громкости на всем пути до оборудования потребителя, включая приставки и AV приемники. Таким образом, он также стимулирует передачи к соответствию EBU R 128. В этот документ включены цифровая и аналоговая передача, ретрансляция, местные рекламные вставки, добавление новых слуг и многое другое.

Вместо измерения громкости отдельных программ, сетевая компания контролирует службу 24 часа, уделяя особое внимание коммутируемым или общим службам. Как и программы, любые метаданные громкости, передаваемые с цифровыми службами, контролируются, чтобы фактический уровень громкости услуги можно было сравнить с указанным уровнем, оба из которых должны быть, конечно, **-23 LUFS!** Раз в сутки эти данные анализируются и если служба отклоняется более чем на 0.9 LU от целевого уровня (точное значение и метод пока обсуждаются), рулевой механизм применяет поправочный коэффициент, чтобы долговременное среднее значение звука оставалось на целевом уровне ± 1 LU.

Коммерческие последствия

Поскольку парадигма регулировки уровня громкости влияет на все этапы вещательного аудио сигнала от сбора до распределения и передачи и поскольку конечная цель состоит в гармонизации уровней громкости звука внутри вещательных каналов, а также между ними для получения одинакового универсального уровня громкости в интересах слушателей, **это изменение повлияет на всех профессионалов звука и измерительное оборудование во всех частях цепи.**

Для многих главный вопрос будет в том, нужно ли заменять все существующие измерители QPPM и ин-структировать весь соответствующий персонал по новому методу работы. Ответ – «в дальней перспективе, несомненно, да», но этот переход не нужно делать сразу, хотя несколько измерителей громкости нужно ввести в эксплуатацию при первой возможности, рядом с QPPM. Замена измерителей может происходить в соответствии с техническими циклами обновления другого оборудования, модернизацией производственного оборудования или отдельными поэтапными проектами (смотря, что осуществимо в разумно сжатые сроки), и персонал можно обучать в соответствующее время.

³ “**M**” и “**S**” используется в стереофонии для обозначения “Mid” и “Side”. Для различия «моментального» и «кратковременного» времени интеграции можно использовать версии “**MLK**” и “**SLK**” (а также “**ILK**”). “**LK**” означает “Level, K-weighted”, и соответствует международному стандарту имен ISO 80000-8.

Те, кто несет ответственность за приобретение оборудования, также должны знать, что защитные ограничители для предотвращения перемодуляции смогут работать в **режиме реальных пиковых значений** и должны быть настроены на максимально допустимый уровень реальных пиковых значений в производстве, а также на выходе центральной аппаратной, в головном узле распределения и в передающем пункте.

В заключение

EBU R 128 и четыре вспомогательных документа обеспечивают средства для окончания «войны громкости» – наконец! Использование звуковой динамики опять становится творчеством. Еще есть чему учиться, и нужно время, чтобы люди привыкли к новым методам работы, но это стоит того.

Свыше 230 участников вступили в группу EBU PLOUD (на август 2010 г.), рефлексоры электронной почты показали невиданный уровень активности, а производители измерительных приборов подготовили технику для показа на IBC 2010, прежде чем была опубликована спецификация.

Пора водить R 128 в действие!



Логотип EBU R 128

Официальное уведомление

Автор особенно хочет поблагодарить Ian Rudd, независимого технологического стратега и медиа консультанта, за существенный редакторский вклад, а также Frans de Jong, EBU, и Andrew Mason, BBC R&D, за ценный вклад в эту статью.

Ссылки

- [1] EBU Technical Recommendation R 128: **Loudness normalisation and permitted maximum level of audio signals**
- [2] ITU-R BS.1770: **Algorithms to measure audio programme loudness and true-peak audio level**
- [3] EBU Tech Doc 3343: **Practical Guidelines for Production and Implementation in accordance with EBU Technical Recommendation R 128 – публикация ожидается**
- [4] EBU Tech Doc 3342: **Loudness Range: A descriptor to supplement loudness normalisation in accordance with EBU R 128**
- [5] EBU Tech Doc 3344: **Practical Guidelines for Distribution of Programmes in accordance with EBU R 128 – публикация ожидается**
- [6] EBU Tech Doc 3341: **Loudness Metering: 'EBU Mode' metering to supplement loudness normalisation in accordance with EBU R 128**

Эта версия: 6 сентября 2010

Опубликовано European Broadcasting Union, Женева, Швейцария

ISSN: 1609-1469

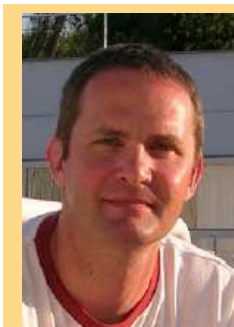
Ответственный редактор: Lieven Vermaele

Редактор: Mike Meyer

E-mail: tech@ebu.ch



Ответственность за мнения, выраженные в статье, лежит исключительно на авторе.



Florian Camerer - старший звукорежиссер ORF, австрийской вещательной корпорации, Вена. Начал работу в области звукового производства, перейдя к монтажу и микшированию звука. Основным полем деятельности были документальные фильмы, где проявил особый интерес к объемному звуку. В 1995 г. сводил первый документальный фильм ORF с многоканальным звуком 5.1 и с тех пор активно занимался объемным звуком, помогая ORF стать первым европейским вещателем, передающим прямой сигнал с объемным звуком 5.1 (New Year's Concert, 2003).

В 2008 г. г-н Camerer попросил EBU создать рабочую группу для изучения вопросов громкости, что привело к созданию возглавляемой им группы PLOUD. PLOUD – крупнейшая и самая активная группа EBU, которая отражена в большом количестве публикаций.

Florian Camerer читает лекции на международной основе по вопросам объемного звука и громкости.