

УДК 681.84

ОЦЕНКА ВЛИЯНИЯ АНАЛОГО-ЦИФРОВОГО ПРЕОБРАЗОВАНИЯ НА ИСКАЖЕНИЯ ПРИ ПЕРЦЕПЦИАЛЬНОМ КОДИРОВАНИИ АУДИОДАННЫХ

И. А. Оболонин, доцент каф. САПР СибГУТИ, к.т.н.

Н. А. Рыговская, аспирант СибГУТИ; n.rygovskaya@vgroup.ru

Ключевые слова: компрессия аудиоданных, искажения сигнала, цифро-аналоговые преобразователи, аналого-цифровые преобразователи, групповое время запаздывания.

Введение. В настоящее время в радиовещании, профессиональной и бытовой аудиотехнике широко используется компрессия цифровых аудиоданных, неизбежно сопровождаемая искажениями качества стереофонического звучания. При этом качество аналого-цифрового преобразования (АЦП) подразумевается идеальным.

В работе рассматриваются три основных фактора АЦП, которые, по мнению авторов, оказывают существенное влияние на качество цифрового звука (с компрессией аудиоданных), при высококачественном стереофоническом воспроизведении.

В психоакустических моделях стандартов MPEG не учтены механизмы временной маскировки сигналов; пространственной демаскировки источников звука, составляющих стереопанораму как по фронту, так и по глубине; особенности восприятия реверберационных составляющих стереофонических сигналов.

Именно эти механизмы пространственного слуха играют наиболее важную роль при стереовоспроизведении. Они определяют восприятие основных признаков качества стереофонического звучания, таких как пространственное восприятие, прозрачность звучания, естественность и богатство тембров инструментов и голосов, восприятие акустической атмосферы первичного помещения. Это приводит к снижению качества звучания, которое отчетливо замечают слушатели.

Однако на качество звучания существенное влияние оказывает и собственно АЦП исходного аудиосигнала.

В предлагаемой работе рассматриваются специфический характер спектра ошибок квантования, фильтрация сигналов при АЦП и цифро-аналоговых преобразованиях (ЦАП) и защита АЦП от перегрузок, вызывающая нелинейные искажения аудиосигнала.

Влияние ошибок квантования. В технической литературе [1] вопрос о спектре ошибок квантования обычно рассматривается для одного частного случая широкополосного звукового сигнала, имеющего равномерное распределение по частоте и уровню. Спектральная плотность мощности шума квантования для такого сигнала имеет равномерное распределение от 0 Гц до частоты Найквиста. Фактически этот шум является белым и, в отличие от искажений цифрового звука, он воспринимается на слух без неприятных ощущений.

Очевидно, что музыкальный сигнал нельзя отождествлять с шумовым, так как он несет звуковую информацию и отличается тем, что ошибки квантования в нем коррелированы с самим сигналом и поэтому на слух они часто воспринимаются как призвуки.

Квантованию подвергается амплитудно-импульсный модулированный сигнал (АИМ-сигнал) с широким дискретным спектром. Ошибки квантования, возникающие

при этой операции, — ошибки квантования АИМ-сигнала. Это следует учитывать, поскольку в литературе под ошибкой квантования часто понимают ошибку, возникающую при квантовании синусоидального сигнала, хотя спектры этих двух ошибок не сопоставимы. Таким образом, цифровой звуковой сигнал отличается от аналогового тем, что он — продукт АИМ.

В [1] показано, что спектр ошибок квантования и их звучание зависят от коэффициента кратности k , частоты дискретизации f_s и частоты сигнала F . Кроме того, спектр ошибок квантования существенно отличается при четных и нечетных значениях k , как на кратных, так и на субкратных частотах.

В [1] также исследовано звучание ошибок квантования в зависимости от вида звуковых тестовых сигналов и коэффициента кратности k . Главное отличие отклонений звукового сигнала от субкратных и кратных частот заключается в том, что у этих сигналов принципиально разные спектры ошибок квантования, и это оказывает влияние на звучание ошибок квантования.

Установлено, что при синусоидальных сигналах независимо от частоты и уровня сигнала ошибки квантования не воспринимаются как шум. Даже при небольших отклонениях сигнала от кратных и субкратных частот, когда частотный интервал между составляющими спектра ошибок квантования во всем диапазоне меньше 1 Гц, их звучание напоминает не шум, а рокот, который в музыкальных сигналах создает впечатление «грязного» звучания.

Влияние фильтрации. На входе АЦП и на выходе ЦАП находятся фильтры нижних частот (ФНЧ), ограничивающие спектр входных частот и устраняющие ВЧ составляющие выходного сигнала.

Подавление сигнала ФНЧ на частоте, равной половине частоты дискретизации, должно быть не менее 60 дБ [2]. При этом крутизна ската ФНЧ получается очень высокой, а это требует использования фильтров высокого порядка. Кроме того, такие фильтры имеют недостатки, главный из которых — нелинейная фазо-частотная характеристика, что приводит к заметным на слух искажениям аудиосигналов, проявляющихся в потере «прозрачности» звучания. Несмотря на то, что повышение разрядности и частоты дискретизации АЦП улучшают качество звука, это не оказывает большого влияния на его качественные показатели. Очевидная причина этого — сложность высококачественной фильтрации. На входе АЦП необходимо применять фильтр с очень высокой крутизной характеристики (для подавления частоты, равной половине частоты выборки или превышающей ее). Фильтры с частотой дискретизации 44,1 кГц (их иногда называют фильтрами типа «кирпичная стена») должны пропускать частоту 20 кГц и при этом отсекают 22,05 кГц. Кроме того, при записи и воспроизведении неизменно добавляются шумы квантования.

Повышение эффективности звукового вещательного сигнала связано с развитием техники АЦП. В частности, появление однобитного цифрового формата (DSD) [3] в каналах записи и передачи сигнала звукового вещания (ЗВ) было обусловлено совершенствованием дельта-сигма преобразователей и увеличением пропускной способности трактов хранения и передачи звуковых сигналов.

Известный метод оптимизации соотношения цена/качество (в понятие цены, кроме стоимости заложена аппаратная сложность, габаритно-весовые показатели, сроки разработки) — использование кодеров (декодеров), в которых АЦП (ЦАП) работают на повышенной частоте, что позволяет значительно снизить требования к крутизне ската аналогового ФНЧ. В этом случае основное затухание на граничной частоте полосы пропускания обеспечивается цифровым фильтром. При выборе структуры цифрового фильтра преимущества имеют нерекурсивные фильтры, позволяющие получить линейную характеристику.

На рис. 1 показаны рассчитанные [4] зависимости группового времени запаздывания от нормированной частоты (ω_n — верхняя граничная частота сигнала равна 1) для ФНЧ Баттерворта 2-го — $\tau(\omega)$, 4-го — $\tau_1(\omega)$, 6-го — $\tau_2(\omega)$ и 8-го — $\tau_3(\omega)$ порядков, а на рис. 2 — для ФНЧ Чебышева 2-го — $\tau_1(\omega)$, 4-го — $\tau_2(\omega)$, 6-го — $\tau_3(\omega)$ и 8-го — $\tau_4(\omega)$ порядков. Для фильтров Баттерворта неравномерность в полосе пропускания не превышает 3,2 мс (фильтр 8-го порядка). При этом обеспечивается рабочее затухание на граничной частоте полосы пропускания A_{pmin} — 5 дБ (передача сигналов ЗВ) и 13 дБ (цифровая звукозапись).

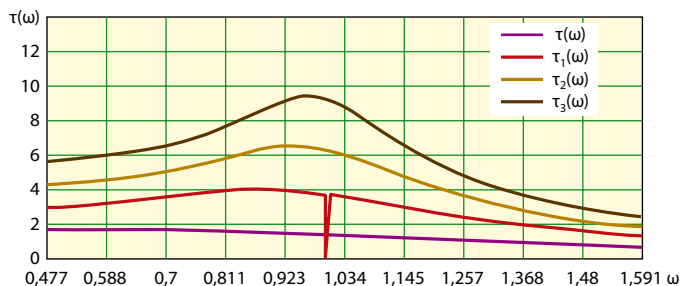


Рис. 1

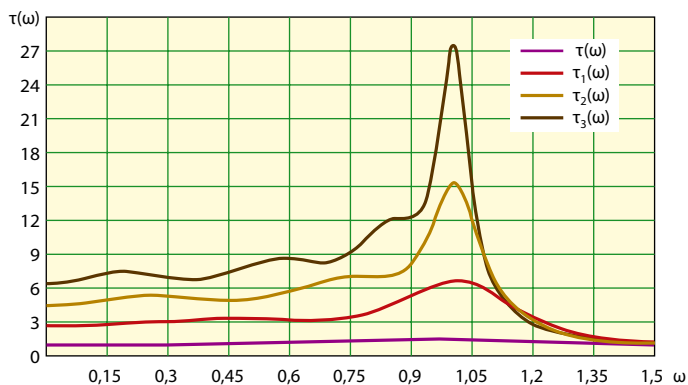


Рис. 2

Для фильтра Чебышева используется фильтр не выше 6-го порядка (неравномерность не превышает 11 мс), что для слуха вполне допустимо. При этом ФНЧ обеспечивает A_{pmin} — 28 дБ (звукозапись) и 10 дБ (передача сигналов ЗВ). Оценивая допустимую неравномерность группового времени запаздывания, авторы руководствовались нормами ГОСТ Р 50757-95.

Влияние защиты АЦП от перегрузок. При компрессии аудиоданных в психоакустических моделях стандартов MPEG и Dolby AC-3 шаг квантования изменяется в соответствии с уровнем кодируемого сигнала в субполосе. Однако кодирование ведется так, что уровень шумов не превышает относительный порог слышимости. В этом случае следует учесть, что при первичном ИКМ-преобразовании уровень аудиосигнала может превышать квазимаксимальный, на основе которого выбирается порог перегрузки АЦП. За счет ограничения уровня входного сигнала АЦП возникают нелинейные искажения. В психоакустической модели осуществляется перераспределение количества бит, отводимых на субполосу в фиксированный интервал времени. При росте уровня сигнала количество бит уменьшается, т. е. растет шаг квантования. Вместе с первичным сигналом (с таким же шагом квантования) квантуются и высшие гармоники, возникшие при превышении квазимаксимального уровня.

Перегрузки АЦП на входе цифровых аудиотрактов недопустимы, поскольку АЦП порождают собственные побочные эффекты (артефакты) и возникающие искажения значительно отличаются от нелинейных искажений в аналоговых трактах (с точки зрения слухового восприятия).

Вероятность превышения квазимаксимального уровня для сигнала с гауссовским законом распределения составляет примерно 0,02, что соответствует появлению при воспроизведении 6 щелчков в течение 5 мин (только за счет перегрузки АЦП) [5]. Нормы допускают не более одного заметного на слух щелчка в час, поэтому необходимы меры, не допускающие перегрузку АЦП.

Простейшее решение этой задачи — применение на входе АЦП жесткого (безынерционного) ограничения на уровне, не превышающем порог перегрузки. Это приводит к появлению нелинейных искажений и потере ВЧ составляющих в области превышения порога ограничения.

Другое решение рассматриваемой задачи — компрессия сигнала в верхней части его динамического диапазона. В этом случае сохраняются ВЧ составляющие исходного сигнала, однако нелинейные искажения неизбежны. От выбора закона компрессии будет зависеть уровень нелинейных искажений при воспроизведении аудиосигналов. Кроме того, следует учитывать вероятность появления того или иного уровня сигнала. На рис. 3 приведена функция распределения уровней N , полученная экспериментальным путем [6] при исследовании записанных на магнитной ленте музыкальных и речевых отрывков сигналов ЗВ. Можно считать, что такое распределение подчиняется гауссовскому закону с параметрами $m = -24$ дБ, $\sigma = 10$ дБ.

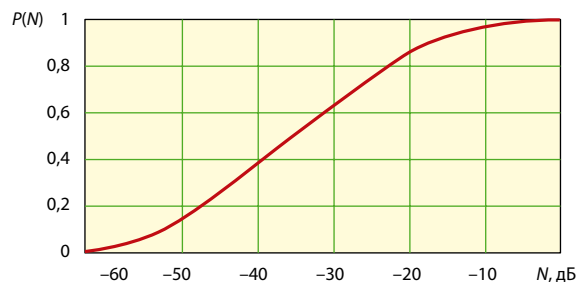


Рис. 3

Известны исследования зависимости заметности искажений перегрузки АЦП от коэффициента перегрузки K_n [6]

$$K_n = 20 \log |X_{max}| / |X_n|,$$

где X_{\max} — максимальный по модулю отсчет; X_n — порог перегрузки АЦП, полученные субъективно-статистическим путем (рис. 4).

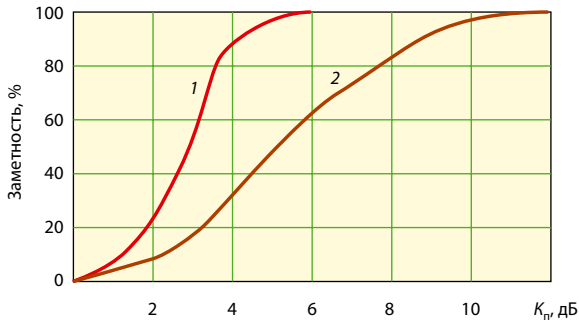


Рис. 4

Из рисунка 4 видно, что искажение перегрузки незаметно при $K_n \approx 1$ дБ, где кривая 1 — усредненное значение для различных передач при радиотрансляции; кривая 2 — для женского сольного пения (критический фрагмент). В аппаратуре часто используют защитный интервал в 3 дБ ($K_n = 3$ дБ), что приводит к уменьшению отношения сигнал/искажения квантования. Кроме того, нет гарантии, что диаграмма уровней в канале стабильна.

На рис. 5 показана амплитудная характеристика преобразователя (на входе АЦП), осуществляющего компрессию сигнала в верхней части его динамического диапазона, по логарифмическому закону. Верхние 5 дБ используются для создания области, где уровень преобразованного сигнала представлен в логарифмическом соотношении. Это позволяет представить в пределах всего лишь 5 дБ фронты сигнала, уровень которых гораздо выше точки, где начинается область перегрузки. Такой прием подобен эффекту компрессии звука, появляющегося при высококачественной записи аудиосигналов на магнитную ленту в аналоговом виде (из-за нелинейности кривой намагничивания).

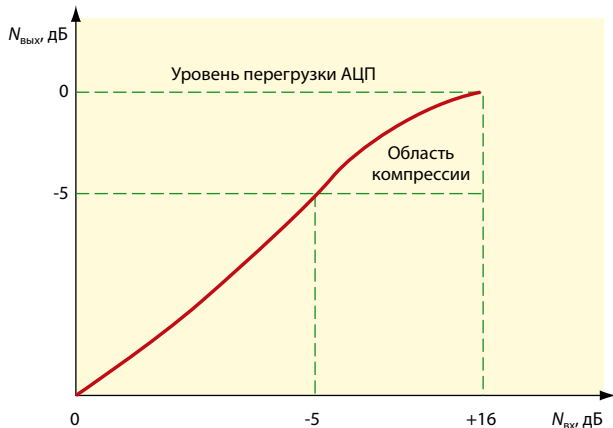


Рис. 5

Использование компрессии сигнала в верхней части позволяет задействовать динамический диапазон, превышающий возможности самих АЦП и сохранить максимальное количество информации об аналоговом сигнале, которая затем кодируется с разрешающей способностью конкретного преобразователя. Кроме того, в этом случае можно сохранить высокочастотные составляющие сигнала.

Оценка нелинейных искажений. Для одночастотного периодического сигнала методом коэффициентов гармо-

ник была получена [7] сравнительная оценка нелинейных искажений, возникающих при безынерционном ограничении и компрессии верхней части динамического диапазона звукового сигнала (по логарифмическому закону). На рис. 6 представлены полученные зависимости коэффициента гармоник для ограничителя $K_G(N)$ (кривая 1) и компрессора верхней частоты динамического диапазона $K_{ГК}(N)$ (кривая 2), от уровня сигнала N .

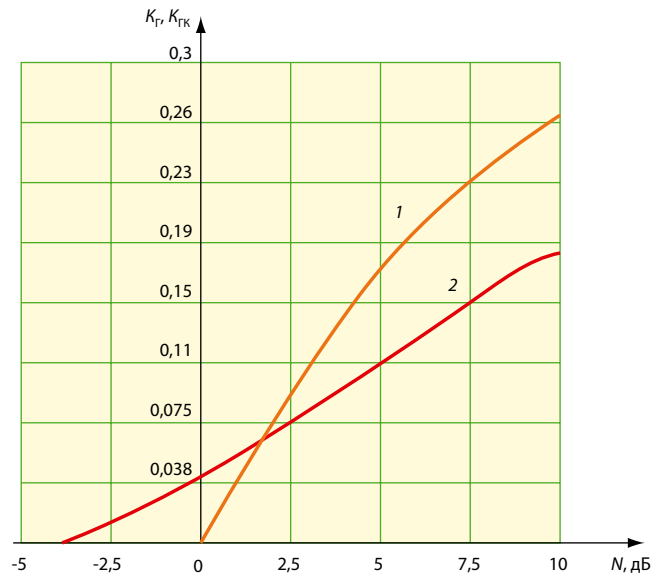


Рис. 6

Вероятность того, что коэффициент гармоник превысит то или иное значение зависит от появления различных уровней сигнала. Вероятность проявления сигнала уровня N (для функции распределения уровней, представленных на рис. 1):

$$P_1(N) = 1 - \left[\frac{1}{\sigma\sqrt{2\pi}} \int_{N_{\min}}^N e^{-\frac{(N-\alpha)^2}{2\sigma^2}} \right],$$

где в соответствии с экспериментальными данными [8] $N_{\min} = -60$ дБ, $\alpha = -24$ дБ, $\sigma = 10$ дБ.

Вероятность того, что K_G превысит то или иное значение, равна произведению K_G для заданного уровня сигнала $K_{ГК}$ на вероятность его превышения:

— для ограничителя

$$P(K_G > K_{ГН}) = K_G(N)P_1(N),$$

— для компрессора

$$P(K_{ГК} > K_{ГКН}) = K_{ГК}(N)P_1(N).$$

Для рассматриваемого примера (музыкальные и речевые отрывки сигналов ЗВ) были рассчитаны вероятности превышения K_G некоторого уровня в диапазоне изменения уровня сигнала от -5 до 10 дБ, не превышают $4 \cdot 10^{-4}$.

Это позволяет сделать вывод о незначительном влиянии нелинейных искажений на качество звучания при защите АЦП от перегрузок с помощью компрессии верхней части динамического диапазона сигнала (для статистик аудиосигналов приведенных в [5]). Однако нельзя гарантировать, что статистики разнообразных аудиосигналов с большой вероятностью будут близки к приведенным в [5].

Известно [6], что отношение сигнал/шум квантования для квазимаксимального аудиосигнала:

$$\frac{P_{\text{сквmax}}}{P_{\text{шкв}}} = 6m - 16, \text{ дБ},$$

где m — число бит на отсчет.

Отношение сигнал/шум для продуктов нелинейности второго и третьего порядка можно получить из выражения:

$$\begin{aligned} \frac{P_{\text{сквmax}}}{P_{\text{ни}}} &= 10 \lg \frac{U_{\text{квmax}}^2 / R}{U_{22}^2 / R + U_{32}^2 / R} = \\ &= 10 \lg \frac{U_{\text{квmax}}^2}{U_{22}^2 + U_{32}^2} \approx 10 \lg \frac{1}{K_{\Gamma 2}}, \text{ дБ}, \end{aligned}$$

где $P_{\text{ни}}$ — мощность сигналов продуктов нелинейности второго и третьего порядков; U_2 и U_3 — напряжения второй и третьей гармоник соответственно. Для примера можно предположить, что сигнал превышает квазимаксимальный на 4 дБ. Тогда отношение сигнал/шум квантования составит

$$\frac{P_c}{P_{\text{шкв}}} = 6 \cdot 6 - 16 = 20 \text{ дБ}.$$

Если для приведенного примера $K_{\Gamma} = 0,1$ то отношение

$$\frac{P_c}{P_{\text{шкв}}} = 10 \lg \frac{1}{K_{\Gamma 2}} = 20 \text{ дБ}.$$

В этом случае уровень шумов квантования равен уровню сигнала продуктов нелинейности.

Как известно, аудиосигнал — нестационарный случайный процесс. Следовательно, нельзя исключить существенного превышения порога перегрузки АЦП (пусть и кратковременного). При этом уровень сигналов продуктов нелинейности становится больше уровня шумов квантования, и их влияние на качество воспроизводимого сигнала увеличивается.

Заключение. Проведенный анализ позволяет сделать вывод о том, что наибольшее влияние на качество воспроизводимого звука при его цифровой передаче или звукозаписи (кроме собственно компрессора цифровых аудиоданных) оказывают ошибки квантования, проявляющиеся в виде звучаний, похожих на звон колокольчиков и рокот. При построении АЦП авторы рекомендуют применять компрессию входного сигнала в верхней части его динамического диапазона. Это позволит устранить потери ВЧ составляющих сигнала и уменьшить нелинейные искаже-

ния, по сравнению с методом ограничения уровня входного сигнала АЦП. Нелинейные искажения зависят от выбора вида компрессий, но ввиду малой вероятности появления больших уровней сигнала решающим здесь является отсутствие перегрузок АЦП, которые приводят к нежелательным последствиям.

Для реализуемых в аналоговой схемотехнике фильтров предварительную фильтрацию предпочтительно выполнять с помощью фильтров Чебышева, обеспечивающих приемлемую (с точки зрения слухового восприятия допустимых искажений) неравномерность группового времени запаздывания и сравнительно большое затухание на граничной частоте полосы пропускания. Для получения линейной фазо-частотной характеристики в АЦП и ЦАП с передискретизацией необходимо использовать нерекурсивные цифровые фильтры.

ЛИТЕРАТУРА

1. Вологдин Э. Как возникают и звучат ошибки квантования // Звукорежисер. — 2007. — № 6, 7. — С. 28—41, 32—40.
2. Золотухин И. П., Изюмов А. А., Райзман М. М. Цифровые звуковые магнитофоны. Томск: Радио и связь, Томский отдел, 1990. — 160 с.
3. Щитов Ю. Direct Stream Digital: одноканальный цифровой формат записи // Звукорежисер. — 1999. — № 2. — С. 23—28.
4. Оболонин И. А. Анализ искажений аудиосигналов при цифровых методах передачи и звукозаписи. // Информатика и проблемы телекоммуникаций: тезисы докл. Российской НТК. — Новосибирск, 24—25 апреля 2008. — С. 233—234.
5. Ковалгин Ю. А. Радиовещание и электроакустика. — М.: Радио и связь, 1999. — 792 с.
6. Выходен А. В., Гитлиц М. В. Радиовещание и электроакустика. — М.: Радио и связь, 1989. — 542 с.
7. Оболонин И. А., Портфиненко А. Н. Оценка искажений при компрессии аудиосигнала на входе цифровых трактов. // Перспективы развития современных средств и систем телекоммуникаций: тезисы докл. 11-й международной НТК. — Екатеринбург, 23—25 апреля 2005. — С. 117—120.
8. Дворецкий И. М., Дриацкий И. Н. Цифровая передача сигналов звукового вещания. // М.: Радио и связь, 1987. — 191 с.
9. Hans M. and Schafer R. W. Lossless Compression of Digital Audio/Client and Media Systems Laboratory, HP Laboratories Palo Alto HPL. — Nov., 1999. — 144.
10. ISO/IEC 14496—3. Information technology — Coding of audio-visual objects. Part 3: Audio.

Получено после доработки 22.01.10

ИНФОРМАЦИЯ

НОВАЯ УСЛУГА ОТ «РОСТЕЛЕКОМА»

«Ростелеком» ввел в опытно-коммерческую эксплуатацию услугу «Интеллектуальная аудиоконференц-связь», которая позволяет пользователю организовать сеанс телефонной связи одновременно между несколькими участниками с возможностью самостоятельно управлять подключениями через Web-интерфейс.

В сеансе аудиоконференцсвязи могут принимать участие до 128 человек, находящихся как на территории Российской Федерации, так и за ее пределами.

В России пользователи начинают сеанс связи, набирая выделенный интеллектуальный номер в коде «8-800», а участни-

ки за рубежом — многоканальный номер +7 (495) 727-45-89.

Заказчик может самостоятельно забронировать время для проведения сеанса аудиоконференцсвязи на сайте «Ростелекома» (www.rt.ru). После ввода даты, времени и продолжительности планируемого сеанса, а также количества участников система автоматически резервирует необходимые ресурсы сети, а заказчик получает PIN-код для доступа участников к данному сеансу. Интеллектуальная система авторизации гарантирует защиту от несанкционированного доступа.

По словам директора Департамента продуктов и маркетинга «Ростелекома»

Р. Левочки, «предоставленная заказчику возможность самостоятельно управлять параметрами аудиоконференцсвязи через Web-интерфейс на всех этапах оказания услуги — от бронирования до проведения самого сеанса связи — является уникальным преимуществом новой услуги».

Цена новой услуги формируется из единовременной платы за выделение интеллектуального номера, ежемесячной абонентской платы и оплаты входящего трафика, сгенерированного участниками сеанса. Услугу оплачивает заказчик, для остальных участников сеанса услуга бесплатна.