



МЕЖДУНАРОДНЫЙ СОЮЗ ЭЛЕКТРОСВЯЗИ

МСЭ-Т

СЕКТОР СТАНДАРТИЗАЦИИ
ЭЛЕКТРОСВЯЗИ МСЭ

P.862.3

(11/2005)

СЕРИЯ P: КАЧЕСТВО ТЕЛЕФОННОЙ ПЕРЕДАЧИ,
ТЕЛЕФОННЫЕ УСТАНОВКИ, СЕТИ МЕСТНЫХ
ЛИНИЙ

Методы для объективной и субъективной оценки
качества

**Руководство по применению объективных
измерений качества, основанных на
Рекомендациях P.862, P.862.1 и P.862.2**

Рекомендация МСЭ-Т P.862.3

РЕКОМЕНДАЦИИ МСЭ-Т СЕРИИ Р

КАЧЕСТВО ТЕЛЕФОННОЙ ПЕРЕДАЧИ, ТЕЛЕФОННЫЕ УСТАНОВКИ, СЕТИ МЕСТНЫХ ЛИНИЙ

Словарь и воздействие параметров передачи на мнение клиента о качестве передачи	Серия	P.10
Абонентские линии и аппараты	Серия	P.30 P.300
Стандарты передачи	Серия	P.40
Аппарат объективного измерения	Серия	P.50 P.500
Объективные электроакустические измерения	Серия	P.60
Измерения, относящиеся к громкости речи	Серия	P.70
Методы объективной и субъективной оценки качества	Серия	P.80 P.800
Аудиовизуальное качество в мультимедийных услугах	Серия	P.900
Характеристики передачи и аспекты КО конечной точки в IP-сети	Серия	P.1000

Для получения более подробной информации просьба обращаться к перечню Рекомендаций МСЭ-Т.

Рекомендация МСЭ-Т Р.862.3

Руководство по применению объективных измерений качества, основанных на Рекомендациях Р.862, Р.862.1 и Р.862.2

Резюме

В настоящей Рекомендации содержатся некоторые важные замечания, которые следует учитывать при объективной оценке качества речи в соответствии с Рекомендациями МСЭ-Т Р.862, Р.862.1 и Р.862.2. Пользователи Рекомендации МСЭ-Т Р.862 должны понимать и следовать указаниям, приведенным в настоящей Рекомендации.

Настоящая Рекомендация представляет собой дополнительное руководство для пользователей Рекомендации МСЭ-Т Р.862, в котором рекомендованы средства оценки качества прослушиваемой речи, с использованием эталонных образцов речевых сигналов и образцов с ухудшенным качеством. Сфера применения Рекомендации МСЭ-Т Р.862 четко определена в самой Рекомендации. В настоящей Рекомендации не расширяется и не сужается эта сфера применения, но в ней содержится необходимая и существенная информация для получения на практике стабильных, надежных и содержательных результатов объективных измерений.

Источник

Рекомендация МСЭ-Т Р.862.3 утверждена 29 ноября 2005 г. 12-й Исследовательской комиссией МСЭ-Т (2005–2008 гг.) в соответствии с процедурой, изложенной в Рекомендации МСЭ-Т А.8. Настоящий текст содержит разъяснения, согласованные 12-й Исследовательской комиссией МСЭ-Т 13 июня 2006 г.

ПРЕДИСЛОВИЕ

Международный союз электросвязи (МСЭ) является специализированным учреждением Организации Объединенных Наций в области электросвязи. Сектор стандартизации электросвязи МСЭ (МСЭ-Т) – постоянный орган МСЭ. МСЭ-Т отвечает за изучение технических, эксплуатационных и тарифных вопросов и за выпуск Рекомендаций по ним с целью стандартизации электросвязи на всемирной основе.

На Всемирной ассамблее по стандартизации электросвязи (ВАСЭ), которая проводится каждые четыре года, определяются темы для изучения Исследовательскими комиссиями МСЭ-Т, которые, в свою очередь, вырабатывают Рекомендации по этим темам.

Утверждение рекомендаций МСЭ-Т осуществляется в соответствии с процедурой, изложенной в Резолюции I ВАСЭ.

В некоторых областях информационных технологий, которые входят в компетенцию МСЭ-Т, необходимые стандарты разрабатываются на основе сотрудничества с ИСО и МЭК.

ПРИМЕЧАНИЕ

В настоящей Рекомендации термин "администрация" используется для краткости и обозначает как администрацию электросвязи, так и признанную эксплуатационную организацию.

Соблюдение положений данной Рекомендации носит добровольный характер. Однако в Рекомендации могут содержаться определенные обязательные положения (например, для обеспечения возможности взаимодействия или применимости), и соблюдение положений данной Рекомендации достигается в случае выполнения всех этих обязательных положений. Для выражения необходимости выполнения требований используется синтаксис долженствования и соответствующие слова (такие, как "должен" и т.п.), а также их отрицательные эквиваленты. Использование этих слов не предполагает, что соблюдение положений данной Рекомендации является обязательным для какой-либо из сторон.

ПРАВА ИНТЕЛЛЕКТУАЛЬНОЙ СОБСТВЕННОСТИ

МСЭ обращает внимание на вероятность того, что практическое применение или реализация этой Рекомендации может включать использование заявленного права интеллектуальной собственности. МСЭ не занимает какую бы то ни было позицию относительно подтверждения, обоснованности или применимости заявленных прав интеллектуальной собственности, независимо от того, отстаиваются ли они членами МСЭ или другими сторонами вне процесса подготовки Рекомендации.

На момент утверждения настоящей Рекомендации МСЭ получил извещение об интеллектуальной собственности, защищенной патентами, которые могут потребоваться для выполнения этой Рекомендации. Однако те, кто будет применять Рекомендацию, должны иметь в виду, что это может не отражать самую последнюю информацию, и поэтому им настоятельно рекомендуется обращаться к патентной базе данных БСЭ по адресу: <http://www.itu.int/ITU-T/ipr/>.

© ITU 2006

Все права сохранены. Никакая часть данной публикации не может быть воспроизведена с помощью каких-либо средств без предварительного письменного разрешения МСЭ.

СОДЕРЖАНИЕ

	Стр.
1	Сфера применения 1
2	Справочные документы 1
3	Определения 2
4	Сокращения 2
5	Условные обозначения 3
6	Общие положения 3
6.1	Критерии испытаний 3
6.2	Приложения 4
7	Характеристики эталонных сигналов 4
7.1	Продолжительность сигналов 4
7.2	Активная речь 5
7.3	Временная структура 5
7.4	Уровень активной речи 5
7.5	Применение искусственного голоса 5
7.6	Требования к записи речи 5
7.7	Смена диктора и изменение содержания речи 6
7.8	Паузы перед началом и после окончания речи 6
7.9	Предварительная фильтрация 6
7.10	Уровень шума 6
7.11	Пути реализации 7
8	Характеристики ухудшенного сигнала, который следует оценить 7
8.1	Различия в продолжительности активной речи между эталонным речевым сигналом и ухудшенным речевым сигналом 8
8.2	Уровень активной речи 8
8.3	Различия в продолжительности пауз перед началом и после окончания речи между эталонным речевым сигналом и ухудшенным речевым сигналом 8
9	Характеристики каналов подачи и извлечения сигнала 8
9.1	Влияние измерительных линий связи и испытательных конфигураций в канале подачи сигнала 9
9.2	Влияние измерительных линий связи и испытательных конфигураций в канале извлечения сигнала 10
10	Анализ результатов 10
10.1	Усреднение результатов измерений 10
10.2	Надежность результатов измерений PESQ 10
10.3	Показатели точности измерений PESQ 11
10.4	Интерпретация результатов измерения точности 12
11	Описание результатов 12

12	Руководство по применению широкополосного дополнения Р.862.2 к Рекомендации Р.862	13
	Приложение I – Контрольные показатели объективного качества, полученные в Рекомендации Р.862 для стандартных кодеков МСЭ-Т/GSM	14
I.1	Контрольные показатели в соответствии с Рекомендацией МСЭ-Т Р.862.1, рассчитанные для следующих условий применения кодека/MNRU с использованием базы данных речевых сигналов из Дополнения В/Р.501:	14
I.2	Предварительная обработка исходного речевого сигнала	16
I.3	Обработка в соответствии с G.711	16
I.4	Обработка в соответствии с G.726	16
I.5	Обработка в соответствии с G.728, G.729, Дополнением А/G.729 и G.723.1	17
I.6	Обработка в соответствии с MNRU	17
	Приложение II – Базы данных испытаний для Р.862/Р.862.1	21
	Приложение III – Отчет об измерениях в соответствии с Р.862/Р.862.1	22
III.1	Отчет и интерпретация усредненных результатов измерений PESQ	22
III.2	Отчет и интерпретация отдельных результатов измерений PESQ	22
	Приложение IV – Метод калибровки патентованных интерфейсов	24
IV.1	Калибровка уровня передачи (на стороне передачи) испытательного оборудования	24
IV.2	Калибровка уровня приема (на стороне приема) испытательного оборудования	24
	БИБЛИОГРАФИЯ	25

Рекомендация МСЭ-Т Р.862.3

Руководство по применению объективных измерений качества, основанных на Рекомендациях Р.862, Р.862.1 и Р.862.2

1 Сфера применения

В настоящей Рекомендации содержатся некоторые важные замечания, которые следует учитывать при объективной оценке качества речи в соответствии с Рекомендациями МСЭ-Т Р.862, Р.862.1 и Р.862.2. Пользователи Рекомендации МСЭ-Т Р.862 должны понимать и следовать указаниям, приведенным в настоящей Рекомендации.

Настоящая Рекомендация представляет собой дополнительное руководство для пользователей Рекомендации МСЭ-Т Р.862, в котором рекомендованы средства оценки качества прослушиваемой речи, с использованием эталонных образцов речевых сигналов и образцов с ухудшенным качеством. Она не может применяться для оценки качества произносимой речи или качества взаимодействия. В ней предполагается, что алгоритм оценки объективного качества точно соответствует Рекомендации МСЭ-Т Р.862. Это может быть подтверждено при помощи проверки на совместимость, выполненной в соответствии с дополнением к Рекомендации МСЭ-Т Р.862.

Сфера применения Рекомендации МСЭ-Т Р.862 четко определена в самой Рекомендации. В настоящей Рекомендации не расширяется и не сужается эта сфера применения, но в ней содержится необходимая и существенная информация для получения на практике стабильных, надежных и содержательных результатов объективных измерений.

Приложения и ограничения, связанные с широкополосным дополнением Рекомендации Р.862, описанным в Рекомендации МСЭ-Т Р.862.2, рассматриваются в разделе 12.

2 Справочные документы

В нижеследующих Рекомендациях МСЭ-Т и других справочных документах содержатся положения, которые, с помощью ссылки в настоящем тексте, составляют положения настоящей Рекомендации. На время публикации указанные здесь издания были действительными. Все Рекомендации и другие справочные документы постоянно пересматриваются; поэтому всем пользователям данной Рекомендации настоятельно рекомендуется изучить возможность использования последних изданий перечисленных ниже Рекомендаций и других справочных документов. Перечень действующих на настоящий момент Рекомендаций МСЭ-Т регулярно публикуется. Ссылка в настоящей Рекомендации на какой-либо документ не придает этому отдельному документу статуса Рекомендации.

- [1] Рекомендация МСЭ-Т Р.501 (2000), *Испытательные сигналы для телефонометрии, Дополнение В – Речевые файлы и шумовые последовательности.*
- [2] Рекомендация МСЭ-Т Р.56 (1993), *Объективное измерение уровня активной речи.*
- [3] МСЭ-Т Рекомендации серии Р – Дополнение 23 (1998), *База данных кодированных речевых сигналов МСЭ-Т.*
- [4] Рекомендация МСЭ-Т Р.50 (1999), *Искусственные голоса.*
- [5] Рекомендация МСЭ-Т Р.800 (1996), *Методы субъективного определения качества передачи.*
- [6] Рекомендация МСЭ-Т Р.830 (1996), *Субъективная оценка качественных показателей цифровых кодеков телефонной полосы и широкополосных цифровых кодеков.*
- [7] Рекомендация МСЭ-Т Р.862 (2001), *Оценка воспринимаемого качества речи (PESQ): Объективные метод оценки сквозного качества речи в узкополосных телефонных сетях и речевых кодеках.*
- [8] Рекомендация МСЭ-Т Р.862.1 (2003), *Функция преобразования для перевода необработанных результатов измерений по Р.862 в MOS-LQO.*

3 Определения

В настоящей Рекомендации определяются следующие термины.

3.1 речь/сигнал источника (source speech/signal): Исходный речевой сигнал без каких-либо ухудшений. Он должен записываться и храниться в соответствии с Рекомендацией МСЭ-Т Р.830. Он может быть или не быть одним из эталонных речевых сигналов, описанных ниже.

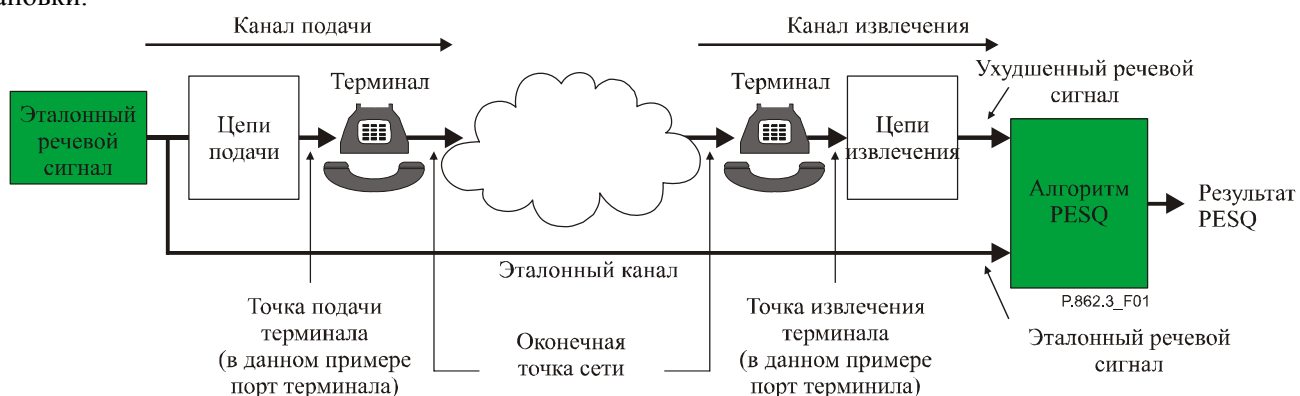
3.2 эталонный речевой сигнал (reference speech/signal): Речевой сигнал, который должен использоваться алгоритмом Рекомендации МСЭ-Т Р.862 в качестве эталона, относительно которого выявляется влияние испытываемой системы.

3.3 входной речевой сигнал (input speech/signal): Сигнал, подаваемый на испытываемую систему в точке введения сигнала. Он получается из эталонного речевого сигнала. Он может быть идентичен эталонному сигналу или может быть результатом его обработки, например, за счет наложения на него шума. Более подробная информация приведена в § 7.10.

3.4 ухудшенный речевой сигнал (degraded speech/signal): Эталонный речевой сигнал, прошедший через испытываемую систему.

3.5 канал подачи сигнала (signal insertion path): Состоит из канала связи (проводного, электронного и т. д.) между эталонным сигналом и алгоритмом по Рекомендации МСЭ-Т Р.862 и входным интерфейсом, который называется точкой подачи сигнала.

На рисунке 1 показан пример испытательной цепи, которая разделена на канал подачи, испытываемую систему и канал извлечения, на рисунке показаны возможные точки введения и извлечения сигнала для случая аппаратных измерений. Конкретные точки введения и извлечения сигнала будут зависеть от конкретной испытываемой системы конфигурации испытательной установки.



Примечание. – В зависимости от того, как сигналы подаются, извлекаются и хранятся, и от того, выполняются ли вычисления результатов измерения в соответствии с Рекомендацией МСЭ-Т Р.862 в реальном времени или с задержкой, эти каналы могут быть электрическими линиями или логическими путями, как, например, в случае, когда выходные

Рисунок 1/Р.862.3 – Пример измерительной установки и терминология

3.6 канал извлечения сигнала (signal capture path): Состоит из канала связи (проводного, электронного и т. д.) между точкой извлечения сигнала (выходным интерфейсом с испытываемой сетью) и алгоритмом по Рекомендации МСЭ-Т Р.862 (см. рис. 1).

3.7 дБов: Значение в дБ относительно точки перегрузки в цифровой системе. Согласно Рекомендации МСЭ-Т G.711, 0 дБм0 в аналоговом представлении соответствует –6,15 дБов и –6,18 дБов для кодеков типа А и типа μ , соответственно.

4 Сокращения

В настоящей Рекомендации используются следующие сокращения:

IRS	Intermediate Reference System	Промежуточная эталонная система
MOS	Mean Opinion Score	Показатель по среднему мнению

MOS-LQO	Mean Opinion Score-Listening Quality Objective (an estimation of subjective listening quality using an objective measurement technique)	Показатель по среднему мнению для объективной оценки прослушиваемой речи (оценка субъективного качества прослушивания с использованием метода объективного измерения)
MOS-LQS	Mean Opinion Score-Listening Quality Subjective (a direct measurement of listening quality using subjective ratings of samples)	Показатель по среднему мнению для субъективной оценки прослушиваемой речи (прямое измерение качества прослушивания с использованием субъективной оценки образцов речи)
PESQ	Perceptual Evaluation of Speech Quality	Оценка воспринимаемого качества речи
RMS	Root Mean Square	Среднеквадратический

5 Условные обозначения

Рекомендуется преобразовывать необработанные результаты по Рекомендации P.862 в формат MOS-LQO (как он определен в Рекомендации МСЭ-Т P.800.1) с использованием соотношения, определенного в Рекомендации МСЭ-Т P.862.1¹. Это предотвратит возможные недоразумения при сравнении и интерпретации результатов из-за кажущейся схожести этих двух шкал.

6 Общие положения

В настоящем разделе приводятся дополнительные замечания относительно сферы применения Рекомендации МСЭ-Т P.862. Сама эта сфера применения достаточно ясно определена в Рекомендации МСЭ-Т P.862.

Надежность и повторяемость результатов зависит от нескольких факторов, например:

- Количество вызовов.
- Количество измерений.
- Продолжительность образцов речи.
- Тип используемой речи, например естественная или искусственная.

Эти факторы и следующие соображения влияют на структуру и сложность испытаний:

- Цель измерений (например, для сопоставительного анализа линий связи, обычного мониторинга или для обнаружения неисправностей).
- Характеристики канала передачи (например, изменяются ли характеристики канала во времени, как это происходит в подвижной связи и в некоторых типах каналов VoIP?).
- Доступное время для измерений (которое может влиять на объем испытаний).

Если предполагается, что некоторые типы линий связи могут быть подвержены влиянию условий "часа пик", может быть важно выполнить несколько последовательных измерений в различное время суток.

Используемая структура испытаний должна быть указана вместе с обработанными значениями результатов измерений в формате MOS-LQO.

6.1 Критерии испытаний

Рекомендация МСЭ-Т P.862 действительна для оценки критериев испытаний, технологий кодирования и приложений, которые перечислены в таблице 1/P.862. В частности, следует проявлять осторожность при выполнении испытаний на работающей сети, поскольку может существовать оборудование, вызывающее ухудшение сигнала, которые не могут быть учтены в Рекомендации МСЭ-Т P.862, например, аномалии, вносимые системами шумоподавления, работающими между точкой подачи сигнала и точкой извлечения сигнала. Известно также, что PESQ недооценивает некоторые искажения линейной частотной характеристики. Это особенно важно, например, для ограничений полосы пропускания менее 300 Гц ... 3,4 кГц.

¹ Подробная процедура получения MOS-LQO описана в разделе 10.

Не рекомендуется применять Рекомендацию МСЭ-Т Р.862 с системами, в которых используются алгоритмы шумоподавления, работающие между точкой подачи сигнала и точкой извлечения сигнала.

6.2 Приложения

Рекомендация МСЭ-Т Р.862 может использоваться как средство для измерений на работающих сетях, когда оценивается система в реальных условиях работы, а не в условиях компьютерного моделирования или на лабораторных испытательных установках.

Измерения на работающих сетях не приведут к получению повторяемых результатов из-за неуправляемых изменений во времени характеристик каналов передачи. Альтернативой является управляемое моделирование сети с точно повторяемыми результатами. В последнем случае следует использовать усреднение.

Измерения на работающих сетях, например испытания подвижной связи в процессе движения, повлияют на структуру и содержание эталонных речевых сигналов. При выполнении испытаний в процессе движения необходимо оценивать качество, очень быстро изменяющееся во времени, для того чтобы получить точную информацию о качестве в каждой географической точке.

Измерения на работающих сетях также требуют оценки качества для каждого образца речи в отдельности, поскольку в условиях изменений во времени параметров работающей сети невозможно выполнить усреднение параметров для каждого условия использования.

Оба вышеприведенных соображения оказывают влияние на стабильность и, возможно, на точность результатов, полученных по Рекомендации МСЭ-Т Р.862. По этой причине пользователи Рекомендации МСЭ-Т Р.862 для измерений на работающих сетях должны быть очень внимательными при проверке результатов и повторять измерения с целью проверки стабильности результатов. Результаты измерения качественных характеристик приведены в § 10.2.

Если испытываемая система содержит широкополосный терминал (например, некоторые приспособления "hands-free" или широкополосный IP телефон), то PESQ будет определять качество, как если бы оно воспринималось для фильтрации на приеме типа IRS.

7 Характеристики эталонных сигналов

Эталонные сигналы определяются и используются как входные сигналы для испытываемой системы и как эталонные входные сигналы для Рекомендации МСЭ-Т Р.862. Характеристики канала подачи сигнала рассматриваются отдельно в разделе 9. Если язык оцениваемой речи включен в базу данных, приведенную в Дополнении В/Р.501, мы рекомендуем использовать ее в качестве испытательных сигналов для улучшения совместимости результатов различных измерений, избегая использования различных эталонных сигналов.

7.1 Продолжительность сигналов

Рекомендация МСЭ-Т Р.862 утверждена в МСЭ-Т для применения с сигналами, которые в большинстве своем имеют продолжительность 8–12 с. Однако, известно, что Рекомендация МСЭ-Т Р.862 может применяться и для образцов речи продолжительностью 30 с [В.1]. Следовательно, рекомендуется, чтобы каждый образец речевого сигнала имел продолжительность от 8 до 30 с. В эту величину включены все паузы до, после и между фрагментами активной речи ².

Для испытаний на работающих сетях могут использоваться более короткие эталонные сигналы, однако в этом случае система не тестируется на полную мощность. Такие укороченные звуковые последовательности должны содержать не менее 3,2 с речи, как определено в § 7.2.

² Эталонное программное обеспечение, представленное в виде Дополнения А/Р.862, имеет следующие ограничения относительно продолжительности сигналов, хотя эти ограничения уже выходят за пределы диапазона, определенного в настоящей Рекомендации: из-за точности, присущей вычислениям с плавающей запятой в Рекомендации МСЭ-Т Р.862, как только обрабатываемый сигнал достигает определенной длины, в расчеты энергии сигнала вкрадываются ошибки. Анализ предполагает, что сигналы, состоящие из более чем миллиона отсчетов, будут создавать проблемы. Шестьдесят секунд монофонического сигнала с частотой дискретизации 16 кГц содержат 960 000 отсчетов, и эта величина будет являться разумным порогом, после которого действует данное предупреждение.

Следует отметить, что из-за нелинейности алгоритма, описанного в Рекомендации МСЭ Т Р.862, результат, полученный с использованием объединенных сигналов, не будет соответствовать простому среднему арифметическому результатов для отдельных образцов речи.

7.2 Активная речь

Активность речи в эталонных речевых сигналах, которая может быть измерена на основе Рекомендации МСЭ-Т Р.56³, должна находиться в пределах от 40% до 80%. Эталонные образцы должны содержать не менее 3,2 с активной речи. Это требование, вместе с рекомендованной длиной сигнального файла, должно гарантировать, что в соответствии с Рекомендацией МСЭ-Т Р.862 имеется достаточно речи для точного расчета, и речь должна содержать паузы для тестирования важных элементов сети.

7.3 Временная структура

Эталонный речевой сигнал должен содержать фрагменты активной речи, разделенные периодами молчания, представляющими собой естественные паузы в речи. Большая часть экспериментов, использованных при калибровке и утверждении Рекомендации МСЭ-Т Р.862, содержала пары сигнальных последовательностей, разделенных молчанием. Хорошими примерами являются речевые сигналы, включенные в Дополнение 23 к Рекомендациями⁴ серии Р, которые длятся 8 с и содержат по два коротких предложения, разделенных периодом молчания, продолжительностью, как минимум, 1 с, и в упомянутое выше Дополнение В/Р.501. Рекомендуется, чтобы эталонный речевой сигнал содержал несколько непрерывных фрагментов активной речи, а не множество коротких фрагментов активной речи типа быстрого устного счета⁵.

7.4 Уровень активной речи

Уровень активной речи, упоминаемый в настоящей Рекомендации, это эквивалентный уровень эталонного сигнала, сохраненного в цифровом виде, измеренный в соответствии с Рекомендацией МСЭ-Т Р.56. Уровень активной речи, относящийся к каналу подачи сигнала измерительной системы, описывается отдельно в разделе 9. Рекомендуется, чтобы все файлы эталонных речевых сигналов хранились с уровнем -30 дБов во избежание ограничения пиков. Отметим, что это – уровень речевого сигнала источника, сохраненного в цифровом формате и что уровень входного сигнала для испытываемой системы должен быть определен отдельно в соответствии с целями объективных измерений (см. Рекомендацию МСЭ-Т Р.830)⁶.

7.5 Применение искусственного голоса

Применение сигналов искусственного голоса требует дополнительных исследований с точки зрения языка и временной структуры мощности сигнала, а также других возможных факторов [В.2].

7.6 Требования к записи речи

В Рекомендациях МСЭ-Т Р.800 и Р.830 содержатся указания для записи речевого материала. В настоящей Рекомендации предполагается, что речевой сигнал источника записан в соответствии с этими указаниями. Отметим, что эталонный речевой сигнал может быть тем же, что и на источнике

³ В Рекомендации МСЭ-Т G.191 приводится программное обеспечение, называемое sv56demo.c, которое измеряет коэффициент активной речи и уровень активной речи в соответствии с Рекомендацией МСЭ-Т Р.56.

⁴ Права на копирование Дополнения не разрешают применять эти сигналы в коммерческих приложениях.

⁵ Эталонное программное обеспечение, представленное в виде Дополнения А/Р.862, ограничивает максимальное число фрагментов активной речи значением 50. Если используются эталонные сигналы с большим числом фрагментов активной речи, необходимо удостовериться, что вариант реализации Рекомендации МСЭ-Т Р.862, используемый для испытаний, пригоден для обработки такого большого числа.

⁶ Типовым номинальным значением уровня активной речи является -20 дБм0, соответствующее примерно -26 дБов. В некоторых конкретных системах, которые подвергаются испытаниям, средний уровень активной речи может существенно отличаться от номинального значение -20 дБм0. В таких случаях среднее измеренное значение может использоваться в качестве уровня активной входной речи. Когда оценивается реакция системы на входной уровень, допускается использовать диапазон значений активной речи, например, -14 , -26 и -38 дБов (что приблизительно равно -8 , -20 и -32 дБм0), как рекомендовано в Рекомендации МСЭ-Т Р.830.

речи, или к нему может быть добавлен низкий уровень шума и/или его спектр может быть изменен (см. § 7.9 и § 7.10).

7.7 Смена диктора и изменение содержания речи

Изменения из-за смены диктора или содержания речи могут регулироваться за счет использования определенного набора образцов речи для всех тестов, результаты которых должны сравниваться. Следовательно, для упрощения выполнения последующих сравнений и интерпретации результатов, полученных из различных лабораторий, полезно использовать базы данных речевых сигналов, содержащиеся в Дополнении В/Р.501.

В ситуациях с моделированием сети рекомендуется, чтобы эталонный речевой сигнал включал как минимум голоса двух женщин и двух мужчин, которые произносят различные предложения. Результаты, полученные для этих различных образцов речи в соответствии с Р.862.1, должны быть впоследствии усреднены для всех условий работы в отдельности.

В ситуациях с испытаниями на работающих сетях может использоваться меньшее количество дикторов, однако это может не позволить максимально протестировать сеть. Если такой формат является необходимым, то в этих коротких эталонных сигналах могут быть включены голоса многих дикторов. В случае когда требуется отдельная оценка для каждого образца речи, образцы, содержащие голоса нескольких дикторов, будут уменьшать зависимость полученных результатов от конкретного образца.

В ходе утверждения Рекомендации МСЭ-Т Р.862 имелось очень небольшое число данных относительно детских голосов и некоторых характеристик речи (например, голосовых/речевых декодерах и т. д.). В условиях ограниченности данных с детскими голосами проблем не возникало. С Рекомендацией МСЭ-Т Р.862 музыке использовать нельзя.

Рекомендуется также использовать несколько различных образцов речи (4–10 предложений) для каждого диктора, для того чтобы учесть фонетические изменения.

7.8 Паузы перед началом и после окончания речи

В измерениях по Рекомендации МСЭ-Т Р.862 для выравнивания уровней использован среднеквадратический (RMS) уровень эталонного и ухудшенного сигналов. Если в начало и конец эталонного сигнала включены длительные периоды молчания, то результат выравнивания уровней может оказаться недостоверным.

Рекомендуется, чтобы минимальная длительность паузы перед началом и после окончания речи составляла 0,5 с, поскольку это время требуется измерительному оборудованию для синхронизации ухудшенного речевого сигнала с эталонным.

Рекомендуется, чтобы максимальная длительность паузы перед началом и после окончания речи составляла 2 с, и это было бы полезным при продолжительных задержках в системе.

7.9 Предварительная фильтрация

Эталонный речевой сигнал, подготовленный в соответствии с §§ 7.1–7.8, должен быть отфильтрован так, чтобы учитывались частотные характеристики передачи абонентского устройства. Следует отметить, что в Рекомендации МСЭ-Т Р.862 предполагается, что эталонный речевой сигнал отражает эти электроакустические характеристики соответствующим образом. Если предположить, что эталонный речевой сигнал подается в сети как выходной сигнал абонентского терминала, МСЭ-Т рекомендует использовать модифицированные характеристики передачи IRS, определенные в Дополнении D/Р.830. Такая фильтрация должна выполняться после того, как учтены соответствующим образом §§ 7.1–7.8.

Следует тщательно координировать фильтрацию, используемую с номинальной частотной характеристикой испытываемой системы, поскольку такая фильтрация зависит от того, где эталонный речевой сигнал подается на оборудование и/или от испытываемой сети (см. раздел 9).

7.10 Уровень шума

Уровень шума в эталонном речевом сигнале должен быть достаточно низким, как предполагается в записях, соответствующих Рекомендациям МСЭ-Т Р.800 и Р.830. Кроме того, можно добавить отрывки полной тишины (например, с нулевой цифровой амплитудой), так чтобы эталонные речевые сигналы

имели бы характеристики, определенные в § 7.1, § 7.2, § 7.3 и § 7.8⁷. В этом случае эталонный речевой сигнал соответствует речевому сигналу источника, как описано в § 7.6.

Однако если в канале измерений, описанном в разделе 9, или в самом испытываемом устройстве ожидается нежелательный шум, то небольшой уровень белого шума примерно -75 дБов должен быть специально добавлен к эталонному сигналу, как указано выше, и сохранен в 16-битовом формате линейной ИКМ. Величина уровня шума должна определяться в полосе $0-4000$ Гц⁸. Шум на этом уровне не будет серьезно влиять на результаты, полученные по Рекомендации МСЭ-Т Р.862, но эффективно устранил влияние такого измерительного шума на конечные результаты [В.3]. Достаточно важно добавить такой уровень шума после предварительной фильтрации, описанной в § 7.9.

Уровень активной речи в канале подачи сигнала измерительной системы, описанный в разделе 9, после такой предварительной фильтрации должен быть откалиброван⁹.

Следует отметить, что предлагаемое введение в эталонный сигнал дополнительного шума приведет к получению более точных результатов, если нежелательный шум в приемном канале является непрерывным фоновым шумом, и это не решает проблем, возникающих с появлением комфортного шума, который вводится только в паузах речи.

7.11 Пути реализации

Многие сигналы, включенные в проверку на совместимость по Р.862, не отвечают перечисленным выше требованиям. Это совершенно не имеет значения для проверки на совместимость, поскольку единственной его целью является доказать корректность реализации. Однако необходимо проявлять осторожность, поскольку реализованный алгоритм также выдает результаты в случаях нарушения требований, определенных в настоящей Рекомендации, поскольку в противном случае проверка на совместимость не может быть применена.

8 Характеристики ухудшенного сигнала, который следует оценить

Ухудшенные сигналы являются выходными сигналами испытываемой системы, которые соответствуют входному испытательному сигналу, включая любые влияния измерительного интерфейса. В настоящем разделе описываются характеристики сигналов, которые хранятся в цифровом виде как выходные сигналы испытываемой системы и предназначены для использования

⁷ Передача цифрового молчания в цифровой канал подачи сигнала, например в телефон ЦСИС, а затем в беспроводную телефонную линию может привести к появлению нежелательных явлений, если кодеки GSM или 3GPP используются в беспроводных сетях. А именно, передача постоянного сигнала с наименьшим положительным значением $G.711$ типа А, равным линейному числу $+8$ ($D5$ в 26-битовой ИКМ), приведет к тому, что речевой кодек будет переустанавливаться раз в 20 мс, т. е. 50 раз/с. Это происходит в результате работы встроенной процедуры, возвращающей кодек в исходное состояние для тестирования самого кодека. Хотя переустановка во время пауз в речи для кодека безопасна, измеренное качество может отличаться от результата для нормальной ситуации, когда кодеки не переустанавливаются. Побочные явления таковы, что кодек не будет использовать дискретную передачу во время пауз в речи (DTX), хотя это допускается сетью. Таким образом, комфортный шум или отсечение пиков речи детектором речевой активности (VAD) не будет измерен. Для решения этой проблемы к испытательному образцу до его передачи в сеть должен быть добавлен фоновый шум с низким уровнем порядка -65 дБм0. Это ломает шаблон постоянного цифрового молчания.

⁸ Когда в эталонном речевом сигнале используется частота дискретизации 16 кГц, необходимо тщательно определять уровень фонового шума.

⁹ Фоновый шум может не вводиться в испытываемую систему, потому что уровень фонового шума становится меньше минимально возможного уровня системы. Например, наименьшие значения для кодирования типа А составляют ± 8 в линейных 16-битовых значениях. Таким образом, наименьший уровень составит -72 дБов. Это значит, что, если входной уровень системы откалиброван по значению, например -30 дБов, то фоновый шум на уровне -75 дБов не может пройти через систему и не решает проблемы нежелательного шума. Если требуется провести испытания при существенно более высоких уровнях, чем номинальный уровень активной речи -26 дБов, и если к эталонному отрывку добавляется фоновый шум -75 дБов, то при помощи Рекомендации МСЭ-Т Р.862 невозможно оценить возможные ухудшения (например, щелчки, шумовые импульсы и т. д.) в паузах речи, поскольку более высокий уровень фонового шума в ухудшенном отрывке может маскировать ухудшения на меньшем уровне. Однако этой проблемы не существует, если использовать номинальный уровень активной речи, равный примерно -26 дБов.

вычислениях на основании Рекомендации МСЭ-Т Р.862. Характеристики канала извлечения сигнала рассматриваются отдельно в разделе 9.

8.1 Различия в продолжительности активной речи между эталонным речевым сигналом и ухудшенным речевым сигналом

Продолжительность активной речи определяется в Рекомендации МСЭ-Т Р.56.

В Рекомендации МСЭ-Т Р.862 для выравнивания уровней использованы RMS уровни эталонного и ухудшенного сигналов. Это означает, что данный алгоритм может выдавать ошибочные результаты, если речевой сигнал пропускается или если периоды тишины добавляются к ухудшенному сигналу или удаляются из него.

Когда из ухудшенного сигнала удаляются фрагменты активной речи или если глушится один или несколько больших отрезков ухудшенного сигнала, этот сигнал будет сдвинут по уровню до значения, превышающего реальное значение.

Когда из ухудшенного сигнала удаляются периоды молчания, этот сигнал будет сдвинут по уровню до значения ниже реального значения.

Это влияет на величину искажений в ухудшенном сигнале и, следовательно, повлияет на результат объективного измерения качества. Если продолжительность речи в эталонном и ухудшенном сигнале отличается более, чем на 25%, эффект может оказаться достаточно большим для того, чтобы заметно исказить результат. Это особенно заметно, если молчанием заменяются продолжительные непрерывные отрезки речи.

8.2 Уровень активной речи

Уровень активной речи определяется в Рекомендации МСЭ-Т Р.56.

Хотя уровень активной речи нормализован в вычислениях значений PESQ, рекомендуется, чтобы цифровой уровень речи, сохраняемый как ухудшенные сигналы для алгоритма PESQ, составлял бы примерно -30 дБов во избежание отсека пиков сигнала и искажений квантования. Следует отметить, что Рекомендация МСЭ-Т Р.862 не может быть использована для оценки влияния уровня приема/прослушивания¹⁰.

8.3 Различия в продолжительности пауз перед началом и после окончания речи между эталонным речевым сигналом и ухудшенным речевым сигналом

В Рекомендации МСЭ-Т Р.862 для выравнивания уровней использованы среднеквадратические (RMS) уровни эталонного и ухудшенного сигналов. Если в начале и в конце ухудшенного сигнала включены продолжительные паузы, то процесс выравнивания уровней может быть субоптимальным. Эта ситуация может превратиться в проблему, если продолжительность эталонного и ухудшенного сигналов отличается более чем на 20%¹¹.

Кроме того, в Рекомендации МСЭ-Т Р.862 не учитываются никакие искажения ухудшенного сигнала, возникшие до начала или после окончания сигнала активной речи. Этот сигнал активной речи определяется из эталонного сигнала как первая и последняя точки, в которых уровень сигнала достигает значения примерно 50 дБ SPL.

9 Характеристики каналов подачи и извлечения сигнала

В настоящем разделе описываются желаемые характеристики каналов подачи и извлечения сигнала при аппаратных измерениях. Схема и условия измерений могут повлиять на результаты, полученные

¹⁰ Если испытывается MNRU, соответствующая Рекомендации МСЭ-Т Р.810, следует осторожно выполнять выравнивание уровней для сохранения реальных уровней речи, исключая шум, добавленный при помощи MNRU.

¹¹ Эмпирические наблюдения предполагают, что результаты, полученные при помощи Рекомендации МСЭ-Т Р.862 для EVRC [B.4], зависят от конкретного выравнивания границ кадра кодирования с входными ИКМ данными. Результат может меняться до 0,25 в зависимости от того, где кончаются границы кадра. В случае EVRC методом получения стабильных результатов будет измерение каждого из 80 возможных вариантов выравнивания и усреднение результатов. Аналогичные ситуации могут встречаться в процессах DSP.

по Рекомендации МСЭ-Т Р.862, если только не были предприняты меры предосторожности для управления факторами, вовлеченными в процесс. Шумы и помехи должны быть в максимально возможной степени исключены из каналов подачи и извлечения сигналов для гарантии того, что они не влияют на результаты.

9.1 Влияние измерительных линий связи и испытательных конфигураций в канале подачи сигнала

Если имеется вероятность использования известного интерфейса, например POTS или ISDN, то этот метод является предпочтительным, и испытательное оборудование должно быть откалибровано так, чтобы работать с этим интерфейсом с рекомендованными номинальными уровнями сигнала.

Если ни один из известных интерфейсов не может быть использован, то точкой подачи сигнала часто является порт оконечного устройства, который является патентованным интерфейсом, и требуемый входной уровень изначально неизвестен. Несмотря на то что стандарты (например, североамериканский стандарт Т1А-810-А) определяют характеристики терминала между акустическим и сетевым интерфейсами, они не определяют промежуточных точек, например, интерфейса портативного терминала. Распределение усиления и фильтрация будут различными для разных производителей и даже для отдельных терминалов. В ряде случаев эти характеристики могут меняться на конечном устройстве. Когда порт носимого терминала используется в качестве испытательного порта по Рекомендации МСЭ-Т Р.862, инженер-испытатель может измерять:

- 1) качественные характеристики оконечного устройства и сети вместе;
- 2) качественные характеристики самого оконечного устройства (соединенного с определенной эталонной сетью); или
- 3) качественные характеристики самой сети с минимальным увеличением искажений от оконечного устройства.

Однако во всех случаях мы стремимся исключить влияние измерительной установки.

Инженер-испытатель должен гарантировать, что уровень активной речи, прикладываемой к патентованному интерфейсу, соответствует желаемому уровню в сети и динамическому диапазону кодека. Это требует соответствующего распределения усиления между точкой подачи сигнала и точкой интерфейса терминала, и сетью в обоих направлениях передачи.

Подавая речевые сигналы на патентованный интерфейс, инженер должен быть проинформирован о фильтрации (например, модифицированном IRS и выравнивании частот приемопередатчика) между акустическим интерфейсом и интерфейсом с сетью. Продавцы терминалов могут реализовывать любую комбинацию акустической, электронной или цифровой фильтрации на любой стороне патентованного интерфейса портативной станции. Следовательно, испытательное оборудование Р.862 может наблюдать либо полную, либо частичную фильтрацию, либо полное отсутствие фильтрации после точки подачи сигнала. Для получения точного результата измеряемая конфигурация должна включать фильтрацию, соответствующую наблюдаемому испытанию. Точно так же фильтрация, применяемая к эталонному сигналу, должна совпадать с фильтрацией, применяемой к сквозной испытательной цепи.

В настоящем параграфе объясняется идеальный метод, который может использоваться для определения входных характеристик для эталонного сигнала, когда существует возможность того, что входной фильтр, используемый при обычной работе терминального оборудования связи, не находится в измерительном канале. Искусственный голос (например, расположенный в модели головы и туловища) должен использоваться для акустической подачи испытательного сигнала на терминал, который должен быть соединен с удаленной контрольной точкой (например, точкой ЦСИС). Акустический уровень должен представлять собой нормальное использование терминального устройства, и уровень фонового шума должен быть менее 35 дБА. Такое нормальное использование может отражать применение либо внутреннего микрофона, либо персонального комплекта "hands-free" и зависит от цели оцениваемой ситуации. Искусственный голос должен быть откалиброван, и положение терминала должно представлять обычное использование. Электрический уровень и значение частоты должны измеряться в контрольной точке в сетевом соединении (например, оконечная точка ЦСИС). Этот процесс должен затем повторяться (с тем же испытательным сигналом) с использованием электрического входного сигнала в точке подачи

сигнала Р.862 и применением оборудования, используемого в процессе испытаний Р.862. Входной сигнал теперь должен быть отрегулирован так, чтобы его электрический уровень и частота соответствовали значениям, полученным во время акустической подачи сигнала. Описанный здесь метод является идеальным и может быть аппроксимирован для множества ситуаций¹². Если этот метод не используется, то рекомендуется, чтобы испытатель учитывал особые пожелания в спецификациях производителя по акустическим и электрическим интерфейсам с терминалами связи.

9.2 Влияние измерительных линий связи и испытательных конфигураций в канале извлечения сигнала

После того как эталонный речевой сигнал прошел через испытываемую систему, из точки извлечения сигнала он должен быть передан на Р.862. Этот канал извлечения сигнала вносит шумы и искажения, которые могут повлиять на результат. Канал извлечения сигнала может содержать нежелательные сигналы контуров заземления, адаптеров электрических проводников переменного тока или других синфазных сигналов. Внутриполосные сигналы адаптеров могут искажать результаты. Кроме того, внеполосный шум с достаточно высокими уровнями может создавать помехи в полосе измерений, если используется недостаточная фильтрация подавления помех.

Для минимизации шумов, вносимых каналами подачи и извлечения сигнала, рекомендуется, чтобы каналы подачи и извлечения сигнала совместно вносили шум с уровнем не более -70 дБov, так, чтобы результирующее значение SNR становилось равным 40 дБ, и результат объективного измерения качества определяется исключительно влиянием испытываемой системы.

Как правило, медленно меняющиеся скорости дискретизации, растяжение во времени и компрессия во времени передаваемого сигнала могут привести к получению чересчур пессимистичных значений из-за неправильной синхронизации во времени.

Если используется аналоговая передача, необходимо проявлять осторожность, чтобы не возникало чересчур больших сдвижек между конверторами AD и DA. Это может произойти с оборудованием пользователя, особенно если оборудование не поддерживает требуемую частоту дискретизации и используется программное преобразование с использованием драйвера звуковой карты.

10 Анализ результатов

10.1 Усреднение результатов измерений

Как отмечено в § 7.7, при выполнении объективных измерений следует использовать голоса, как минимум, двух женщин и двух мужчин. Перед вычислением средних значений или иных статистических показателей отдельные результаты измерений должны быть сначала переведены в область MOS-LQO (в соответствии с Рекомендацией МСЭ-Т Р.862.1) и затем усреднены по дикторам и образцам речи. Поскольку алгоритм, определенный в Рекомендации МСЭ-Т Р.862, не линеен, результаты, полученные из объединенных сигналов, не будут совпадать с усредненными значениями результатов, полученных для каждого сигнала в отдельности.

Как отмечено в § 6.2, имеется два типа приложений Р.862, которые требуют различных подходов к анализу. В первом случае усреднение по дикторам и образцам речи должно быть выполнено до продолжения анализа. Такой анализ пригоден для управляемого моделирования сетей с точно повторяемыми результатами. При выполнении измерений на реально работающих сетях требуется оценка качества для каждого образца в отдельности из-за неуправляемости изменяющихся во времени каналов передачи.

10.2 Надежность результатов измерений PESQ

Для испытаний, проверок и калибровки (Р.862.1) по алгоритму Р.862 используется большое количество баз данных. Например, таких, как описанные в Рекомендациях МСЭ-Т Р.862 и Р.862.1 базы данных, содержащие образцы речевых сигналов, наговоренные различными дикторами –

¹² Такая аппроксимация зависит от измерительного оборудования. Один из возможных методов описан в Приложении IV.

мужчинами/женщинами, на различных языках, и содержащие отрывки, качество которых было ухудшено как искусственным моделированием, так и в естественных условиях работы сети. Кроме того, учтены условия работы сетей в фиксированных, беспроводных и VoIP приложениях. Более подробно эти базы данных испытательных сигналов описаны в Приложении II.

Следует отметить, что результаты измерений по P.862/P.862.1 имеют надежность 95% и их точность известна и управляема, когда этот алгоритм используется для того же типа приложений, для работы с которым он был предварительно отлажен, проверен и одобрен. Другими словами, сценарии измерений должны представлять собой статистически тот же тип совокупности выборки, для которого алгоритм P.862/P.862.1 был предварительно отлажен, проверен и откалиброван, для того чтобы заявленные значения точности оставались действительными. Если же этот алгоритм используется для оценки качества речи в новых технологиях и/или с использованием других типов кодеков и/или в новых работающих сетях, точность и надежность результатов становится неизвестной и непредсказуемой.

10.3 Показатели точности измерений PESQ

Для оценки качественных показателей с использованием алгоритмов P.862/P.862.1 на базах данных, описанных в § 10.2, используются три статистических параметра – коэффициент корреляции, погрешность предсказания и распределение остаточной ошибки. Как сказано в § 10.1, способы анализа зависят от типа приложения, управляемого моделирования сети и условий испытаний на работающих/действующих сетях.

Для всех условий моделирования сетей для расчета статистических параметров использовались значения, полученные при оценке речевых сигналов, наговоренных как минимум двумя мужчинами и двумя женщинами, усредненные по типам условий работы. Для баз данных, составленных по работающим сетям, статистические параметры рассчитывались с использованием объективных и субъективных оценок для каждого образца речи.

Результаты оценки качественных показателей представлены в таблицах 1 и 2. Рассчитаны также 95-процентные критические пределы доверительных интервалов для значений коэффициента корреляции и погрешности предсказания для того, чтобы определить нижнюю границу 95% корреляции и верхнюю границу 95% погрешности предсказания.

Результаты приведены по типам приложений (например, моделированные условия работы беспроводной сети и сети VoIP и реальные условия работы беспроводной сети и сети VoIP). Следовательно, эти значения точности выражают качественные показатели алгоритма PESQ, если он используется для любого из приложений, упомянутых в § 10.2.

Таблица 1/P.862.3 – Доверительные интервалы для коэффициента корреляции и предсказания ошибок

Приложение	N	Параметр	P.862 (необработанные результаты PESQ)	P.862.1 (калиброванные результаты PESQ)
Данные при условии моделирования сетей (беспроводная связь, VoIP и фиксированные приложения)	1 357	R	0,956	0,956
		CI95%-нижний предел	0,940	0,940
		PE	N.A.	N.A.
		CI95%-верхний предел	N.A.	N.A.
Данные реальных измерений (Беспроводные приложения: GSM US и EU, CDMA-US, TDMA-US, iDEN-US, AMPS-US; и VoIP приложение)	1 135	R	0,925	0,926
		CI95%-нижний предел	0,916	0,917
		PE	0,479	0,462
		CI95%-верхний предел	0,492	0,475

Таблица 2/Р.862.3 – Распределение остаточных ошибок

Приложение	MOS буферы	<0,25	<0,5	<0,75	<1	<1,25	<1,5	<1,75	<2
Данные реальных измерений (Беспроводные приложения: GSM US и EU, CDMA-US, TDMA-US, iDEN-US, AMPS-US; и VoIP приложение)	Р.862 CDF (%)	32,51	66,52	90,84	97,97	99,38	99,91	99,91	100
	Р.862 вероятность (%)	32,51	34,09	24,32	7,14	1,41	0,53	0	0,09
	Р.862.1 CDF (%)	40,44	70,48	90,33	97,71	99,3	99,7	99,91	100
	Р.862.1 вероятность (%)	40,44	30,04	19,82	7,4	1,59	0,44	0,18	0,09

10.4 Интерпретация результатов измерения точности

По определению, алгоритм Р.862/Р.862.1 является механизмом оценки субъективного мнения о качестве речи, обеспечиваемого испытываемой сетью. Следовательно, следует отметить, что результаты любых измерений качества речи, выполняемых по алгоритму PESQ, будут зависеть от точности значений, представленных в таблицах 1 и 2.

Следует отметить, что, как указано в § 10.2, значения точности сохраняются действительными до тех пор, пока в сценариях проведения измерений представлены статистически те же наборы отрывков речевых сигналов, которые представлены в § 10.2.

Нижняя граница 95-процентного доверительного интервала коэффициента корреляции показывает, что, как ожидается, измерения по Р.862/Р.862.1 продемонстрируют корреляцию с субъективным мнением, которое больше или, как минимум, равно нижнему пределу 95-процентного доверительного интервала коэффициента корреляции, вне зависимости от того, использовалось ли моделирование или испытания на работающих сетях, и вне зависимости от типа испытываемой сети (такой как, беспроводная, фиксированная, VoIP) (таблица 1).

Распределение остаточной ошибки (таблица 2) представляет собой кумулятивную функцию плотности (CDF) абсолютных ошибок между MOS и результатами Р.862/Р.862.1, и она показывает вероятность того, что абсолютная ошибка ниже, чем некоторое значение. Например, вероятность того, что абсолютная ошибка ниже чем 0,5 MOS, превышает 70%, тогда как вероятность того, что ошибка ниже чем 0,75 MOS, превышает 90%. В таблице 2 также показана функция плотности вероятности (PDF) абсолютной ошибки. Как ожидается, в соответствии с CDF функция PDF показывает, что меньшие абсолютные ошибки имеют более высокую вероятность появления, чем более высокие значения.

11 Описание результатов

Как сказано в § 10.4, в зависимости от типа приложения, т. е. проводились ли измерения в процессе моделирования или на работающей сети, результаты измерений Р.862/Р.862.1 должны представляться с учетом точности алгоритма, указанной в § 10.3 (таблицы 1 и 2).

Рекомендуется использовать коэффициент корреляции как информативный статистический параметр качества Р.862/Р.862.1 для конкретного приложения. Рекомендуется использовать погрешность предсказания вместе с распределением остаточной ошибки для составления отчета о результатах измерений Р.862.1 для конкретного приложения.

Как правило, должны быть указаны усредненные максимальные и минимальные значения PESQ, а также количество измерений, использованных для расчета среднего значения. Некоторые подробные указания по составлению отчетов об измерениях PESQ приведены в Приложении III. Кроме того, количество измерений, по итогам которых был получен данный результат PESQ или некий диапазон результатов, может быть показано графически в виде частотного распределения. В тех случаях, когда испытываемая система показывает относительно стабильное качество прослушивания, для принятия решения о необходимости проведения дальнейших измерений с целью достижения требуемой

точности может использоваться стандартная девиация. Этот подход не пригоден для испытываемых систем, параметры которых значительно меняются во времени (например, сети VoIP или сети подвижной связи).

12 Руководство по применению широкополосного дополнения Р.862.2 к Рекомендации Р.862

В принципе, указания, содержащиеся в настоящей Рекомендации, применимы как для Рекомендации МСЭ-Т Р.862, так и для широкополосного расширения Р.862.2. Тем не менее для широкополосного расширения Рекомендации МСЭ-Т Р.862 требуются некоторые дополнительные указания.

Приведенные далее указания главным образом относятся к применению характеристик передачи IRS, которые должны применяться к входному или эталонному сигналу. Для широкополосного расширения не рекомендуется использовать фильтрации ни для речевого сигнала, ни для какого-либо окружающего шума. Это указано в § 6.2, § 7.9 и § 7.10.

Что касается вычисления уровня речевой активности в соответствии с Рекомендацией МСЭ-Т Р.56, рекомендуется использовать широкополосный вариант Р.56. Это указано в § 7.2, § 7.4, § 8.1 и § 8.2.

Предлагаемое в § 7.10 добавление низкоуровневого шума для широкополосного расширения не рассматривалось и не может быть рекомендовано.

В разделах 10 и 11 описывается точность метода Р.862. Приведенные цифры применимы только к узкополосным приложениям Р.862.

В § 3.6 описывается уровень 0 дБм0 в соответствии с Рекомендацией МСЭ-Т G.711. Необходимо отметить, что эта ссылка на G.711 применима только к узкополосным приложениям.

Оба метода – Р.862.1 и Р.862.2 – относятся к шкале MOS-LQO, определенной в Рекомендации МСЭ-Т Р.800.1. Необходимо учитывать, что термин MOS-LQO в будущем может быть распространен на узкополосные и широкополосные случаи. Отметим, что результаты, полученные по Рекомендации МСЭ-Т Р.862.1, относятся только к узкополосному варианту. Результаты, полученные по Рекомендации МСЭ-Т Р.862.2, применимы к широкополосным или смешанным широкополосным и узкополосным приложениям. В результате прямое сравнение результатов MOS-LQO по Р.862.1 и Р.862.2 не является возможным.

Приложение I

Контрольные показатели объективного качества, полученные в Рекомендации МСЭ-Т Р.862 для стандартных кодеков МСЭ-Т/GSM

I.1 Контрольные показатели в соответствии с Рекомендацией МСЭ-Т Р.862.1, рассчитанные для следующих условий применения кодека/MNRU с использованием базы данных речевых сигналов из Дополнения В/Р.501:

- G.711 типа μ и типа А;
- G.726 на скоростях 16, 24, 32 и 40 кбит/с;
- G.728;
- G.729;
- Дополнение А/G.729;
- G.723.1 при 5,3 и 6,3 кбит/с;
- GSM-AMR при 4,75, 5,15, 5,9, 6,7, 7,4, 7,95, 10,2 и 12,2 кбит/с;
- GSM-EFR;
- GSM-FR;
- GSM-HR;
- MNRU ($Q = 5, 10, 15, 20, 25, 30, 35, 40$ и 45 дБ).

На рисунке I.1 показана предварительная обработка в соответствии с эталонными речевыми сигналами. На рисунках I.2, I.3 и I.4 показаны процедуры кодирования в кодеках G.711, G.726 и других кодеках соответственно. На рисунке I.5 показана процедура обработки MNRU. В таблицах I.1, I.2 и I.3 приведены контрольные значения, которые были получены путем преобразования необработанных результатов значений PESQ¹³ в значения MOS-LQO с использованием Рекомендации МСЭ-Т Р.862.1 для стандартных кодеков МСЭ-Т и MNRU. Обработка сигнала в соответствии с G.711, G.726 и MNRU была выполнена с использованием программного обеспечения, содержащегося в Рекомендации МСЭ-Т G.191.

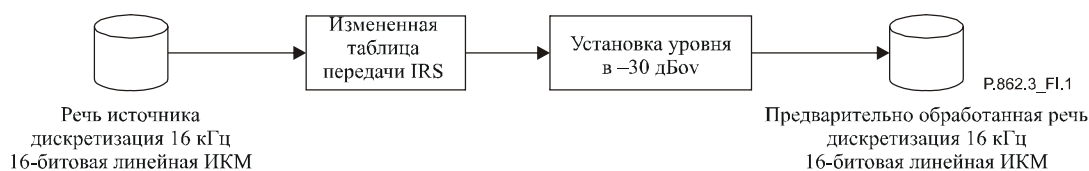


Рисунок I.1/Р.862.3 – Предварительная обработка исходного речевого сигнала

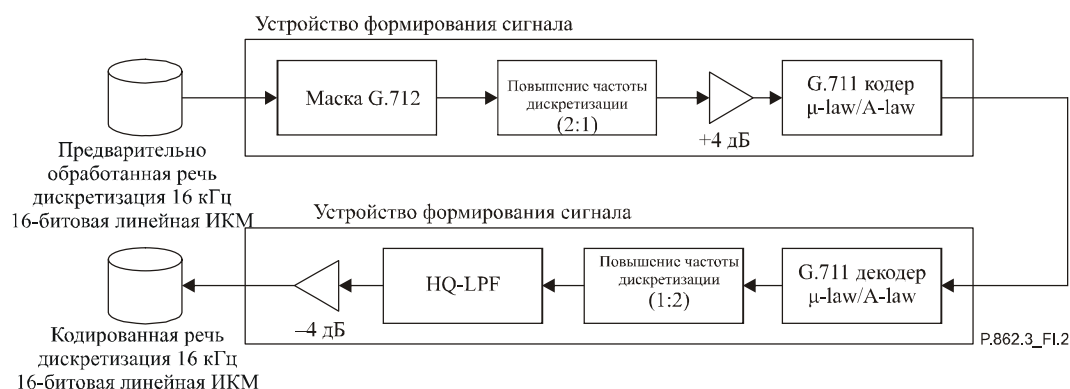


Рисунок I.2/Р.862.3 – Обработка в соответствии с G.711

¹³ В кодеках GSM-AMR, GSM-EFR, GSM-FR, GSM-HR и G.711 типа А необработанные значения PESQ были рассчитаны при частоте дискретизации 16 кГц.

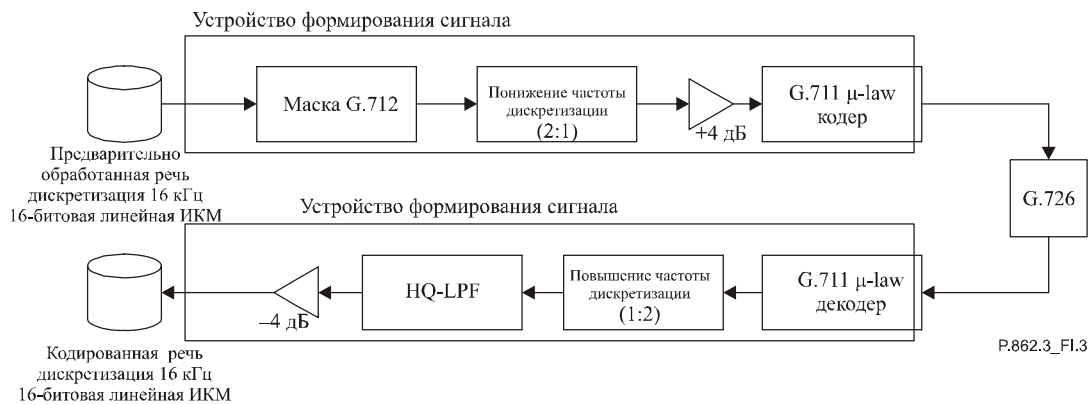


Рисунок I.3/P.862.3 – Обработка в соответствии с G.726

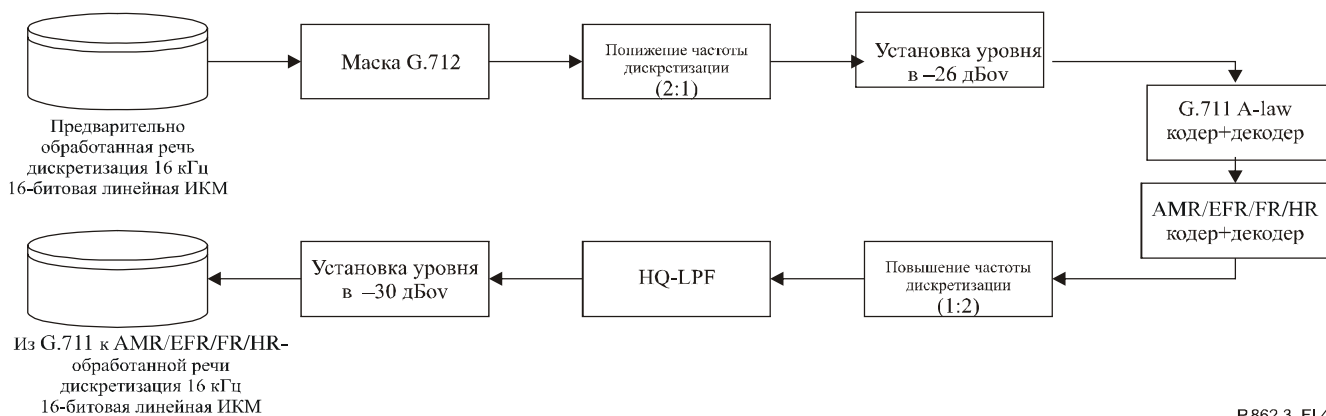
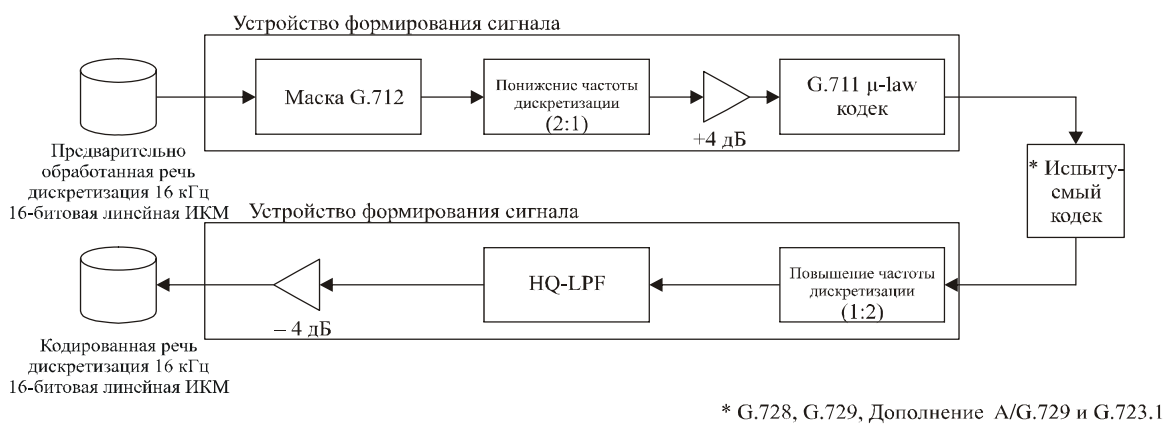


Рисунок I.4/P.862.3 – Обработка в соответствии с G.728, G.729, Дополнением A/G.729 G.723.1 и GSM-AMR/EFR/FR/HR

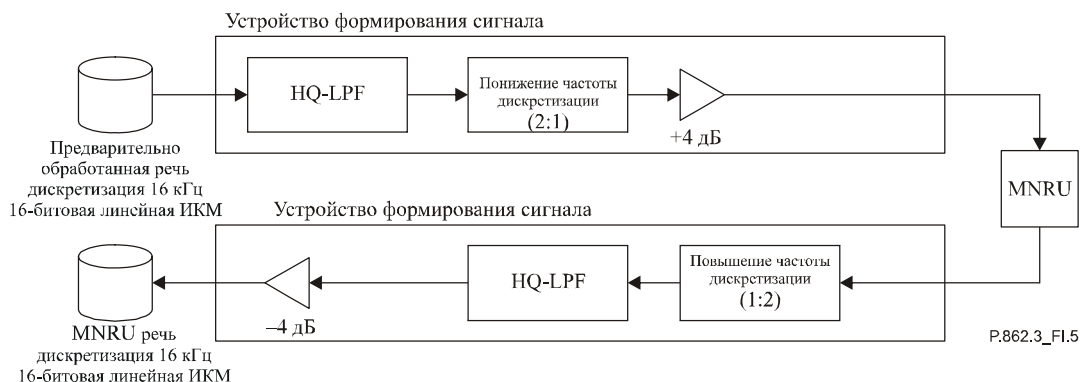


Рисунок I.5/P.862.3 – Обработка в соответствии с MNRU

Обработка сигнала, показанная на рисунках I.1–I.5, описывается далее.

I.2 Предварительная обработка исходного речевого сигнала

На этапе предварительной обработки речевого сигнала источника проходит через модифицированный IRS фильтр на стороне передатчика, и его уровень приводится к значению -30 дБов, как показано на рисунке I.1. Предварительно обработанные речевые файлы являются входными сигналами для устройства формирования сигнала. С использованием программы, представленной в Рекомендации МСЭ-Т G.191, которая далее называется STL2000, предварительная обработка речи может быть выполнена посредством следующих команд:

```
$ filter -q mod IRS16 file.inp file.irs
$ sv56demo -q file.irs file.pre 256 1 0 -30
```

I.3 Обработка в соответствии с G.711

Предварительно обработанная речь маскируется в соответствии с Рекомендацией МСЭ-Т G.712 и передискретизируется с понижением частоты дискретизации. Речевой сигнал с пониженной частотой дискретизации затем доводится до уровня -26 дБов, и полученный выходной сигнал становится входным сигналом для кодера G.711. Команды обработки приведены ниже.

```
$ filter -q -down PCM file.pre tmpfile1
$ sv56demo -q tmpfile1 g711.inp 256 1 0 -26 8000
```

Обработка кодирования и декодирования в кодеке по Рекомендации МСЭ-Т G.711 типа μ выполняется с использованием следующей команды STL2000:

```
$ g711demo u lili g711.inp g711.dec
```

Декодированная речь передискретизируется с повышением частоты дискретизации от 8 до 16 кГц, и уровень сигнала доводится до -30 дБов. Речевой сигнал, кодированный по G.711, может быть получен при помощи следующих команд STL2000:

```
$ filter -q -up HQ2 g711.dec tmpfile2
$ scaledemo -q -gain 0.63095 tmpfile2 g711.out
```

I.4 Обработка в соответствии с G.726

В ходе обработки по G.726 выходной сигнал кодера G.711 типа μ становится входным сигналом кодера G.726, и его выходной сигнал становится входным сигналом декодера G.711 типа μ . Речевой сигнал, кодированный по G.726, может быть получен при помощи следующих команд STL2000:

```
$ filter -q -down PCM infile tmpfile1
$ sv56demo -q tmpfile1 tmpfile2 256 1 0 -26
$ g711demo u lilo tmpfile2 tmpfile3
$ g726demo u lolo {40/32/24/16} tmpfile3 tmpfile4
```

```
$ g711demo u loli tmpfile4 tmpfile5
$ filter -q -up HQ2 tmpfile5 tmpfile6
$ scaledemo -q -gain 0.63095 tmpfile6 outfile
```

I.5 Обработка в соответствии с G.728, G.729, Дополнением A/G.729 и G.723.1

В этих кодеках входной сигнал испытываемого кодека является выходным сигналом кодека G.711 типа μ , т. е. сигналом, который на рисунке I.2 обозначен как *g711.dec*. Когда выходные сигналы испытываемого кодека обозначаются как {g728|g729|g729A|G7231}.dec, кодированный речевой сигнал получается при помощи следующих команд STL2000.

```
$ filter -q -up HQ2 {g728|g729|g729A|g7231}.dec tmpfile2
$ scaledemo -q -gain 0.63095 tmpfile2 {g728|g729|g729A|g7231}.out
```

I.6 Обработка в соответствии с MNRU

Предварительно обработанный речевой сигнал передискретизируется с понижением частоты дискретизации без изменения частотной характеристики входного сигнала, его уровень доводится до -26 дБов, и выходной сигнал становится входным сигналом для MNRU. Выходной сигнал MNRU передискретизируется с повышением частоты дискретизации от 8 до 16 кГц, и уровень сигнала доводится до -30 дБов. Речевой сигнал MNRU с уровнем Q -дБ может быть получен при помощи следующих команд STL2000:

```
$ filter -q -down HQ2 infile tmpfile1
$ sv56demo -q tmpfile1 tmpfile2 256 1 0 -26 8000
$ mnrudemo tmpfile2 tmpfile3 128 1 0 Q
$ filter -q -up HQ2 tmpfile3 tmpfile4
$ scaldemo -q -gain 0.63095 tmpfile3 mnruQ.out
```

Таблица I.1/Р.862.3 – Эталонные значения по Рекомендации Р.862.1 для стандартных кодеков МСЭ-Т

Язык	Название файла	G.711		G.726				G.728	G.729	G.729A	G.723.1	
		Типа μ	Типа А	16 кбит/с	24 кбит/с	32 кбит/с	40 кбит/с				5.3 кбит/с	6.3 кбит/с
Американский английский	Female 1 (0.00- 7.97 s).wav	4,46	4,28	2,50	3,34	3,89	4,18	3,95	3,95	3,80	3,65	3,81
	Female 2 (0.00- 8.06 s).wav	4,45	4,42	3,12	3,76	4,07	4,33	4,27	4,08	3,99	3,67	3,80
	Male 1 (0.00- 8.44 s).wav	4,49	4,36	2,86	3,82	4,25	4,35	4,19	4,17	4,13	3,90	3,97
	Male 2 (0.00- 7.96 s).wav	4,47	4,31	2,97	3,80	4,21	4,40	4,22	4,15	4,06	3,95	4,07
	усредненное значение	4,47	4,34	2,86	3,69	4,11	4,32	4,16	4,09	4,00	3,80	3,92
Китайский	Female 1 (0.00-10.87 s).wav	4,46	4,48	2,38	3,42	4,21	4,34	3,86	3,72	3,65	3,10	3,33
	Female 1b (0.00-13.39 s).wav	4,42	4,43	2,29	3,34	4,07	4,26	3,98	3,80	3,75	3,26	3,49
	Female 2 (0.00-13.32 s).wav	4,50	4,50	2,26	3,29	4,02	4,37	4,06	3,88	3,75	3,33	3,53
	Female 2b (0.00-13.39 s).wav	4,50	4,50	2,38	3,30	4,03	4,33	4,03	3,87	3,72	3,39	3,59
	Male 1 (0.00-12.15 s).wav	4,44	4,48	2,64	3,58	4,23	4,28	4,19	3,83	3,75	3,36	3,51
	Male 1a (0.00-12.91 s).wav	4,52	4,51	2,84	3,77	4,21	4,37	4,18	4,06	4,00	3,65	3,89
	Male 2 (0.00-12.50 s).wav	4,49	4,48	2,74	3,74	4,30	4,44	4,18	3,99	3,89	3,62	3,78
	Male 2b (0.00-12.82 s).wav	4,50	4,40	2,89	3,90	4,29	4,43	4,16	3,95	3,89	3,35	3,55
усредненное значение	4,48	4,47	2,55	3,55	4,18	4,35	4,08	3,89	3,80	3,38	3,59	
Английский	Female 1 (0.00- 8.00 s).wav	4,50	4,49	2,69	3,37	3,88	4,21	3,89	3,72	3,58	3,42	3,59
	Female 2 (0.00- 8.00 s).wav	4,48	4,46	2,81	3,44	3,92	4,24	4,00	3,91	3,80	3,60	3,67
	Male 1 (0.00- 8.00 s).wav	4,50	4,45	3,03	3,53	3,96	3,99	4,08	3,88	3,88	3,67	3,81
	Male 2 (0.00- 8.00 s).wav	4,51	4,48	2,94	3,79	4,24	4,34	4,05	3,73	3,60	3,64	3,83
	усредненное значение	4,50	4,47	2,87	3,54	4,01	4,20	4,01	3,81	3,72	3,59	3,73
Французский	Female 1 (0.00-10.04 s).wav	4,50	4,47	3,06	3,84	4,28	4,42	4,21	3,85	3,77	3,59	3,69
	Female 2 (0.00-10.04 s).wav	4,51	4,48	2,76	3,64	4,15	4,41	4,03	3,77	3,64	3,39	3,56
	Male 1 (0.00-12.18 s).wav	4,50	4,46	3,09	3,79	4,18	4,32	4,09	3,84	3,82	3,45	3,60
	Male 2 (0.00-10.04 s).wav	4,52	4,48	3,33	3,92	4,28	4,40	4,25	4,00	3,91	3,70	3,88
	усредненное значение	4,51	4,47	3,06	3,80	4,22	4,39	4,15	3,87	3,79	3,54	3,69
Немецкий	female1 (0.00- 8.00 s).wav	4,49	4,48	2,68	3,46	4,02	4,27	4,04	3,86	3,69	3,54	3,75
	female2 (0.00- 8.00 s).wav	4,48	4,46	2,84	3,65	4,24	4,40	4,13	4,07	3,89	3,61	3,82
	male1 (0.00- 8.00 s).wav	4,50	4,47	2,99	3,72	4,27	4,41	4,09	3,95	3,87	3,56	3,84
	male2 (0.00- 8.00 s).wav	4,50	4,46	2,86	3,47	4,03	4,35	4,12	4,07	4,01	3,75	3,91
	усредненное значение	4,50	4,47	2,84	3,58	4,14	4,36	4,09	3,99	3,87	3,62	3,83
Итальянский	Female 1 (0.00-20.60 s).wav	4,49	4,40	2,38	3,23	3,81	4,25	3,80	3,75	3,62	3,28	3,49
	Female 2 (0.00-21.78 s).wav	4,48	4,39	2,72	3,68	4,14	4,34	4,16	3,95	3,87	3,57	3,75
	Male 1 (0.00-18.13 s).wav	4,50	4,44	2,61	3,50	4,01	4,33	4,12	3,88	3,81	3,59	3,75
	Male 2 (0.00-20.86 s).wav	4,51	4,43	3,05	3,94	4,28	4,41	4,18	4,12	4,05	3,73	3,95
	усредненное значение	4,49	4,41	2,69	3,60	4,07	4,33	4,07	3,93	3,84	3,55	3,74
Японский	Female 1 (0.00- 7.60 s).wav	4,46	4,36	2,22	2,97	3,65	4,11	3,76	3,70	3,61	3,25	3,40
	Female 2 (0.00- 7.31 s).wav	4,48	4,38	2,45	3,41	4,12	4,40	3,96	3,82	3,73	3,34	3,59
	Male 1 (0.00- 7.13 s).wav	4,47	4,39	2,36	3,06	3,61	4,20	3,89	3,88	3,74	3,38	3,53
	Male 2 (0.00- 7.45 s).wav	4,49	4,42	2,95	3,90	4,32	4,45	4,31	4,08	4,00	3,83	3,93
	усредненное значение	4,47	4,39	2,49	3,35	3,95	4,30	4,00	3,87	3,78	3,46	3,62
Испанский (США)	Female 1 (0.00- 8.00 s).wav	4,47	4,40	2,33	3,05	3,73	4,17	4,02	3,84	3,71	3,42	3,64
	Female 2 (0.00- 8.00 s).wav	4,40	4,31	2,30	2,92	3,48	4,04	3,77	3,65	3,43	3,07	3,22
	Male 1 (0.00- 8.00 s).wav	4,46	4,30	2,86	3,64	4,19	4,36	4,05	3,83	3,76	3,69	3,77
	Male 2 (0.00- 8.00 s).wav	4,49	4,42	2,76	3,72	4,22	4,40	4,09	3,86	3,85	3,60	3,72
	усредненное значение	4,46	4,36	2,56	3,34	3,93	4,25	3,99	3,80	3,69	3,45	3,60

ПРИМЕЧАНИЕ. – Ячейки серого цвета показывают речевые сигналы, не соответствующие требованиям Рекомендации МСЭ-Т Р.501.

Таблица I.2/Р.862.3 – Эталонные значения по Рекомендации Р.862.1 для MNRU

Язык	Название файла	MNRU								
		5 дБ	10 дБ	15 дБ	20 дБ	25 дБ	30 дБ	35 дБ	40 дБ	45дВ
Американский английский	Female 1 (0.00- 7.97 s).wav	1,80	2,39	3,03	3,68	4,15	4,37	4,47	4,50	4,52
	Female 2 (0.00- 8.06 s).wav	2,09	2,67	3,24	3,80	4,16	4,35	4,41	4,44	4,44
	Male 1 (0.00- 8.44 s).wav	1,93	2,58	3,30	4,01	4,32	4,46	4,51	4,52	4,53
	Male 2 (0.00- 7.96 s).wav	1,99	2,65	3,34	3,90	4,23	4,39	4,45	4,47	4,48
	усредненное значение	1,95	2,57	3,23	3,85	4,22	4,39	4,46	4,48	4,49
Китайский	Female 1 (0.00-10.87 s).wav	1,41	1,84	2,45	3,19	3,83	4,20	4,40	4,49	4,52
	Female 1b (0.00-13.39 s).wav	1,35	1,74	2,36	3,07	3,66	4,02	4,27	4,36	4,38
	Female 2 (0.00-13.32 s).wav	1,46	1,91	2,54	3,33	3,99	4,33	4,46	4,51	4,54
	Female 2b (0.00-13.39 s).wav	1,52	2,02	2,72	3,51	4,10	4,35	4,42	4,44	4,45
	Male 1 (0.00-12.15 s).wav	1,72	2,29	3,02	3,71	4,16	4,33	4,37	4,38	4,38
	Male 1a (0.00-12.91 s).wav	1,78	2,44	3,22	3,91	4,26	4,43	4,51	4,54	4,54
	Male 2 (0.00-12.50 s).wav	1,67	2,22	3,02	3,78	4,19	4,36	4,46	4,51	4,52
	Male 2b (0.00-12.82 s).wav	1,81	2,46	3,25	3,91	4,29	4,42	4,46	4,47	4,47
	усредненное значение	1,57	2,10	2,82	3,56	4,08	4,31	4,42	4,47	4,48
Английский	Female 1 (0.00- 8.00 s).wav	1,97	2,60	3,27	3,86	4,22	4,41	4,49	4,52	4,53
	Female 2 (0.00- 8.00 s).wav	1,90	2,51	3,16	3,64	4,06	4,34	4,45	4,50	4,52
	Male 1 (0.00- 8.00 s).wav	2,41	3,06	3,65	4,11	4,38	4,47	4,51	4,52	4,53
	Male 2 (0.00- 8.00 s).wav	1,90	2,55	3,32	3,90	4,27	4,45	4,51	4,54	4,54
	усредненное значение	2,03	2,68	3,35	3,89	4,24	4,42	4,49	4,52	4,53
Французский	Female 1 (0.00-10.04 s).wav	1,93	2,55	3,22	3,82	4,23	4,42	4,50	4,53	4,54
	Female 2 (0.00-10.04 s).wav	1,74	2,31	3,02	3,70	4,15	4,40	4,50	4,53	4,54
	Male 1 (0.00-12.18 s).wav	2,07	2,76	3,47	3,96	4,23	4,41	4,49	4,53	4,54
	Male 2 (0.00-10.04 s).wav	2,36	3,14	3,88	4,30	4,46	4,51	4,53	4,54	4,54
	усредненное значение	2,01	2,68	3,41	3,96	4,27	4,44	4,51	4,53	4,54
Немецкий	female1 (0.00- 8.00 s).wav	1,70	2,28	3,00	3,65	4,19	4,41	4,49	4,53	4,54
	female2 (0.00- 8.00 s).wav	1,74	2,26	2,94	3,64	4,13	4,33	4,47	4,51	4,53
	male1 (0.00- 8.00 s).wav	1,88	2,47	3,17	3,75	4,20	4,42	4,51	4,53	4,54
	male2 (0.00- 8.00 s).wav	2,09	2,71	3,40	3,97	4,33	4,46	4,51	4,53	4,54
	усредненное значение	1,84	2,43	3,13	3,76	4,21	4,41	4,49	4,53	4,54
Итальянский	Female 1 (0.00-20.60 s).wav	1,52	2,01	2,69	3,44	4,01	4,33	4,46	4,51	4,53
	Female 2 (0.00-21.78 s).wav	1,73	2,25	2,92	3,58	4,12	4,38	4,47	4,51	4,53
	Male 1 (0.00-18.13 s).wav	1,79	2,43	3,24	3,94	4,33	4,46	4,51	4,53	4,54
	Male 2 (0.00-20.86 s).wav	2,09	2,86	3,65	4,16	4,40	4,49	4,52	4,53	4,54
	усредненное значение	1,76	2,37	3,13	3,80	4,23	4,42	4,49	4,52	4,53
Японский	Female 1 (0.00- 7.60 s).wav	1,49	1,94	2,57	3,32	3,92	4,27	4,43	4,50	4,52
	Female 2 (0.00- 7.31 s).wav	1,48	1,91	2,51	3,27	3,99	4,35	4,48	4,52	4,53
	Male 1 (0.00- 7.13 s).wav	1,55	2,02	2,65	3,39	4,05	4,36	4,47	4,52	4,53
	Male 2 (0.00- 7.45 s).wav	1,74	2,27	3,01	3,77	4,24	4,45	4,51	4,53	4,54
	усредненное значение	1,56	2,03	2,68	3,44	4,05	4,36	4,47	4,52	4,53
Испанский (США)	Female 1 (0.00- 8.00 s).wav	1,49	1,95	2,51	3,16	3,81	4,24	4,43	4,50	4,53
	Female 2 (0.00- 8.00 s).wav	1,61	2,08	2,62	3,19	3,81	4,19	4,40	4,49	4,52
	Male 1 (0.00- 8.00 s).wav	2,02	2,70	3,53	4,14	4,39	4,48	4,51	4,52	4,53
	Male 2 (0.00- 8.00 s).wav	1,65	2,24	3,04	3,73	4,23	4,44	4,51	4,53	4,54
	усредненное значение	1,67	2,23	2,92	3,59	4,08	4,35	4,46	4,51	4,53

ПРИМЕЧАНИЕ. – Ячейки серого цвета показывают речевые сигналы, не соответствующие требованиям Рекомендации МСЭ-Т Р.501.

Таблица I.3/Р.862.3 – Эталонные значения по Рекомендации Р.862 для стандартных кодеков GSM

Язык	Название файла	AMR								EFR	FR	HR
		12,2 кбит/с	10,2 кбит/с	7,95 кбит/с	7,4 кбит/с	6,7 кбит/с	5,9 кбит/с	5,15 кбит/с	4,75 кбит/с			
Американский английский	Female 1 (0.00- 7.97 s).wav	3,87	3,75	3,60	3,61	3,52	3,43	3,33	3,18	3,94	3,03	3,20
	Female 2 (0.00- 8.06 s).wav	4,13	4,07	3,96	3,92	3,79	3,72	3,62	3,49	4,08	3,70	3,54
	Male 1 (0.00- 8.44 s).wav	4,10	4,03	3,94	3,97	3,88	3,84	3,64	3,50	4,19	3,69	3,46
	Male 2 (0.00- 7.96 s).wav	4,10	4,06	4,03	4,01	3,91	3,85	3,73	3,70	4,20	3,59	3,72
	усредненное значение	4,05	3,98	3,88	3,88	3,78	3,71	3,58	3,47	4,10	3,51	3,48
Китайский	Female 1 (0.00-10.87 s).wav	3,94	3,81	3,46	3,50	3,27	3,15	2,99	3,00	3,98	3,18	2,94
	Female 1b (0.00-13.39 s).wav	3,97	3,82	3,52	3,45	3,36	3,22	3,00	2,95	4,04	3,02	2,95
	Female 2 (0.00-13.32 s).wav	4,08	3,99	3,71	3,62	3,53	3,38	3,21	3,07	4,12	3,07	3,01
	Female 2b (0.00-13.39 s).wav	4,08	4,01	3,74	3,74	3,62	3,40	3,21	3,13	4,10	3,03	3,04
	Male 1 (0.00-12.15 s).wav	3,94	3,84	3,66	3,68	3,54	3,45	3,22	3,22	4,06	3,50	3,26
	Male 1a (0.00-12.91 s).wav	4,23	4,14	3,95	3,92	3,78	3,65	3,34	3,35	4,20	3,65	3,36
	Male 2 (0.00-12.50 s).wav	4,07	3,93	3,78	3,77	3,62	3,46	3,22	3,21	4,15	3,72	3,30
	Male 2b (0.00-12.82 s).wav	4,16	4,15	3,94	3,94	3,80	3,64	3,40	3,27	4,23	3,64	3,44
	усредненное значение	4,06	3,96	3,72	3,70	3,57	3,42	3,20	3,15	4,11	3,35	3,16
Английский	Female 1 (0.00- 8.00 s).wav	4,00	3,82	3,63	3,62	3,42	3,31	3,21	3,11	3,99	3,27	3,07
	Female 2 (0.00- 8.00 s).wav	3,81	3,78	3,65	3,62	3,56	3,49	3,36	3,34	3,78	3,31	3,28
	Male 1 (0.00- 8.00 s).wav	4,01	3,88	3,75	3,67	3,46	3,57	3,23	3,01	4,03	3,54	3,37
	Male 2 (0.00- 8.00 s).wav	4,06	3,83	3,73	3,68	3,53	3,48	3,18	2,98	4,10	3,75	3,49
	усредненное значение	3,97	3,83	3,69	3,65	3,49	3,46	3,25	3,11	3,98	3,47	3,30
Французский	Female 1 (0.00-10.04 s).wav	4,11	4,01	3,79	3,83	3,66	3,42	3,33	3,30	4,04	3,49	3,37
	Female 2 (0.00-10.04 s).wav	3,91	3,85	3,57	3,54	3,40	3,31	3,22	2,98	3,83	3,19	3,24
	Male 1 (0.00-12.18 s).wav	4,00	3,88	3,71	3,72	3,54	3,37	3,23	3,14	4,07	3,49	3,32
	Male 2 (0.00-10.04 s).wav	4,11	4,05	3,85	3,91	3,75	3,57	3,37	3,28	4,17	3,84	3,30
	усредненное значение	4,03	3,95	3,73	3,75	3,59	3,42	3,29	3,18	4,03	3,50	3,31
Немецкий	female1 (0.00- 8.00 s).wav	4,08	3,98	3,65	3,60	3,54	3,36	3,15	3,06	4,06	3,40	3,20
	female2 (0.00- 8.00 s).wav	4,21	4,14	3,93	3,88	3,76	3,63	3,53	3,47	4,17	3,54	3,37
	male1 (0.00- 8.00 s).wav	4,12	4,08	3,90	3,88	3,78	3,66	3,55	3,50	4,19	3,82	3,43
	male2 (0.00- 8.00 s).wav	4,17	4,07	3,97	3,92	3,77	3,69	3,65	3,56	4,21	3,70	3,38
	усредненное значение	4,14	4,07	3,86	3,82	3,72	3,58	3,47	3,40	4,16	3,62	3,35
Итальянский	Female 1 (0.00-20.60 s).wav	3,80	3,67	3,51	3,41	3,34	3,24	3,12	2,92	3,81	2,83	3,00
	Female 2 (0.00-21.78 s).wav	4,09	4,04	3,88	3,86	3,74	3,61	3,40	3,27	4,14	3,29	3,32
	Male 1 (0.00-18.13 s).wav	4,03	3,95	3,82	3,78	3,64	3,49	3,32	3,20	4,13	3,37	3,17
	Male 2 (0.00-20.86 s).wav	4,23	4,15	4,00	4,07	3,89	3,84	3,60	3,47	4,27	3,58	3,47
	усредненное значение	4,04	3,95	3,80	3,78	3,65	3,55	3,36	3,22	4,08	3,27	3,24
Японский	Female 1 (0.00- 7.60 s).wav	3,89	3,75	3,53	3,41	3,40	3,28	3,08	3,10	3,87	2,92	2,92
	Female 2 (0.00- 7.31 s).wav	3,92	3,85	3,54	3,62	3,42	3,27	3,20	3,08	3,99	2,82	3,09
	Male 1 (0.00- 7.13 s).wav	3,87	3,81	3,59	3,50	3,42	3,28	3,19	3,12	3,91	2,89	2,91
	Male 2 (0.00- 7.45 s).wav	4,18	4,08	3,92	3,94	3,84	3,73	3,59	3,44	4,21	3,63	3,55
	усредненное значение	3,97	3,87	3,64	3,61	3,52	3,39	3,27	3,19	4,00	3,06	3,12
Испанский (США)	Female 1 (0.00- 8.00 s).wav	4,03	3,96	3,63	3,68	3,50	3,30	3,23	3,16	3,96	3,03	3,20
	Female 2 (0.00- 8.00 s).wav	3,63	3,48	3,23	3,26	3,05	3,03	2,88	2,80	3,73	2,68	2,72
	Male 1 (0.00- 8.00 s).wav	3,98	3,61	3,67	3,49	3,51	3,38	3,23	3,15	4,15	3,44	3,19
	Male 2 (0.00- 8.00 s).wav	4,04	3,87	3,71	3,58	3,52	3,28	3,11	3,14	4,09	3,48	3,36
	усредненное значение	3,92	3,73	3,56	3,50	3,40	3,25	3,11	3,06	3,99	3,16	3,12

ПРИМЕЧАНИЕ. – Ячейки серого цвета показывают речевые сигналы, не соответствующие требованиям Рекомендации МСЭ-Т Р.501.

Приложение II

Базы данных испытаний для P.862/P.862.1

Базы данных испытаний включали в себя описания условий работы сетей, которые соответствовали фиксированным, беспроводным и VoIP приложениям. В испытания и анализ по алгоритмам P.862 и P.862.1 был включен большой набор кодеков, используемых различными технологиями (например, GSM-FR, GSM-EFR, GSM-AMR, CDMA-EVRC, IS136-ACELP, G.711, G.726, G.728, G.729, JDC-HR).

В таблице II.1 описано содержание всех баз данных, по которым проверялись результаты PESQ (P.862 и P.862.1).

Таблица II.1/P.862.3 – Обзор типов и содержаний баз данных

Описание	Условия
8 кбит/с характеристический канал, взаимодействующий со стандартами, дополнение 23 к рекомендациям серии P, доп. 1	кодер/декодер + транс.; кодеки: G.711, G.726, G.728, G.729, Is-54, GSM-FR, JDC-HR.
8 кбит/с характеристический канал, ошибки и шумы, дополнение 23 к рекомендациям серии P, доп. 3	кодер/декодер + транскодер + ErrorPatterns + BGN, кодеки: G.729.
Реально действующие беспроводные сети	– Беспроводные сети: IS-136, CDMA, iDEN, AMPS; GSM-US, GSM-Europe; – Кодеки: IS-54, 8 кбит/с ACELP, 13 кбит/с QCELP, GSM-FR и EFR, CDMA- EVRC.
Кодеки, шаблоны ошибок, перекодирование, шумы	– кодер/декодер + ErrorPatterns + BGN; кодеки: G.711, G.726, G.728, G.729, GSM-FR; – кодер/декодер + ErrorPatterns (C/I levels) + BGN; кодеки: G.711, G.723 + перекодирование, 8 кбит/с ACELP, EVRC, GSM-EFR и FR; – кодер/декодер + trans + ErrorPatterns ("плохие" и "хорошие" условия) + BGN, ATM/ISDN/POTS; кодеки: G.729, G.728, GSM-FR, GSM-HR.
Тест на фоновый шум, GSM и фиксированные сети	– кодер/декодер + VAD + VQE (NR); кодеки: GSM-FR, G.729; – кодер/декодер + ErrorPatterns; кодеки: GSM-FR, GSM-AMR-HR.
DTX, ошибки в кадре/пакете и тестирование VAD, AMR, GSM и фиксированные сети	кодер/декодер + ErrorPatterns; кодеки: GSM-FR, G.726, G.728, G.729, AMR.
AMR + шаблоны ошибок в "перескакивающих" и "неперескакивающих" GSM каналах	кодер/декодер + ErrorPatterns; кодеки: AMR475, AMR590, AMR740, AMR122.
AMR + шаблоны ошибок с разной и одинаковой защитой от ошибок	кодер/декодер + ErrorPatterns; кодеки: AMR515, AMR740, AMR102, AMR122, AMR475, AMR590, AMR670, AMR795.
VoIP шаблоны ошибок и потери пакетов; шумы	кодер/декодер + транс + ErrorPatterns + PacketLoss + BGN.
VoIP сценарии испытаний при потере пакетов	VoIP: PktInsert, PktDelete, PktMute.
VoIP кодеки	VoIP: G.723.1, G.728, G.711, G.729, G.726, PDC-HR.
Условия работы VoIP	PLC и управление буфером фазового дрожания (G.711, G.729, G.723 в 6,3 и 5,3 кбит/с), переменная длина пакета.
Реальные сети VoIP	VoIP: УПАТС Интернет, IP шлюзы.

Приложение III

Отчет об измерениях в соответствии с P.862/P.862.1

Результаты измерения P.862/P.862.1 должны быть представлены с учетом точности алгоритма, указанного в § 10.3 (таблицы 1 и 2).

III.1 Отчет и интерпретация усредненных результатов измерений PESQ

Как отмечено в § 10.1, в случае проведения испытаний на управляемой моделируемой сети рекомендуется усреднить результаты для каждого из условий, используя, как минимум, четырех дикторов.

Ожидается, что усредненные результаты по P.862.1 будут определены с погрешностью предсказания, равной верхнему пределу ± 95 -процентно доверительного интервала.

В случае испытаний на работающих сетях существуют ситуации, когда желательно грубо оценить качество речи, обеспечиваемое сетью в пределах определенного района испытаний и/или во время временного интервала. В таком случае результаты, полученные по P.862.1, усредняются по трассе движения и/или по временному отрезку. Усредненные результаты, полученные по P.862.1, как ожидается, будут определены с погрешностью предсказания, равной максимальному значению верхнего предела ± 95 -процентно доверительного интервала. Следовательно, предполагая, например, что испытываемой сетью является сеть CDMA, можно ожидать, что средняя погрешность предсказания усредненных результатов измерения по P.862.1 будет меньше или равна 0,462 MOS.

Тем не менее следует отметить, что для обоих представленных выше случаев – измерения на модели сети и на работающей сети, существует риск в 5%, что ошибка измерений будет больше чем 95-процентный верхний предел погрешности предсказания. Кроме того, как отмечено в § 10.2, сценарии измерений должны быть аналогичными тем, что представлены в таблице 1.

III.2 Отчет и интерпретация отдельных результатов измерений PESQ

Как отмечено в § 10.1, в случае испытаний на работающей сети требуется оценивать качество для каждого образца речи, т. к. каналы передачи изменяются во времени.

Существует две рекомендованные процедуры отчета, выбор которых зависит от типа измерений PESQ, которые используются для анализа качества передачи речи в сети.

Первый тип – это усредненная оценка по P.862.1, которая рассмотрена в разделе III.1 и является грубой оценкой качества речи в пределах определенной области или в течение определенного периода времени.

Другой тип – отдельные измерения по P.862.1, когда требуется выявить и устранить неисправности в сети.

При выполнении отдельных измерений рекомендуется рассчитать гистограмму результатов измерений P.862.1 на шкале MOS. При выполнении анализа по этой процедуре рекомендуется определить порог субъективного качества речи и установить минимальный требуемый процент результатов, которые могут располагаться выше этого порога или, что эквивалентно, минимальную требуемую вероятность, с которой результаты измерений по P.862.1 в испытываемой сети будут оказываться выше этого порога. Таким образом, качество речи, основанное на отдельных оценках по P.862.1, представляет собой вероятность или процент оценок, которые располагаются выше установленного порога субъективного качества речи.

При оценке этой вероятности или процента на результаты оказывают влияние два типа ошибок. Первая ошибка обусловлена вычислениями вероятности или процента. Вторая ошибка определяется распределением остаточной ошибки в алгоритме P.862/P.862.1. Из-за ошибки измерений по P.862.1 не все точки, которые должны быть определены выше (и/или ниже) установленного порога MOS, окажутся там на самом деле. Как отмечено в § 10.4, распределение остаточной ошибки (см. таблицу 2) показывает, что вероятность появления меньших абсолютных ошибок выше, чем более высоких абсолютных ошибок. Следовательно, когда оценка, полученная в измерениях по

P.862.1, находится с одной стороны порога MOS, вероятность того, что она появится с другой стороны этого порога, тем выше, чем ближе к порогу находится реальная оценка. Например, пусть порог MOS=3 и измерения проводятся на сети CDMA, тогда вероятность того, что оценки между 3 и 3,1, полученные по P.862.1, окажутся ниже порога MOS=3, составит примерно 40,44%. Вероятность того, что оценки между 3,1 и 3,2, полученные по P.862.1, окажутся ниже порога MOS=3, составит 30,04%. Аналогично, вероятность того, что оценки между 2,9 и 3 окажутся ниже порога MOS=3, составит 40,44%, и вероятность того, что оценки между 2,8 и 2,9 окажутся ниже порога MOS=3, составит 30,04%.

Таким образом, рекомендованный метод для расчета результатов измерений по P.862.1 имеет следующий вид:

Ошибка измерений по P.862.1 из-за ошибки оценки вероятности того, что оценки, полученные по P.862.1, оказываются выше установленного порога MOS, определяется как стандартная девиация биномиального распределения с вероятностью появления:

$$p = \frac{n}{N},$$

где n – число оценок выше порога, а N – общее число измерений по P.862.1. Следовательно, ошибка, обусловленная вероятностью оценки $Error1$, определяется соотношением:

$$Error1 = \pm \sqrt{\frac{p \times (1-p)}{N}}$$

и характеризуется 95-процентным доверительным интервалом:

$$\pm z_{\alpha} \times Error1,$$

где: $z_{\alpha} = 2$ – квантиль 95-процентной вероятности (нормальное распределение достаточно хорошо описывает характеристику отдельных оценок PESQ).

Рекомендуется рассчитывать ошибку измерений по P.862/P.862.1 из-за остаточной ошибки алгоритма PESQ – $Error2$ как стандартная девиация биномиального распределения, которое описывает остаточную ошибку, представленную в таблице 2.

Следовательно, когда используются отдельные оценки по P.862/P.862.1, рекомендуется вычислять общую ошибку измерений качества речи как квадратный корень суммы квадратов двух вышеуказанных ошибок – $Error1$ и $Error2$.

Приложение IV

Метод калибровки патентованных интерфейсов

ПРИМЕЧАНИЕ. – Описанный здесь метод предназначен для ситуаций, когда неизвестны точные требуемые уровни сигналов и не могут использоваться методы, рекомендованные в разделе 9. Нельзя ожидать, что с использованием этого метода будет достигнута максимальная точность. Важно также, что все регулировки уровня либо выполняются в аналоговой области, либо с использованием достаточной длины слов в цифровой области (как минимум 24 бита).

IV.1 Калибровка уровня передачи (на стороне передачи) испытательного оборудования

Выходной уровень измерительного оборудования должен быть отрегулирован до уровней примерно в середине рабочего диапазона AGC. Это может быть выполнено при помощи измерения уровня сигнала на приемном терминале и настройки ослабления уровня сигнала на передаче с доведением его до среднего значения диапазона, где он будет оставаться постоянным. И наоборот, телефон может использоваться на приеме, и уровень на передаче выставляется в середине диапазона, который определяет комфортный уровень прослушивания на приеме. Второй метод также предпочтителен в ситуациях, когда нет AGC.

IV.2 Калибровка уровня приема (на стороне приема) испытательного оборудования

Уровень на приеме должен быть отрегулирован таким образом, чтобы ослабление между записанным файлом и эталонным файлом было бы равным 0 дБ.

БИБЛИОГРАФИЯ

- [B.1] Объективная оценка качества по Рекомендации МСЭ-Т Р.862 с использованием длинных эталонных отрывков речи (NTT), COM12-D008, январь 2005 г.
- [B.2] Объективная оценка качества с использованием искусственных голосовых сигналов (NTT), COM12-D145, сентябрь 2003 г.
- [B.3] Добавление уровня шума к эталонной речи, используемое в Рекомендации МСЭ-Т Р.862, COM12-D011, январь 2005 г.
- [B.4] ANSI/TIA-127-A-2004 (2004), *Улучшенный кодек с переменной скоростью, для речевой услуги вариант 3 для широкополосных цифровых систем с расширением спектра*

СЕРИИ РЕКОМЕНДАЦИЙ МСЭ-Т

Серия А	Организация работы МСЭ-Т
Серия D	Общие принципы тарификации
Серия E	Общая эксплуатация сети, телефонная служба, функционирование служб и человеческие факторы
Серия F	Нетелефонные службы электросвязи
Серия G	Системы и среда передачи, цифровые системы и сети
Серия H	Аудиовизуальные и мультимедийные системы
Серия I	Цифровая сеть с интеграцией служб
Серия J	Кабельные сети и передача сигналов телевизионных и звуковых программ и других мультимедийных сигналов
Серия K	Защита от помех
Серия L	Конструкция, прокладка и защита кабелей и других элементов линейно-кабельных сооружений
Серия M	Управление электросвязью, включая СУЭ и техническое обслуживание сетей
Серия N	Техническое обслуживание: международные каналы передачи звуковых и телевизионных программ
Серия O	Требования к измерительной аппаратуре
Серия P	Качество телефонной передачи, телефонные установки, сети местных линий
Серия Q	Коммутация и сигнализация
Серия R	Телеграфная передача
Серия S	Оконечное оборудование для телеграфных служб
Серия T	Оконечное оборудование для телематических служб
Серия U	Телеграфная коммутация
Серия V	Передача данных по телефонной сети
Серия X	Сети передачи данных, взаимосвязь открытых систем и безопасность
Серия Y	Глобальная информационная инфраструктура, аспекты межсетевых протоколов и сети последующих поколений
Серия Z	Языки и общие аспекты программного обеспечения для систем электросвязи