РАЗВИТИЕ ИНТРУЗИВНЫХ ПСИХОАКУСТИЧЕСКИХ МЕТОДОВ ОБЪЕКТИВНОГО ОЦЕНИВАНИЯ КАЧЕСТВА ПЕРЕДАЧИ РЕЧЕВОГО СИГНАЛА В СИСТЕМАХ БЕСПРОВОДНОЙ СВЯЗИ

М.В. Илюшин

Академия ФСО России, Орел, Российская Федерация

Рассмотрена задача объективного оценивания качества передачи речевого сигнала в системах беспроводной связи. Представлены основные этапы эволюционного развития активных (интрузивных) методов оценивания качества восприятия речевой информации, передаваемой в телекоммуникационных сетях. Подробно рассмотрен алгоритм POLQA, представляющий собой новейшую технологию тестирования качества восприятия восстановленной речи для стационарных, мобильных и IP-сетей, которые используют широкий спектр технологий кодирования, передачи и повышения качества речевого сигнала. Представленные результаты подтверждают перспективность использования интрузивных методов на основе психоакустической модели в задаче получения объективных оценок качества восприятия речи.

Ключевые слова: речевой сигнал, качество восприятия речи, среднее значение оценок, POLQA, алгоритм оценивания качества передачи речи, психоакустическая модель.

DEVELOPMENT OF INTRUSIVE PSYCHOACOUSTIC METHODS PROGRESSION OF OBJECTIVE ASSESSMENT OF SPEECH SIGNAL TRANSMISSION QUALITY IN WIRELESS COMMUNICATION SYSTEMS

M. V. Ilyushin

Academy FGS of Russia, Orel, Russian Federation

The objective of assessing the quality of voice transmission in wireless communication systems is considered. The main stages of the evolutionary development of active (intrusive) methods for assessing the quality of perception of speech information transmitted in telecommunication networks are presented. The POLQA algorithm, which is the latest technology of testing the quality of perception restored speech for stationary, mobile and IP-networks that use a wide range of technologies of encoding, transmission and improving the quality of the speech signal, is considered in detail. The presented results confirm the prospects of using intrusive methods based on the psychoacoustic model in the task of obtaining objective assessments of speech perception quality.

Keywords: speech signal, speech perception quality, average value of assessments, POLQA, algorithm of evaluation of speech signal transmission quality, psychoacoustic model.

Речь, являясь естественной опорой коммуникации и наиболее универсальным инструментом человеческого общения, не только сохраняет свою значимость в период активной эволюции современного общества, но и приобретает еще большую важность, обусловленную постоянным возникновением новых видов сервиса, основанных на синтезе исключительных свойств речи и растущих возможностей современных цифровых инфокоммуникационных технологий. Высокие темпы развития показывают технологии беспроводного доступа, являясь неотъемлемой частью глобальной информационной инфраструктуры, основная задача которой состоит в передаче различной информации от любого пользователя к любому другому пользователю, независимо от их географического месторасположения. Речевые сообщения по-прежнему составляют значительную долю информации, циркулирующей в сетях радиосвязи с подвижными объектами.

Современное состояние и перспективы развития инфокоммуникационных технологий в бизнесе и обществе характеризуются повышением требований к качеству предоставляемых услуг связи, определяя тем самым основные направления конкурентной борьбы операторов сетей связи. Важным остается вопрос: как оценивать качество передачи речевого сигнала (РС)?

Ответом на заданный вопрос может послужить ряд документов [1–5], предписывающих вычислять оценку как среднее значение оценок *MOS*, полученных либо с помощью прибора, реализующего алгоритм определения средней абонентской оценки в соответствии с руководящими документами, при объективном методе оценивания либо по результатам статистических испытаний при субъективном методе. Субъективно-статистические методы дают наиболее точную оценку, но являются трудоемкими, дорогостоящими и исключают возможность автоматизации процесса. Поэтому в последнее время наметилась тенденция применения анализаторов систем связи, реализующих тот или иной объективный алгоритм оценивания качества передачи РС.

Особый интерес для исследований представляют активные (интрузивные) психоакустические методы оценивания качества звучания восстановленной речи, основанные на сравнении эталонного РС и искаженного при его передаче по сети связи или преобразовании каким-либо устройством. Указанные методы направлены на имитацию с помощью технических и программных средств алгоритма слуховой перцепции человека. Мерой оценки качества в этом случае является разница между «внутренними представлениями» обоих сигналов, формируемыми с помощью функций от параметров восприятия [6–9].

В 1996 году МСЭ-Т была разработана рекомендация Р.861 [10], регламентирующая алгоритм измерения качества восприятия речи PSQM (Perceptual Speech Quality Measure) для анализа узкополосных (300–3400 Гц) кодеков РС с учетом свойств слуха. Использование алгоритма PSQM для оценки работы реальной сети было весьма затруднительным, так как метод не учитывал влияние задержек, их возможную флуктуацию (джиттер), потерю пакетов, а также клиппирование по уровню, что обычно является причиной больших искажений сигнала. Такие факторы как внешний шум и фильтрация в аналоговых устройствах учитывались также слабо. Следствием указанных недостатков явились низкие коэффициенты корреляции оценок PSQM с оценками, полученными в результате субъективных экспертиз.

С появлением технологии VoIP возникла необходимость в учете основных факторов, влияющих на качество передачи речи. В 1997 году алгоритм PSQM был модифицирован разработчиками компании Opticom и получил наименование PSQM+. Указанный алгоритм стал экспериментальной версией PSQM, созданной для оценки качества передачи речи в системах сотовой связи стандарта GSM. Это проявлялось в том, что при тестировании исходного и восстановленного РС в условиях идеального канала связи PSQM+ и PSQM показывали одинаковые значения оценки MOS.

Результатом очередного совершенствования PSQM стал алгоритм PSQM99, разработанный компанией KPN Research в 1999 году. PSQM99 отличался от своих предшественников асимметричной обработкой, шкалой измерения результатов и более высокой корреляцией оценок PSQM99 с оценками экспертов. Однако слабый учет условий функционирования реальных систем передачи речевой информации не позволил получить алгоритму PSQM99 широкое распространение.

Система измерения путем анализа восприятия PAMS (Perceptual Analysis Measurement System) была разработана в 1998 году (последняя, третья, версия PAMS вышла в декабре 1999 года) и оптимизирована для работы с реальными системами VoIP. Главное отличие психоакустической модели алгоритма PAMS от модели, использованной в PSQM, состоит в выравнивании эталонного (входного) и искаженного (выходного) сигналов как по уровню, так и во времени, что позволило удалить негативное влияние больших задержек и джиттера.

С учетом недостатков предшествующих алгоритмов, обусловленных невозможностью анализа сигналов с переменной задержкой и кратковременными искажениями, в ос-

новые рекомендации Р.862 [11] был принят алгоритм PESQ (Perceptual Evaluation of Speech Quality), ориентированный на анализ РС в сетях с КП. В алгоритме PESQ учитываются следующие причины ухудшения качества сигнала: искажение его при кодировании, ошибки при передаче, потеря пакетов, время задержки передачи пакетов и флуктуация этого времени, фильтрация сигнала в аналоговых сетевых компонентах.

В PESQ было заложено множество новых разработок, что выгодно отличает его от предыдущих перцептуальных алгоритмов оценивания качества передачи речи. Эти инновации дают возможность уверенно использовать PESQ как для определения качества сквозной (end-to-end) передачи речи, так и для оценки влияния на качество связи отдельных элементов сетевого оборудования, включая речевые кодеки. Сравнительный анализ некоторых психоакустических методов оценивания качества передачи РС для англоязычных речевых баз представлен в таблице [12].

Результаты корреляционного анализа объективных и субъективных оценок

Вид сети \ Алгоритм	PESQ	PAMS	PSQM	PSQM+
Сеть с подвижными объектами	0,962	0,954	0,924	0,935
Стационарная сеть	0,942	0,936	0,881	0,897
VoIP сеть	0,918	0,916	0,674	0,726

Анализ таблицы показывает, что алгоритм PESQ показывает сравнительно высокие коэффициенты корреляции с оценками субъективных англоязычных тестов. В процессе разработки стандарта PESQ специалистами МСЭ-Т были отобраны лучшие методы определения качества речевой связи с точки зрения корреляции даваемых ими результатов с оценками MOS при различных условиях связи, что является гарантией хорошей работы стандартизированного алгоритма при тестировании обычных (фиксированных и мобильных) сетей и систем передачи пакетных данных. Совершенствование алгоритма привело к появлению оценок MOS-LQO_n [13] и MOS-LQO_w [14] при анализе РС в узкой (0,3–3,4 кГц) и широкой (0,05–7 кГц) полосах частот соответственно.

Указанные оценки наиболее близки к субъективным оценкам MOS и получаются в результате однозначных преобразований исходных оценок PESQ MOS (рис. 1).

Рис. 1. Зависимости оценок MOS-LQO_n и MOS-LQO_w от оценок PESQ MOS

Как видно из рис. 1, оценка PESQ MOS не превышает значение 4,5, которое является максимальным показателем в субъективных тестах MOS при анализе узкополосного РС. Такая оценка указывает на отсутствие каких-либо искажений в сравниваемых сигналах. Оценки MOS-LQO_n и MOS-LQO_w не превышают значения 4,5 и 4,6 соответственно.

Исходный код интрузивного алгоритма PESQ имеется в свободном доступе и реализован в виде отдельной фазы измерений в анализаторе систем связи AnCom TDA-9, что облегчает задачу оценивания качества передачи речи [15].

Вместе с тем в PESQ не учитываются влияние на качество связи изменений уровня сигнала в сети, эхо-сигнала, растяжения и сжатия РС во временной области и круговой задержки. По этой причине развитием алгоритма PESQ стала рекомендация P.863 [16], регламентирующая алгоритм объективной оценки качества прослушивания POLQA (Perceptual Objective Listening Quality Assessment), который был призван устранить присущие его предшественникам недостатки.

POLQA – это новейшая технология тестирования качества восприятия восстановленной речи для стационарных, мобильных и IP-сетей, которые используют широкий спектр технологий кодирования, передачи и повышения качества речевого сигнала. Технология POLQA сохраняет основную идею интрузивных психоакустических методов – преобразование результата сравнения двух сигналов во внутреннее представление с использованием модели человеческого восприятия (рис. 2).

Рис. 2. Блок-схема общего алгоритма POLQA

Входными данными для алгоритма (рис. 2) являются два сигнала, представленные двумя векторами данных, содержащими 16-битные выборки ИКМ. Первый вектор содер-

жит отсчеты исходного (неискаженного) РС, в то время как второй вектор содержит отсчеты искаженного сигнала. Алгоритм POLQA включает блоки временного выравнивания, оценивания частоты дискретизации, результаты работы которого используются для компенсации различий в частотах дискретизации входных сигналов, и базовой модели ядра, выполняющий расчет MOS. На первом этапе определяется задержка между двумя входными сигналами и оценивается частота дискретизации двух сигналов относительно друг друга. Оценка частоты дискретизации основана на информации о задержке, рассчитанной с помощью временного выравнивания. Если частота дискретизации отличается более чем на 1%, для сигнала с большей частотой дискретизации она снижается. После каждого шага результаты сохраняются вместе со средним показателем надежности задержки, который является показателем качества оценивания задержки. В дальнейшем выбирается результат, который дал самую высокую общую надежность. После определения правильной задержки и компенсации различий в частотах дискретизации сигналы и информация о задержке передаются в базовую модель, которая рассчитывает чувствительность и раздражения от искажений и отображает их на шкале MOS [17].

Целью модели МСЭ-Т Р.863 является прогнозирование общего качества речевой связи для узкополосных (300–3400 Гц), широкополосных (50–7000 Гц), сверхширокополосных (50–14000 Гц) и полнополосных (20–20000 Гц) сигналов, которые воспринимаются пользователем. Она включает в себя все компоненты обработки речи, которые обычно рассматриваются для телекоммуникаций в идеальных и реальных условиях.

На рис. 3 представлены результаты вычисления функциональной взаимосвязи (линейной регрессии) между оценками MOS-LQO_n и MOS-LQS для англоязычной речевой базы в условиях искажений, характерных для мобильных и VoIP сетей [9; 16].

Рис. 3. Результаты вычисления функциональной взаимосвязи между оценками MOS-LQO_n и MOS-LQS

Представленные результаты (рис. 3) позволяют сделать вывод о том, что алгоритм POLQA показывает хорошую корреляцию с субъективными тестами и является перспективным интрузивным методом на основе психоакустической модели для получения объективных оценок качества восприятия РС, передаваемого в системах беспроводной связи.

СПИСОК ЛИТЕРАТУРЫ

1. ГОСТ Р 53532–2009. Качество услуг связи. Показатели качества услуг телефонной связи в сети общего пользования. Общие требования.
2. Правила применения комбинированных станций, использующих технологию коммутации пакетов информации.

3. Методика оценки качества услуг подвижной радиотелефонной связи.
4. ГОСТ Р 50840-95. Передача речи по трактам связи. Методы оценки качества, разборчивости и узнаваемости.
5. ГОСТ Р 51061-97. Системы низкоскоростной передачи речи по цифровым каналам. Параметры качества речи и методы измерений.
6. Качество передачи речи и его оценка / М. В. Илюшин [и др.]. Орел: Академия ФСО России, 2015. 103 с.
7. Рекомендация МСЭ-Т Р.800.1.
8. *Полторак В.П., Моргаль О.М., Заика Ю.А.* Оценка качества передачи речи в IP-телефонии // Молодой ученый. 2014. № 4 (63). С. 121–123.
9. ViSQOL: an objective speech quality model / A. Hines [et al.] // EURASIP: Journal on Audio, Speech and Music Processing. 2015. Vol. 13. P. 1–18.
10. Рекомендация МСЭ-Т Р.861.
11. Рекомендация МСЭ-Т Р. 862.
12. PESQ: An Introduction [Электронный ресурс]. Режим доступа: <http://www.psytechnics.com> (дата обращения: 04.06.2019).
13. Рекомендация МСЭ-Т Р.862.1.
14. Рекомендация МСЭ-Т Р.862.2.
15. Кочеров А.В. Качество и устойчивость ТфОП это просто // Т-Comm. 2010. С. 16–21.
16. Рекомендация МСЭ-Т Р.863.
17. POLQA Application Guide: Head acoustics documentation, 2012. 29 p.