

IP-телефонное соединение.
Особенности передачи речи по IP.
Кодеки.

ТфОП и IP

Факторы качества обслуживания	Факторы качества передачи речи	
Сети ТфОП и IP-телефонии	Общие факторы ТфОП и VoIP	Дополнительные параметры для сетей IP
<p><i>Услуги</i> – например, телефонное соединение, дополнительные услуги, переадресация вызова, голосовая почта и т.д.</p> <p><i>Доступность</i> – занятость сети</p> <p><i>Надежность</i> – потеря вызовов, неправильный набор номера</p> <p><i>Задержка после набора номера</i></p>	<p>Громкость</p> <p>Задержка</p> <p>Эхо</p> <p>Четкость:</p> <ul style="list-style-type: none">- <i>разборчивость</i>- <i>шум</i>- <i>ослабление</i>	<p>Четкость задержки-дрожания:</p> <ul style="list-style-type: none">- <i>потеря пакетов</i>- <i>полоса пропускания</i>- <i>сжатие</i>

IP-телефонное соединение

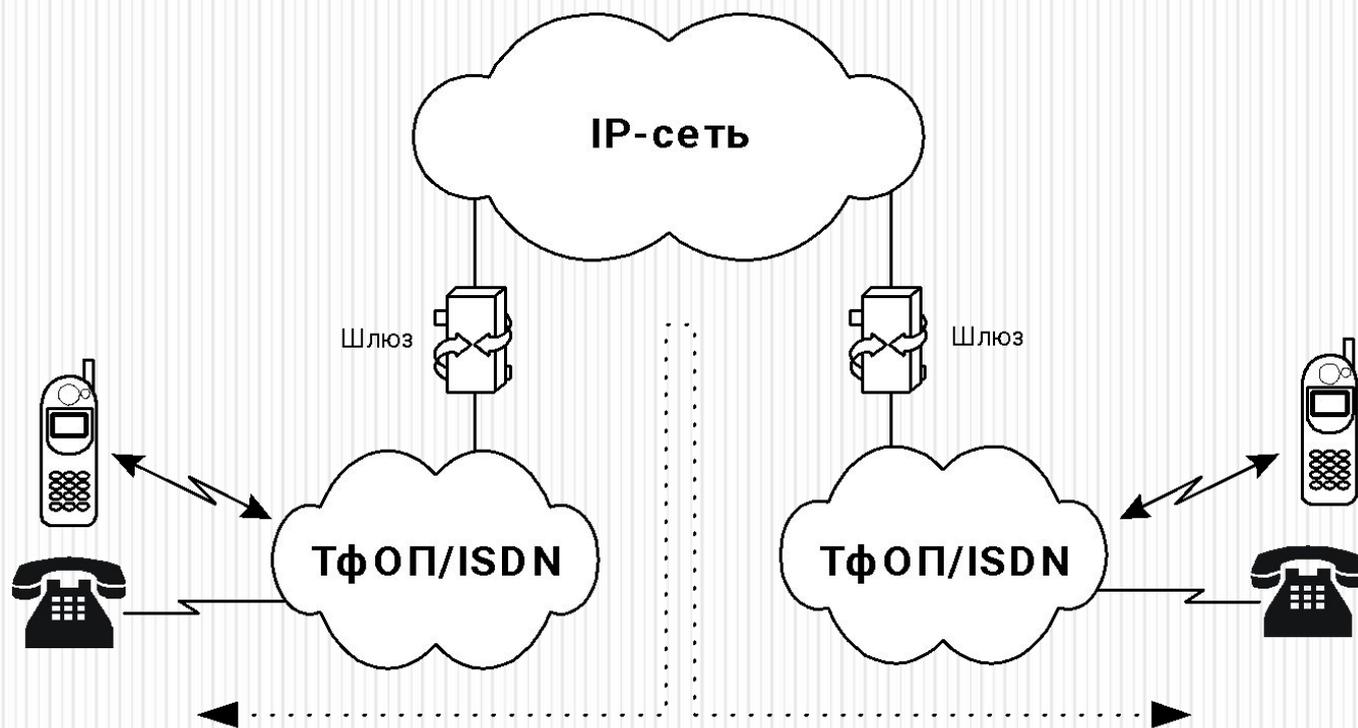
Что такое соединение в IP сети?

- Логическая связь двух устройств
- Устройства знают адреса IP друг друга
- Устройства знают номера портов взаимодействующих приложений
- Устройства знают какие модули надо подключить к портам транспортных протоколов (кодеки, мультиплексоры, приложения)
- Контролируется и управляется качество передачи пользовательских данных
- OKS7, SIP (-T, -I), BICC, H.323, MGCP, H.248/Megaco, SIGTRAN, EDSS-1

Сценарии IP-телефонии

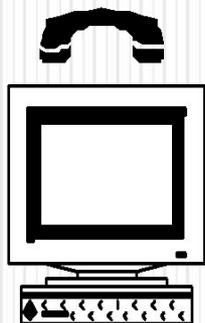
- Телефон-телефон
- Телефон – компьютер
- Компьютер – телефон
- Компьютер-компьютер

Соединение абонентов ТфОП через транзитную IP-сеть по сценарию «телефон-телефон»

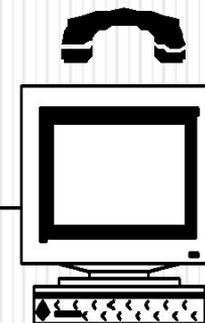


Сценарий “компьютер-компьютер”

Н.323-терминал



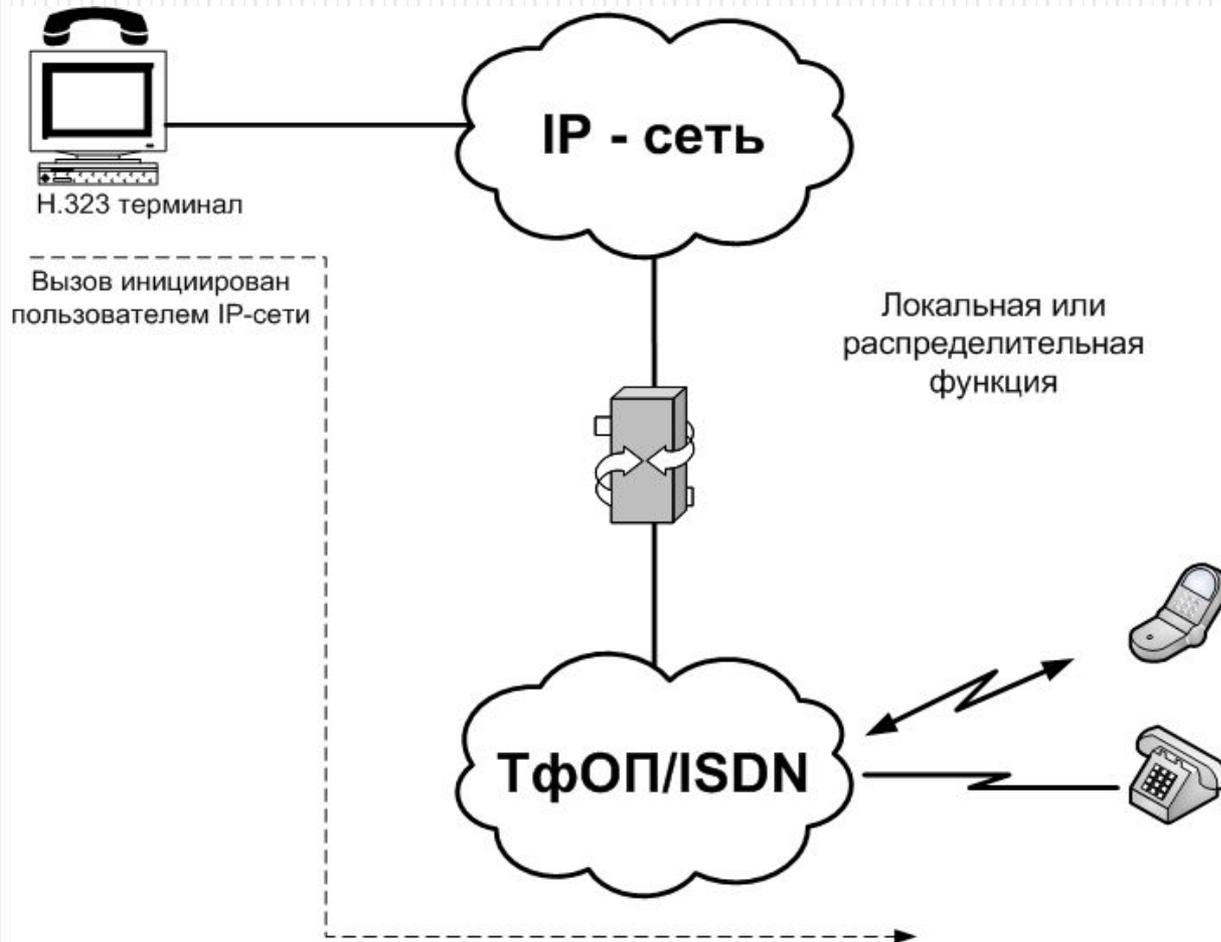
Н.323-терминал



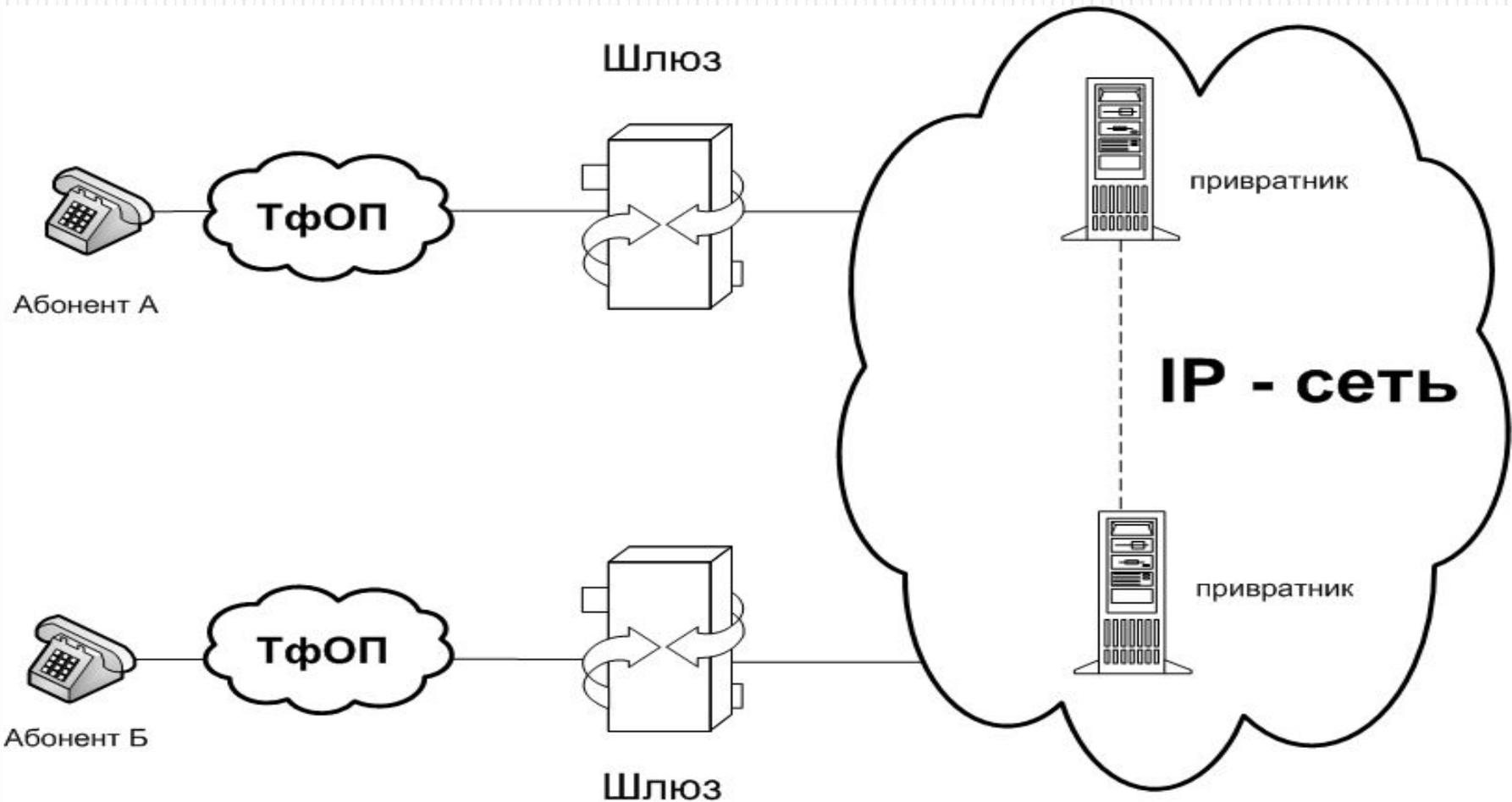
Сценарий “компьютер-компьютер”



Сценарий «компьютер – телефон»



Установка соединения в IP сети



Установление соединения в IP сети

1. Абонент А набирает местный номер доступа к шлюзу
2. Шлюз запрашивает у специального сервера данные о вызывающем абоненте
3. Сервер рассматривает информацию АОН для того, чтобы убедиться, что абоненту А разрешено пользоваться данной услугой, и затем передает к шлюзу сообщение аутентификации пользователя
4. Абонент А набирает телефонный номер вызываемого абонента Б

СЕТИ

5. Шлюз консультируется с привратником о возможных способах маршрутизации вызова
6. Привратник просматривает адрес E.164 на фоне таблицы маршрутизации и передает к исходящему шлюзу IP-адрес встречного (входящего) шлюза
7. Исходящий шлюз направляет вызов N.323 по IP – сети к входящему шлюзу
8. Входящий шлюз направляет вызов по сети ТфОП к вызываемому абоненту
9. Шлюзы посылают на специальный сервер данные о начале/окончании установления соединения для начисления платы за связь

Особенности передачи речи по IP

Особенности передачи речи по IP-сетям

- Задержки и вариация задержки
- Процент потерь пакетов
- Критичность эхоподавления
- Новые принципы кодирования речи
- Проблемы передачи DTMF и FAX

Задержки

- *Задержка (время запаздывания) - Промежуток времени, затрачиваемый на то, чтобы речевой сигнал прошел расстояние от говорящего до слушающего*

Задержки

1-й уровень - до 200 мс - отличное качество связи. Для сравнения, в ТФОП допустимы задержки до 150-200 мс.

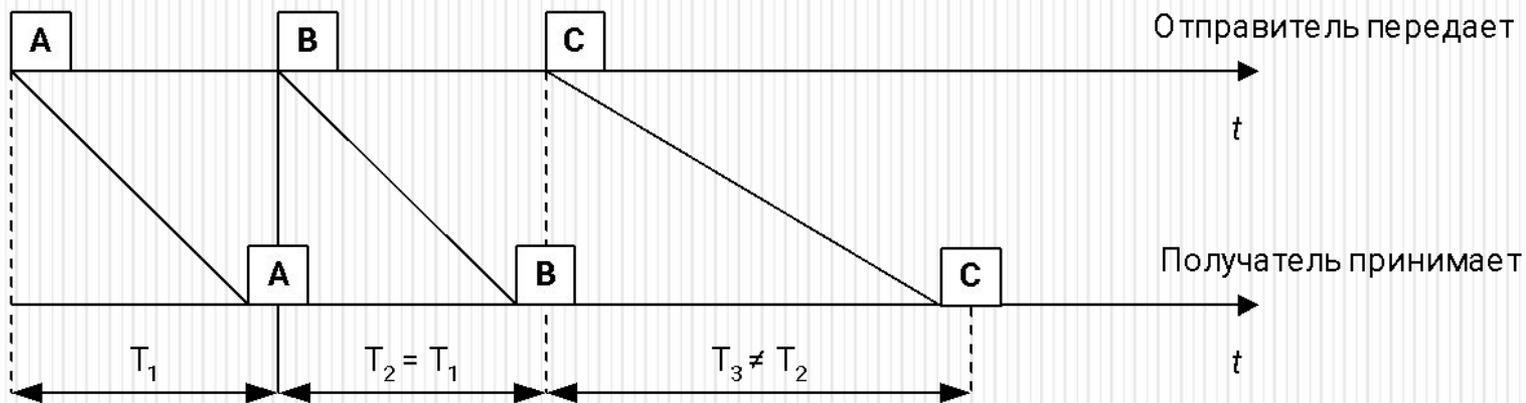
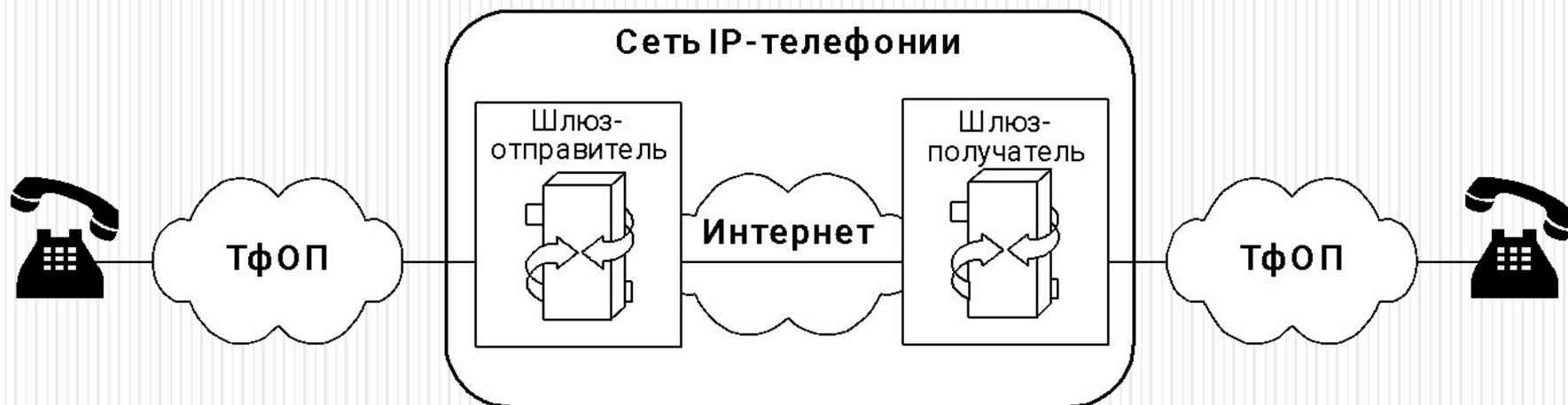
2-й уровень - до 400 мс - считается хорошим качеством связи. Если задержки постоянно удерживаются на верхней границе 2-го уровня (на 400 мс.), то не рекомендуется использовать эту связь для деловых переговоров.

3-й уровень - до 700 мс - считается приемлемым качеством связи для ведения неделовых переговоров. Такое качество связи возможно также при передаче пакетов по спутниковой связи.

Причины возникновения задержек и способы их уменьшения

- Влияние сети (сетевые узлы)
- Влияние операционной системы (DSP)
- Влияние джиттер-буфера (динамическая длина буфера)
- Влияние кодека (выбор кодека)

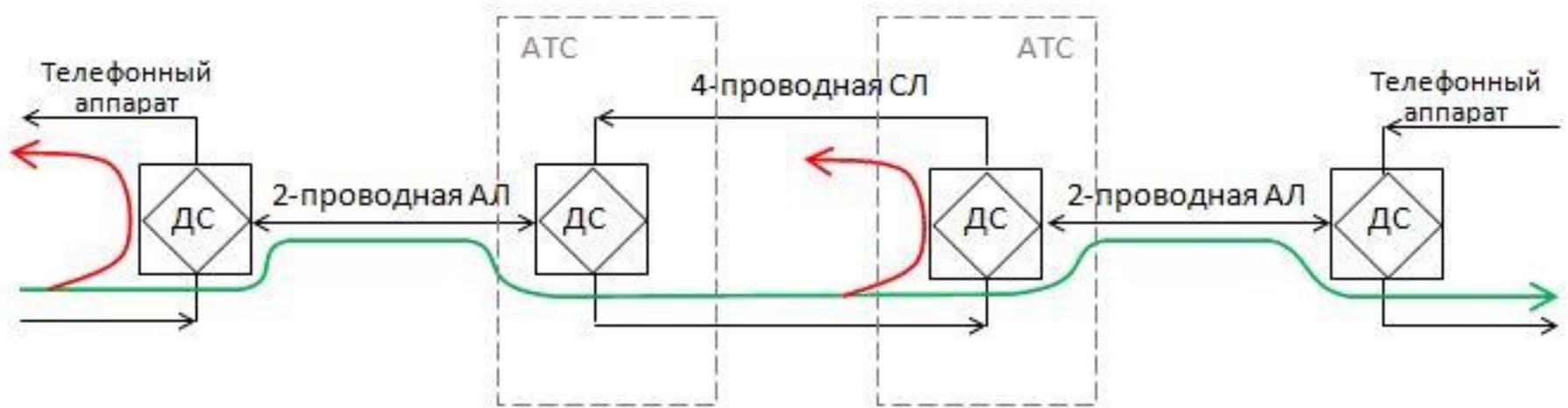
Вариация задержки (джиттер)



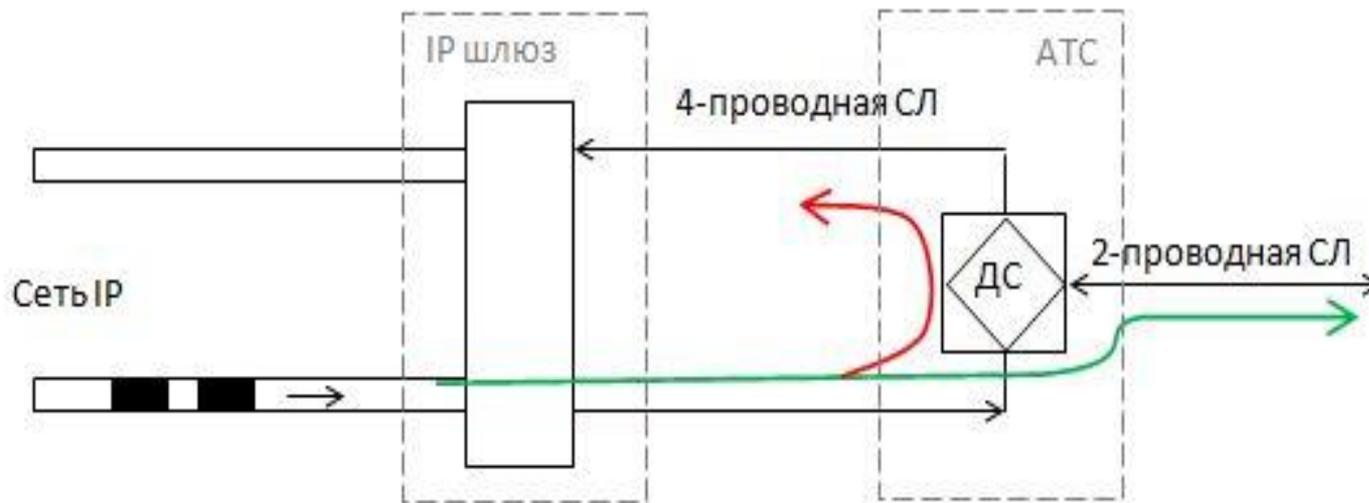
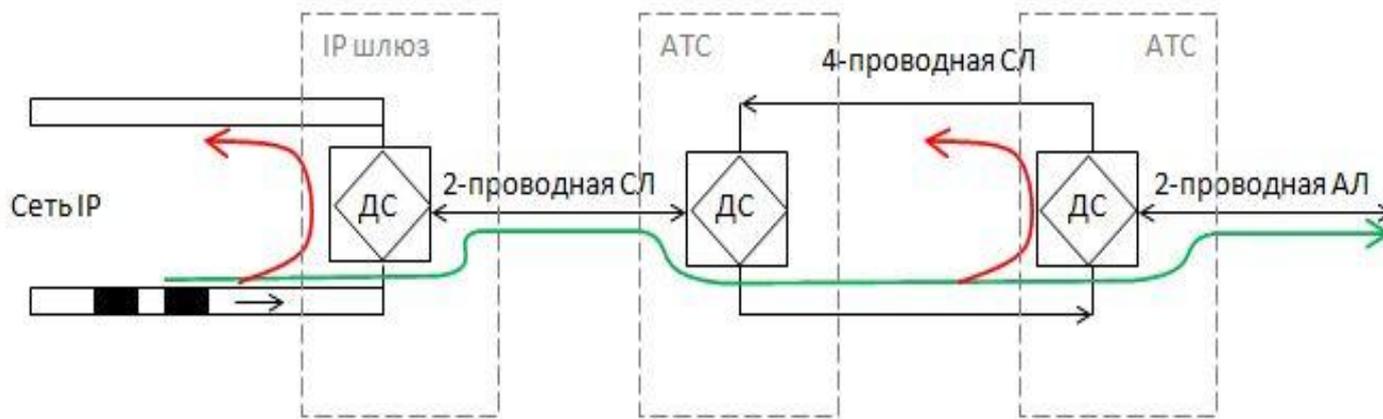
Причины возникновения и ограничение эффектов эха

- Причины возникновения эха
 - Электрическая природа эхо (отражения в дифсистеме)
 - Акустическая природа эхо
- Устройства ограничения эффектов эха
 - Эхозаградители отключают канал передачи, когда в канале приема присутствует речевой сигнал
 - Эхокомпенсаторы моделируют эхосигнал для последующего его вычитания из принимаемого сигнала

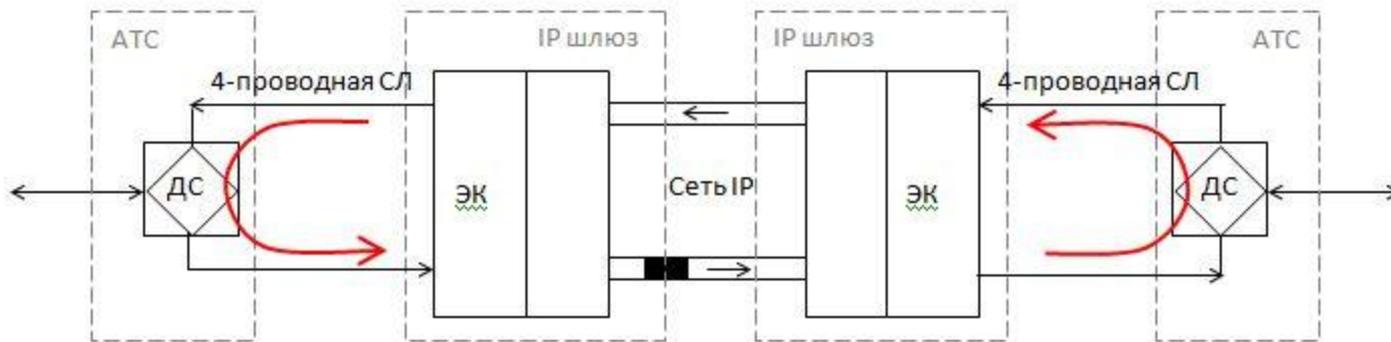
Возникновение эха в традиционных телефонных сетях.



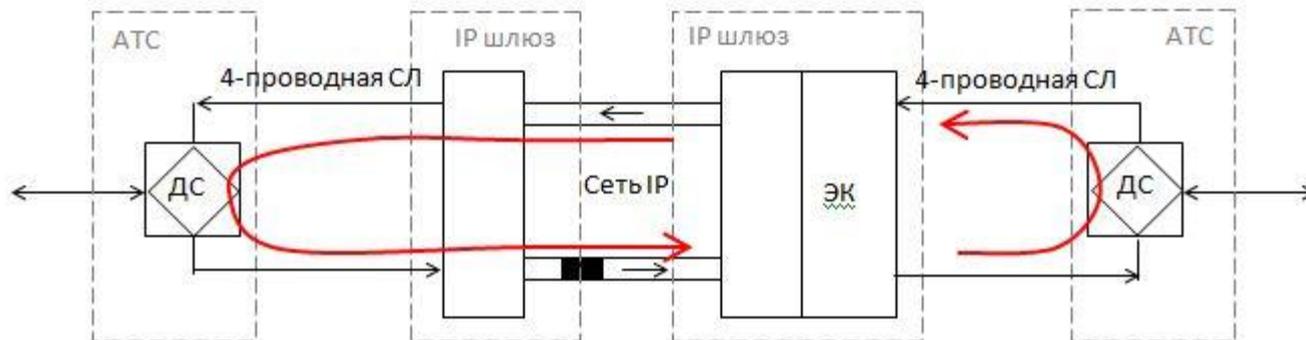
Возможные отражения электрического сигнала



Способы установки эхокомпенсаторов



«предпочтительный способ»



«централизованный способ»

Использование полосы пропускания

□ Подавление периодов молчания (VAD, CNG, DTX)

- ✓ Детектор речевой активности (Voice Activity Detector)
- ✓ Поддержка прерывистой передачи (Discontinuous Transmission)
- ✓ Генератор комфортного шума (Comfort Noise Generator)

VAD (Voice Activity Detector)

- Необходим для определения периодов времени, когда пользователь говорит
- Оценивает энергию входного сигнала и, если она превышает некоторый порог, активизирует передачу
- Чтобы избежать клиппирования начальной части периода активности, необходимо сохранение в памяти нескольких миллисекунд информации, чтобы иметь возможность запустить передачу до начала периода активности

Прерывистая передача

Discontinuous Transmission, DTX

- Позволяет кодеку прекратить передачу пакетов в тот момент, когда VAD обнаружил период молчания
- Некоторые кодеры не прекращают передачу полностью, а переходят в режим передачи гораздо меньшего объема информации, нужной для того, чтобы декодер на удаленном конце мог восстановить фоновый шум

Генератор комфортного шума

Comfort Noise Generator, GNG

- Генерация фонового шума
- Простейшие кодеки просто прекращают передачу в период молчания, и декодер генерирует какой-либо шум с уровнем, равным минимальному уровню, отмеченному в период речевой активности
- Более совершенные кодеки предоставляют удаленному декодеру информацию для восстановления шума с параметрами, близкими к фактически наблюдавшимся

Оценка MOS

- Рекомендация ITU-T P.800, которая описывает способы получения численных оценок качества речевой информации в сети, участке сети или отдельном оборудовании.
- ***MOS (Mean Opinion Score)*** – средняя оценка качества большой группы слушателей по 5-ти бальной шкале. Для прослушивания экспертам предлагаются различные звуковые фрагменты:
 - Речь
 - Музыка
 - Речь на фоне различного шума
 - И т.д.

MOS (продолжение)

Оценки интерпретируются следующим образом:

- **4-5** – высокое качество, аналогично качеству передачи речи в ISDN или еще выше
- **3,5-4** – качество ТфОП (toll quality)
- **3-3,5** – качество речи по прежнему удовлетворительно, но его ухудшение хорошо заметно на слух
- **2,5-3** - речь разборчива, но требует концентрации внимания для понимания

Недостатки оценки MOS

- Субъективность оценки
- Не учитывается влияние:
 - Задержки
 - Вариации задержки
 - Потери пакетов
- Нет возможности производить оценку на этапе планирования

Объективный метод оценки качества

- Рекомендация G.107
- В основе метода - E-модель - вычислительная модель, которая на основе более чем 20 параметров терминалов, линий связи, оборудования и условий разговора определяет оценку качества, так называемый R-фактор

E - модель

- В основе E-модели лежит принцип аддитивности факторов, ухудшающих качество передаваемой речи.
- R-фактор вычисляется по следующей формуле:

$$R = R_0 - I_s - I_d - I_{e-eff} + A$$

- Фактор R_0 представляет собой базовое отношение сигнал-шум, включающее в себя шумы от различных источников.

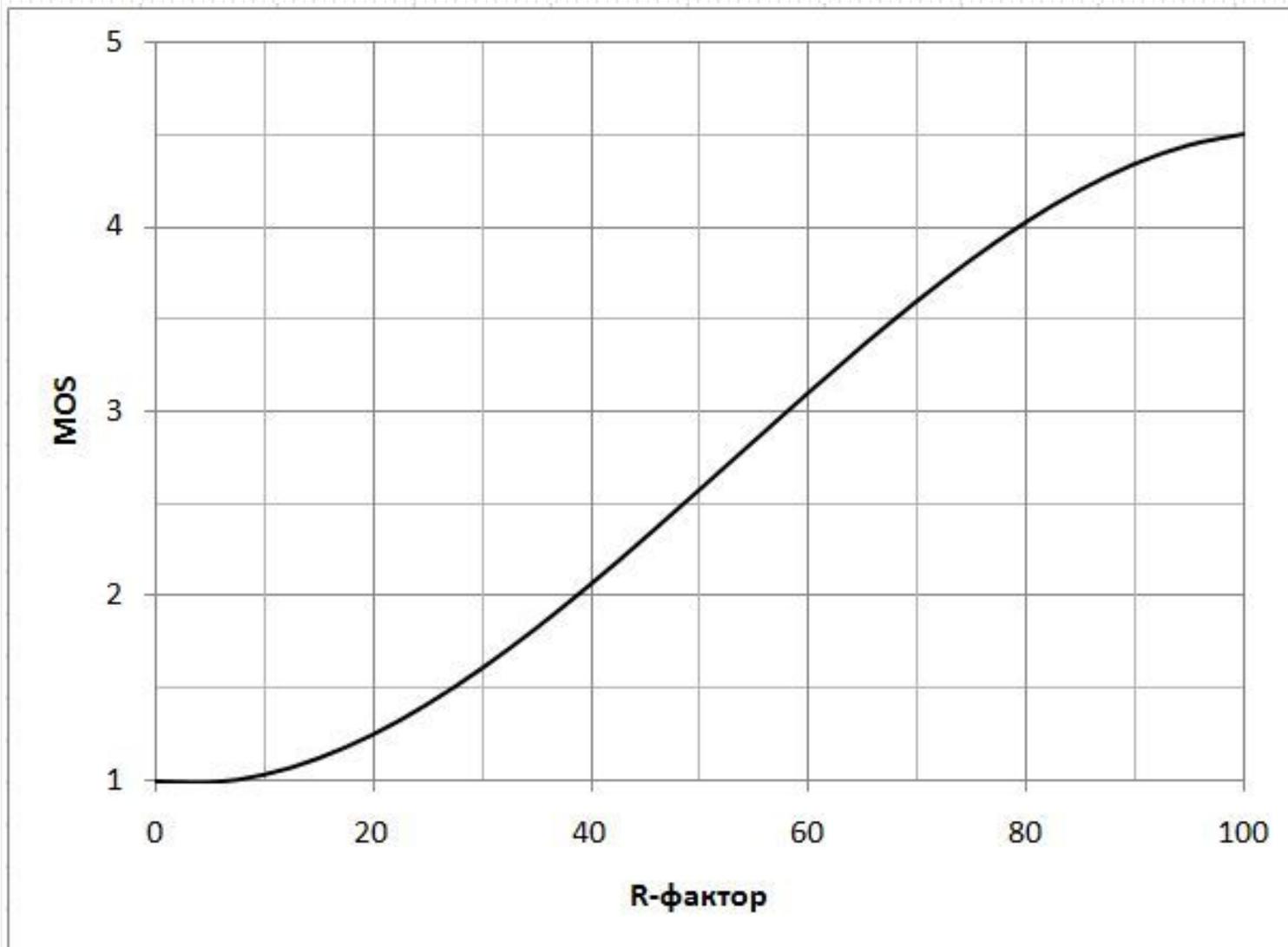
E-модель

- I_s отражает ухудшение качества из-за факторов, которые могут произойти более или менее одновременно с передачей речи (шумы квантования, неоптимальный местный эффект, тихий звук)
- I_d выражает ухудшение качества речи, вызванное большой задержкой и эффектами эха
- Коэффициент снижения эффективности оборудования I_{e-eff} отражает ухудшения качества, вызванные кодеками и потерей пакетов.

E-модель

- Коэффициент выигрыша A служит для компенсации коэффициентов ухудшений. Если пользователю предоставляются дополнительные преимущества можно увеличить значение R -фактора, так как пользователь может ожидать плохого качества от данного вида связи и готов с ним мириться в обмен на предоставляемые преимущества.
- Значения R -фактора лежат в пределах от 0 до 93,2, чем больше это значение, тем лучше качество передачи речи.

Соотношение между оценкой MOS и R-фактором



Передача сигналов DTMF и FAX по IP-сетям

- Обнаружение на передающем конце
- При помощи протокола сигнализации (SIP, H.248): задержки
- При помощи RTP (RFC 2833 “RTP Payload for DTMF Digits, Telephony Tones and Telephony Signals”): реальное время

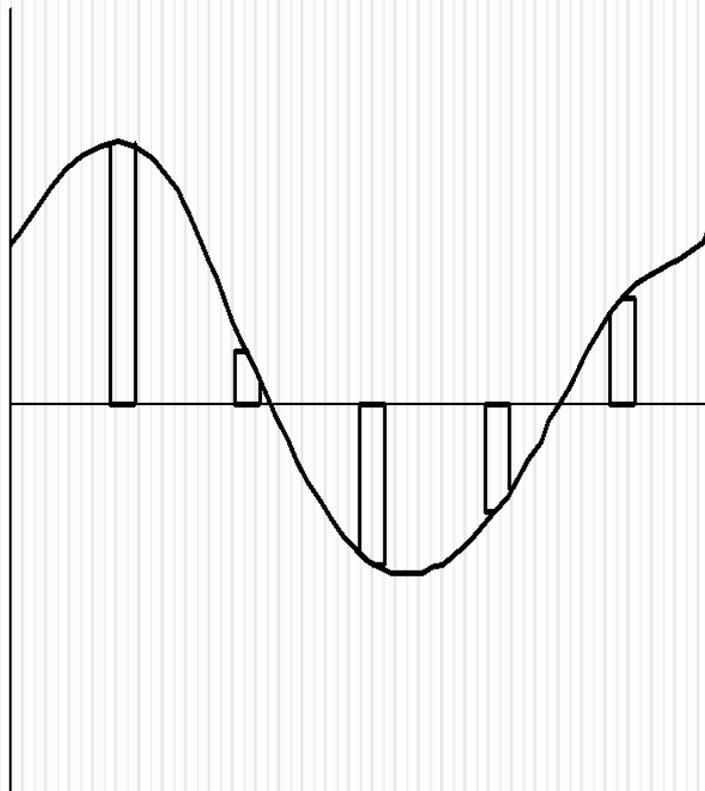
- FAX:
 - *Рекомендация T.37: отложенная доставка факсимильной информации (Store & Forward Fax)*
 - *Рекомендация T.38: доставка факсимильной информации в реальном времени (Fax Relay)*

Кодеки

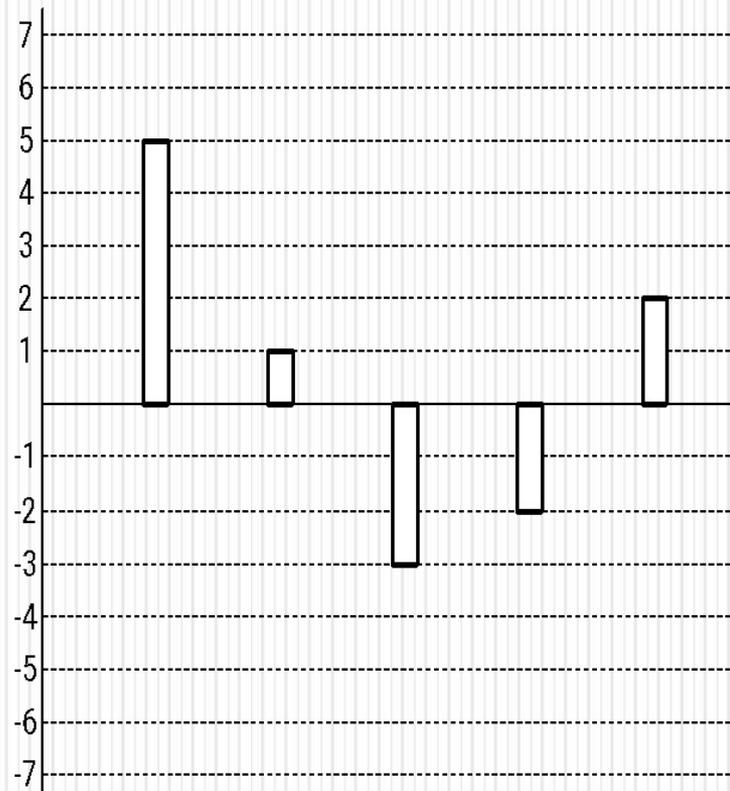
Преобразование речевого сигнала

- Дискретизация
- Квантование амплитуды
- Кодирование квантованных амплитуд

сигнала



Дискретизация



Квантование

Дискретизация

- аналоговый сигнал превращается в последовательность импульсов (отсчетов), величина которых равна амплитуде аналогового сигнала в определенные моменты времени
- Теорема Котельникова:
Аналоговый сигнал может быть восстановлен из последовательности отсчетов с частотой f , которая больше или равна умноженной на 2 максимальной частоте, используемой в спектре сигнала F :

$$f \geq 2F$$

Дискретизация

- Полоса частот речевого сигнала равна 0,3-3,4 кГц, поэтому частота дискретизации аналогового сигнала была выбрана равной 8 кГц (с запасом).
- Эта частота дискретизации является общемировым стандартом .

Квантование

- непрерывная величина амплитуды сигнала накладывается на дискретную шкалу квантования и округляется до ближайшего меньшего значения
- используется нелинейная (логарифмическая) шкала квантования, то есть квантованию подвергается не амплитуда сигнала, а ее логарифм
- 2 закона квантования:
 - **A-закон** (применяется в европейских странах и России)
 - **μ-закон** (применяется в Северной Америке)

Квантование

- Оба закона позволяют закодировать каждый отсчет сигнала 8 битами.
- Дискретизация по времени происходит с частотой 8 кГц, то есть отсчеты генерируются каждые 125 мкс, пропускная способность, необходимая для передачи речи в цифровом виде равна:

$$\frac{8 \text{ бит}}{125 \times 10^{-6} \text{ сек}} = 64 \quad /$$

Кодирование

- Сжатие оцифрованных отсчетов до минимально возможного числа двоичных битов в секунду
- Выполняется после дискретизации и квантования сигнала
- Существенно снижает нагрузку на сеть

Кодирование

- Значения соседних отсчетов как правило мало отличаются одно от другого, что позволяет с довольно высокой точностью предсказывать значение любого отсчета на основе значений нескольких предшествующих ему отсчетов.
- Эта закономерность используется двумя способами:
 - Изменение параметров квантования в зависимости от характера сигнала
 - Линейное предсказание (дифференциальное кодирование)

Линейное предсказание

- Кодируется разность между входным сигналом и «предсказанной» величиной, вычисленной на основе нескольких предыдущих значений сигнала
- Пусть:
 - отсчеты входного сигнала
 - a_i коэффициенты предсказания
- Предсказанное значение в момент времени i представляет собой линейную комбинацию нескольких предыдущих отсчетов:

$$\hat{y}(i) = a_1 y(i-1) + a_2 y(i-2) + \dots + a_p y(i-p)$$

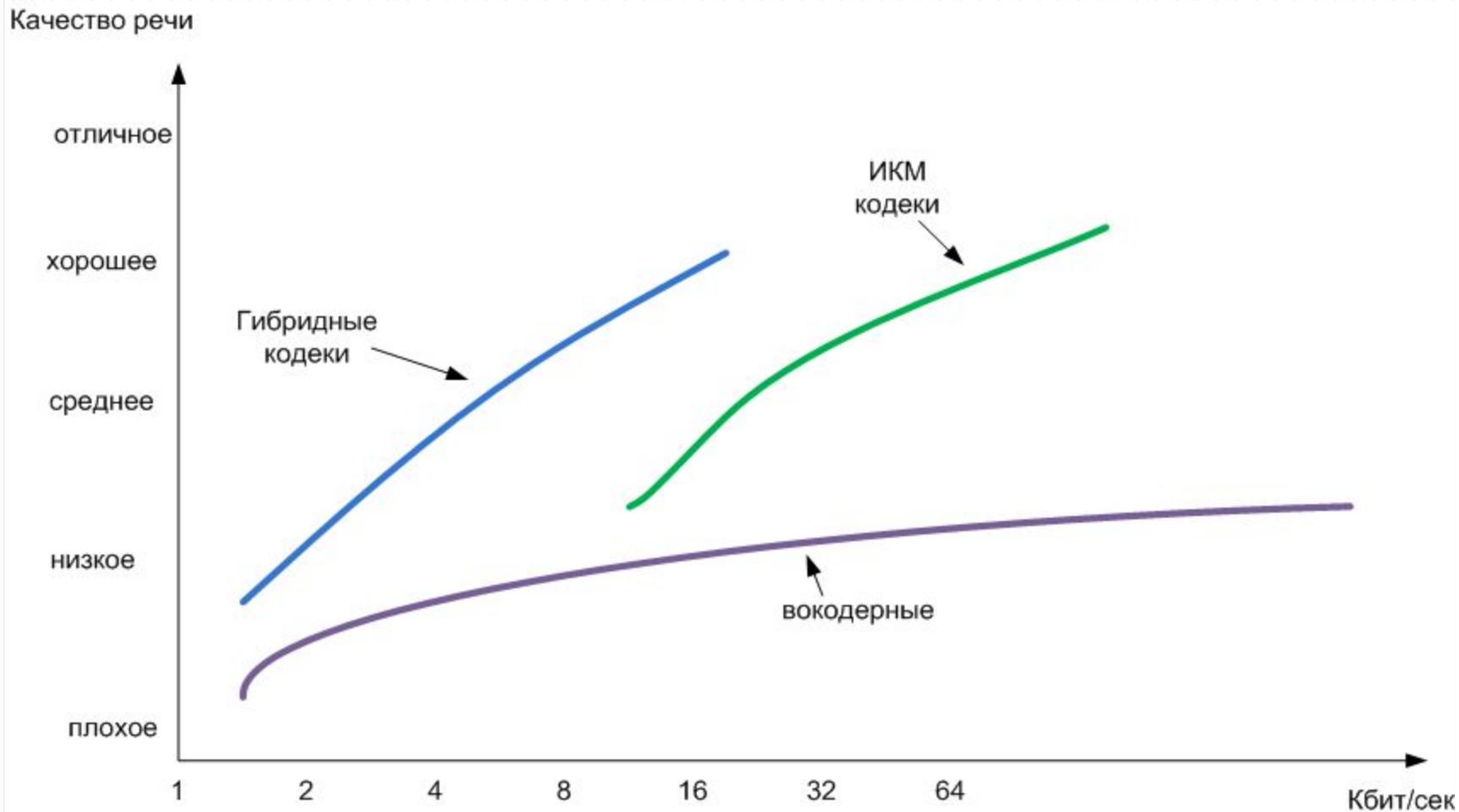
Линейное предсказание

- Разность $e(i) = y(i) - \hat{y}(i)$ имеет меньший динамический диапазон и может кодироваться меньшим числом битов, что позволяет снизить требования к полосе пропускания

Кодеки

- Кодеки с ИКМ и адаптивной дифференциальной ИКМ, появившиеся в конце 50-х годов и использующиеся сегодня в системах традиционной телефонии. В большинстве случаев представляют сочетание АЦП/ЦАП
- Кодеки с вокодерным преобразованием речевого сигнала. Используют принцип гармонического синтеза сигнала на основе информации о его вокальных составляющих – фонемах
- Комбинированные (гибридные) кодеки, которые сочетают в себе технологию вокодерного преобразования/синтеза речи, но оперируют уже с цифровым сигналом специализированных DSP

Усредненная субъективная оценка качества кодирования речи для различных типов кодеков



Критерии выбора кодека:

- Качество речи
- Использование полосы пропускания канала
- Размер кадра
- Чувствительность к потерям кадров

Стандартизированные кодеки

□ *Кодек G.711*: скорость 64 Кбит/с, оценка MOS - 4.2

□ *Кодек G.723.1*: предусмотрено 2 режима работы:

✓ 6.4 Кбит/с, оценка составляет 3.9

✓ 5.3 Кбит/с, оценка составляет 3.7

□ *Кодек G.726*: скорость 40, 32, 24 или 16 Кбит/с, оценка - 4.3

□ *Кодек G.728*: скорости 16 Кбит/с, оценка - 4.3

□ *Кодек G.729*: скорость передачи 8 Кбит/с, оценка – 4,1

□ *Кодек GSM Full Rate*: скорость 13 Кбит/с, оценка – 3,7

Сравнительные характеристики кодеков

Кодек	Скорость потока, Кбит/сек.	Размер кадра, мс.	Сложность (загрузка процессора)
G.711	64	0,125	Низкая
G.728	16	0,625	Очень высокая
GSM FR	13	20	средняя
GSM EFR	12,2	20	Высокая
G.729.a	8	10	Высокая/очень высокая
G.723.1	5,3/6,4	30	высокая

Сравнение кодеков

Единицы MOS (Mean Opinion Score) или единицы рейтинга R (Quality Rating). Соединение с качеством $R < 50$ не рекомендуется МСЭ-Т. Высшему качеству $R = 100$ соответствует $MOS = 4,5$.

