

Építsünk IP telefont!

Moldován István
Sonkoly Balázs

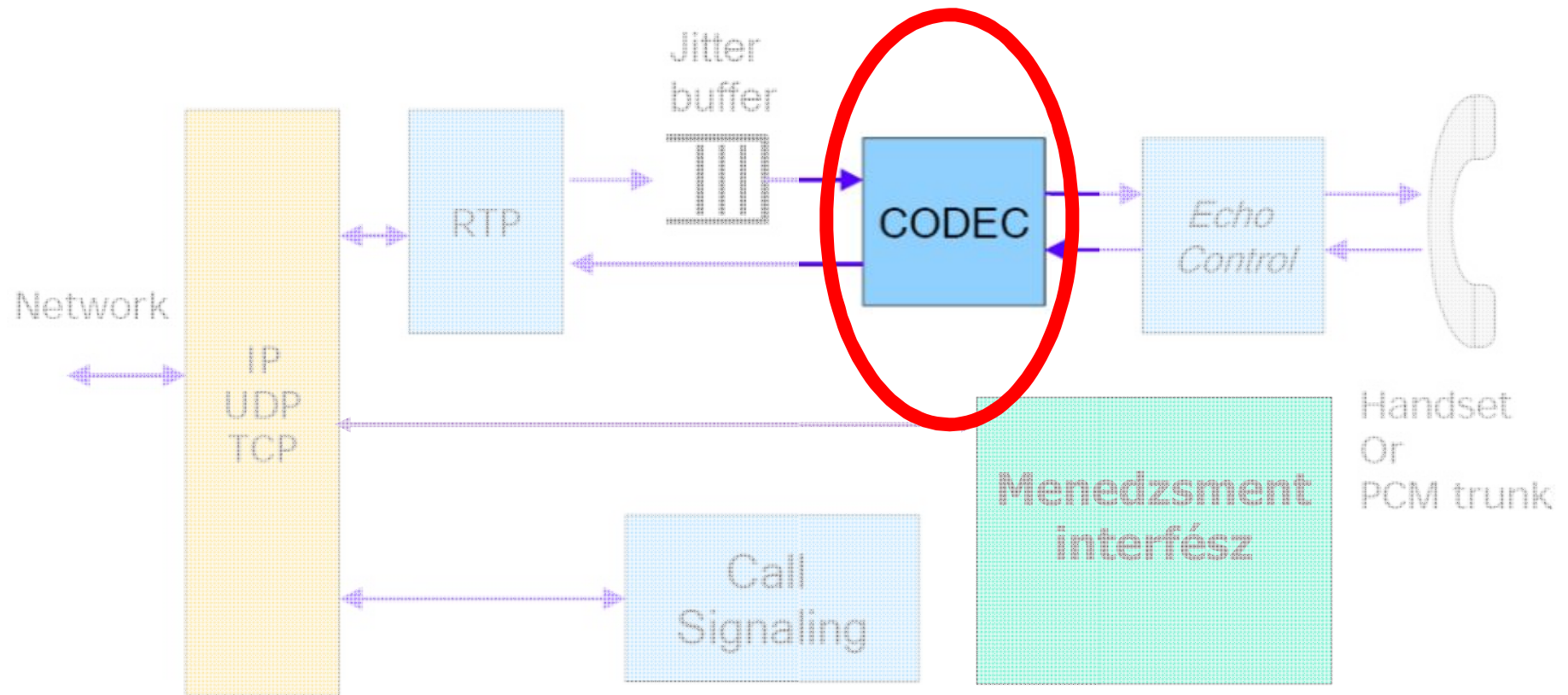


BUDAPESTI MŰSZAKI ÉS GAZDASÁGTUDOMÁNYI EGYETEM
TÁVKÖZLÉSI ÉS MÉDIAINFORMATIKAI TANSZÉK

Egy IP telefon felépítése



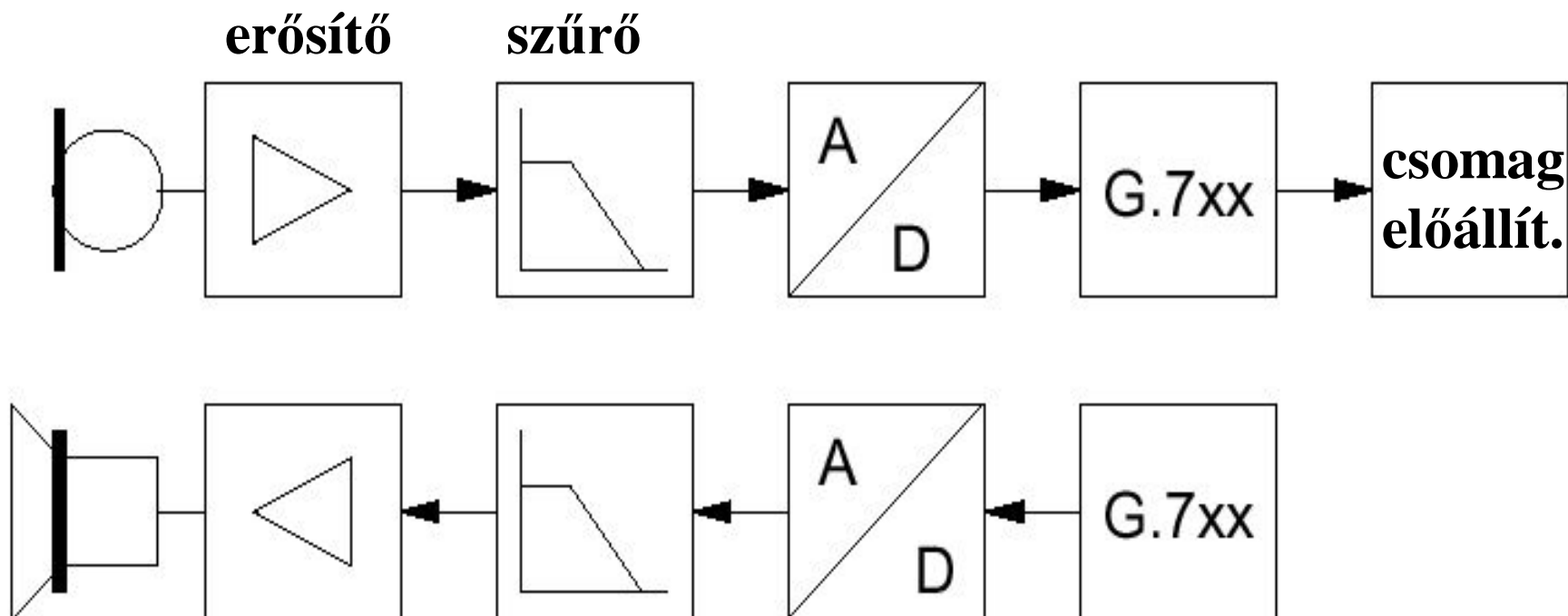
BME-TMIT



A digitális hang



BME-TMIT



A digitális hang minősége



BME-TMIT

Hang minőség különböző mintavételezés esetén

Elnevezés	frekvencia	mintavételezés	AD/DA bits	Alkalmazás
Telefon	300-3400 Hz	8 kHz	12–13	PSTN
Széles sávú	50-7000 Hz	16 kHz	14–15	konferencia
high-quality	30-15000 Hz	32 kHz	16	FM, TV
	20-20000 Hz	44.1 kHz	16	CD
	10-22000 Hz	48 kHz	24	pro-audio

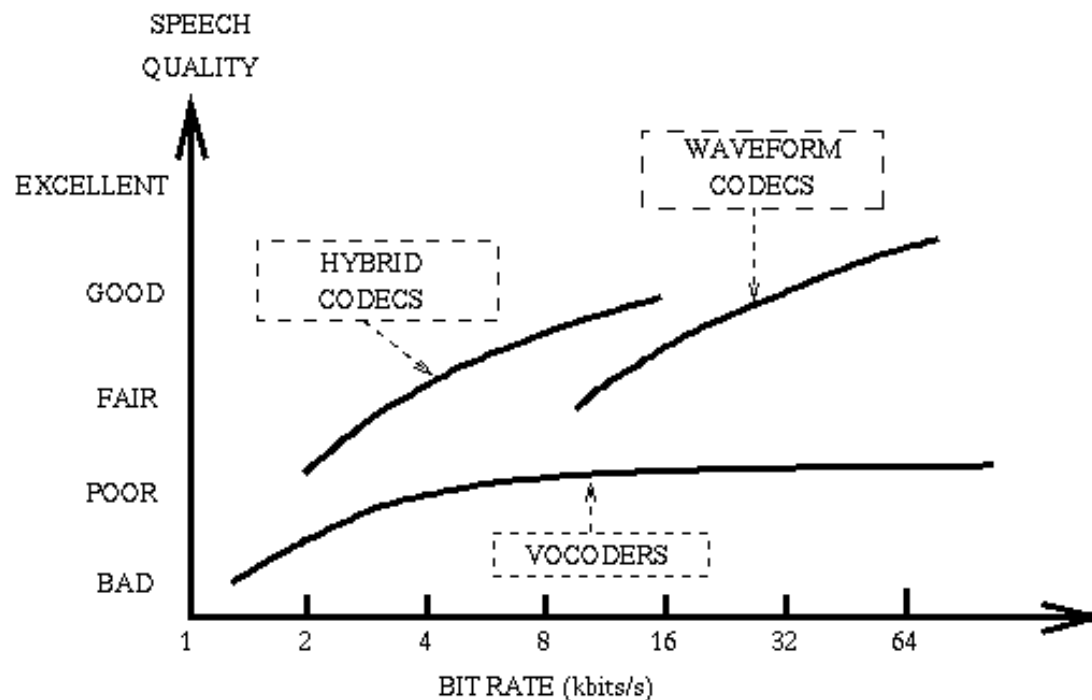
- Audio kódolási lehetőségek:
 - Compadding: logaritmikus skála szerinti (pl. u-law)
 - Waveform: a minták közti korrelációt használja ki (ADPCM)
 - Modell alapú: hang modellezés, paraméterek kinyerése (vocoderek)
 - Subband: a bemenő jelet alsávokra bontja, majd ezeket kódolja (mp3)
- Az újabb kódolási eljárások kihasználják az emberi hallás tökéletlenségét

Hibrid kódolók



BME-TMIT

- Megpróbálja a legjobbat felhasználni a módszerekből
 - Waveform kódolás
 - + modell alapú kódolás
 - Egész jó minőség alacsony sávszélességigénnyel



Kódolók kiértékelése



BME-TMIT

- Paraméterek
 - bitsebesség
 - minőség
 - Késleltetés: algoritmikus, feldolgozás
 - Csomagvesztésre való érzékenység
 - Komplexitás: MIPS, lebegővesszős implementáció, kódolás/dekódolás
 - Tandem használat: több egymásutáni kódolás
 - Nem hang átvitele: DTMF hangok, zene, fax, stb.

Minőségi metrikák

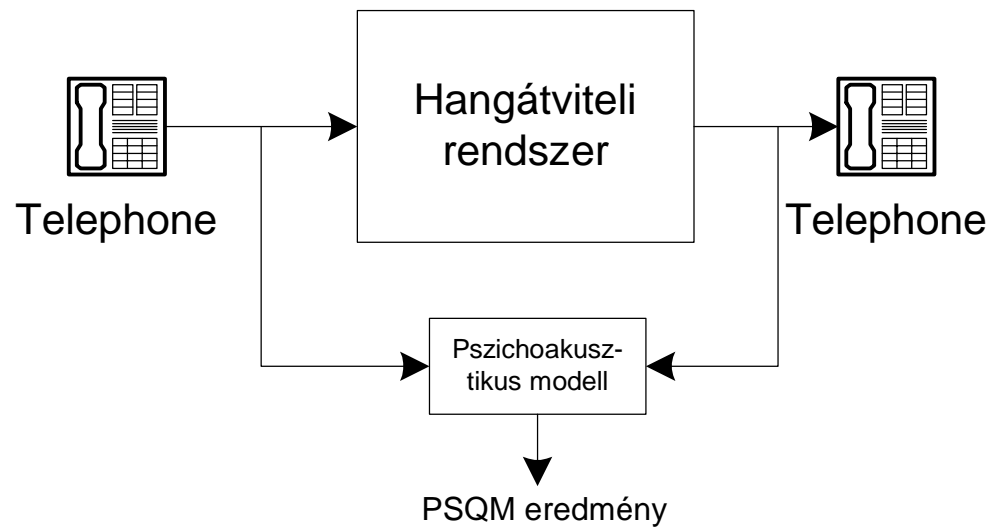


BME-TMIT

- Szubjektív illetve objektív
- MOS: szubjektív
 - 0-5 terjedő pontozás, legalább 30 résztvevő
 - MOS > 4 : "Toll Quality"
 - Az ITU P.800 ajánlása
- PSQM, PESQ: objektív metrikák
 - Olcsóbb módszer
 - Főleg eszköz-minőség ellenőrzésére használják

- PSQM – Perceptual Speech Quality Measurement
 - ITU-T P.861
 - Tükrözi az emberi döntést
 - Algoritmikus összehasonlítása az átvitt és eredeti jelnek
 - Figyelembe veszi a beszélő típusát, hangerőt, késleltetést, környezeti zajt, aktív/csend kereteket, vágást

- PSQM működés



- PSQM+

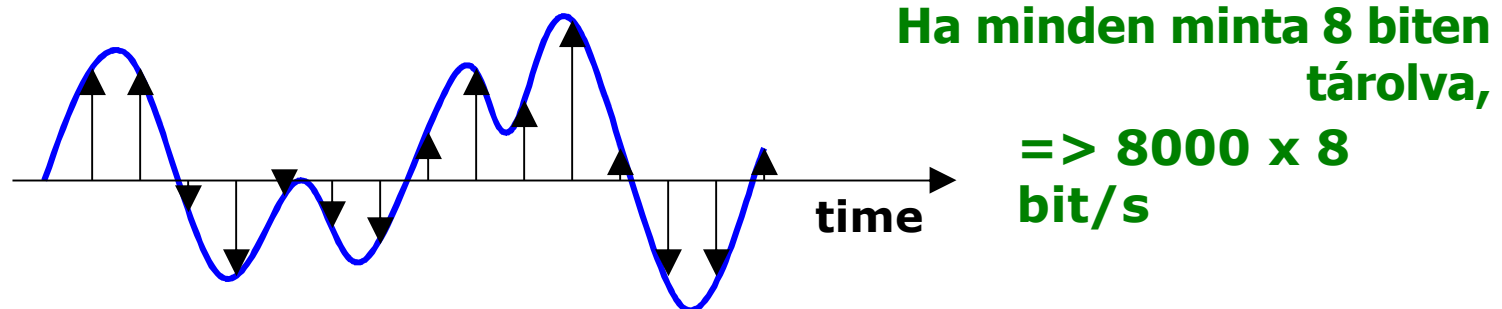
- Jobban közelíti a MOS értékét

Mintavételezés



BME-TMIT

- A-D átalakítás
 - Diszkrét minták, minden minta a hang egy pillanatnyi értékét tárolja valamennyi biten (8, 12, 16)
 - A jel visszaalakítható ha legalább kétszeres mintavételezéssel digitalizálták
- Az emberi beszéd
 - 300-3800 Hz
 - 8000 minta per másodperc



**Ha minden minta 8 biten
tárolva,**

**=> 8000 x 8
bit/s**

- A pillanatnyi értéket valamennyi biten tároljuk
- Kvantálási hiba
 - A különbség az eltárolt érték és az analóg jel értéke között
 - Több biten tárolva kisebb a hiba
 - Uniform kvantálással a hangosabb beszélők jobban érthetőek
 - Nem uniform kvantálással egyenletesebben oszlik el a jel-zaj viszony
- Általában 12 bites mintákat használnak
 - 16 biten tárolva

VAD, CNG



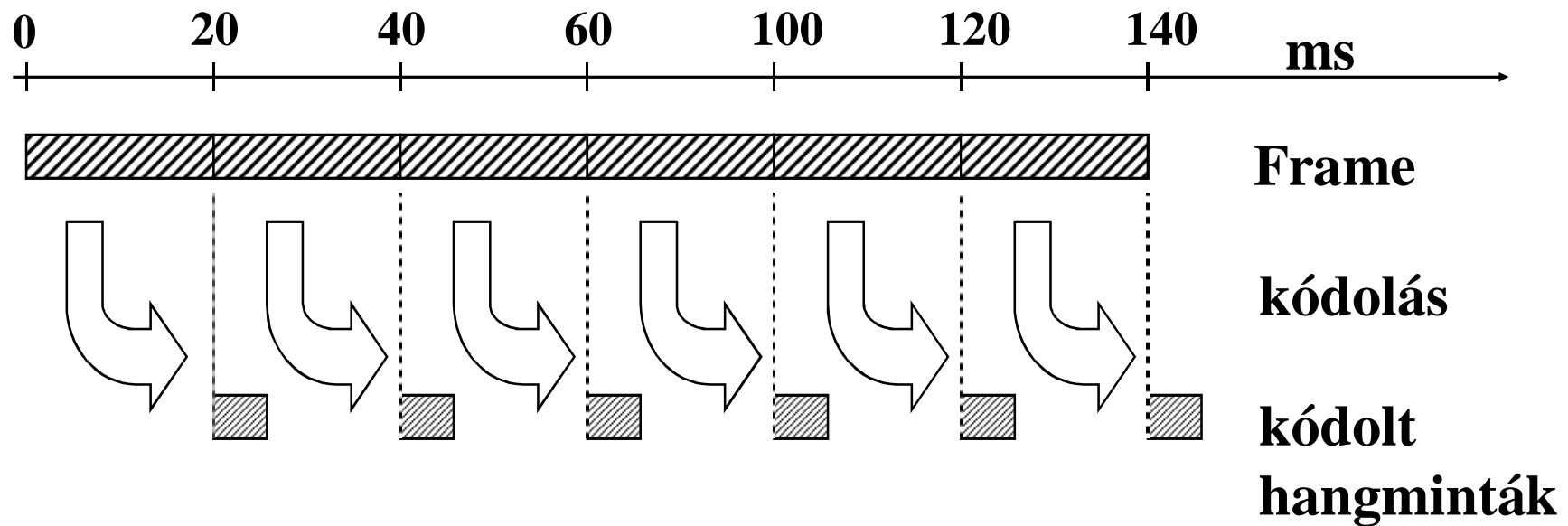
BME-TMIT

- Voice Activity Detection
 - Beszéd szünetekben nem küld adatot
 - Kevesebb sávszélességet igényel, nem terheli fölöslegesen a hálózatot
 - Zavaró lehet
- Comfort Noise Generator
 - Zavaró ha a teljesen süket a vonal, ezért zajt generál
 - Megpróbálja a készülék zaját utánozni
 - Bizonyos kódolók támogatják – zaj leírók

Kódolás



BME-TMIT



- GSM kódolás:
 - Frame: 8kHz mintavétel, 16 bites minták: 320 byte
 - GSM: 22 byte

Kódolók paraméterei



BME-TMIT

- Általában fix sáv szélességet használnak
 - Keret méret - kódolási eljárástól függően a kódoló fix méretű hangmintát képes elkódolni
 - Pl: GSM – 20 ms, G.729 – 10 ms, G.723.1 – 30 ms
 - A kódolt hangminta mérete is fix:
 - Pl: GSM – 22byte, G.729 – 10byte, G.723.1 – 24byte
 - Egy csomag több keretet tartalmazhat
 - Nagyobb késleltetés, kisebb overhead
- Algoritmikus komplexitás
 - Késleltetés
- Minőség: MOS



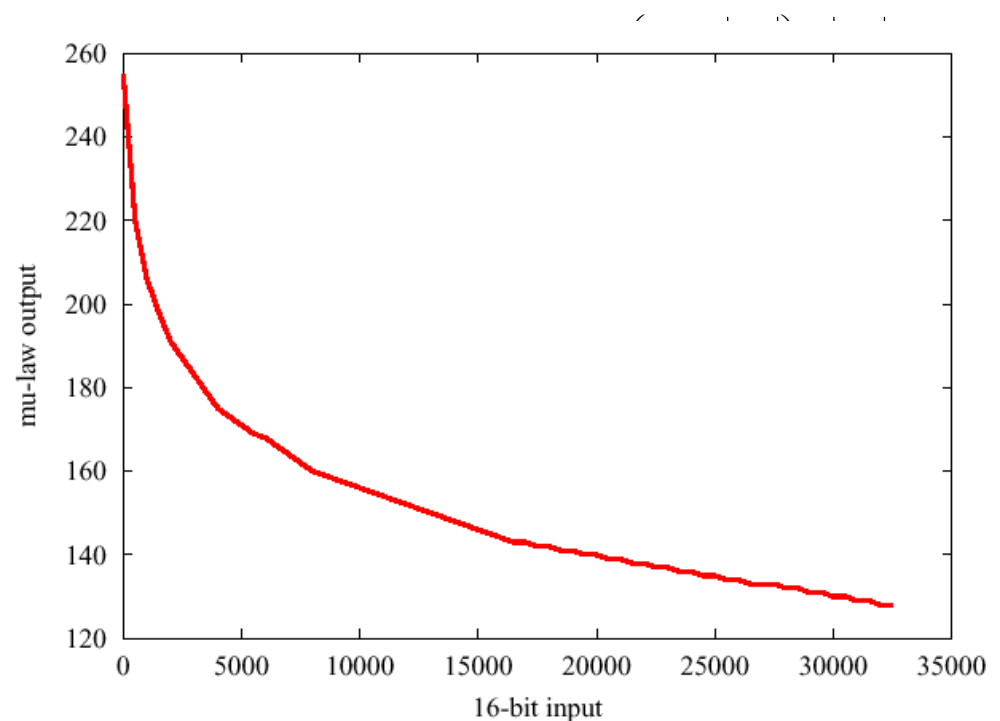
A fontosabb kódolók

G.711



BME-TMIT

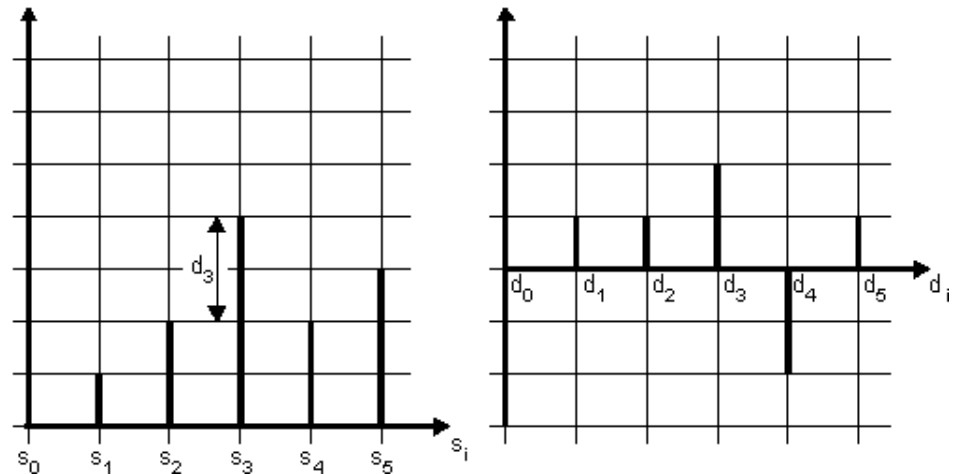
- A legközönségesebb kódoló
 - A vonalkapcsolt telefonhálózatokban elterjedt
 - PCM, Pulse-Code Modulation
- Uniform kvantálás
 - $12 \text{ bits} * 8 \text{ k/sec} = 96 \text{ kbps}$
- Non-uniform kvantálás
 - 64 kbps
 - mu-law
 - Amerikában
 - A-law
 - Európa, kicsit barátságosabb a halk jelekkel
 - MOS - 4.3



DPCM



- DPCM, Differenciális PCM
 - Csak az aktuális jel és az előző minta közti eltérést tárolja
 - A hang relatív lassan változik
 - Lehetséges előre prediktálni a következő minta értékét
 - A vevő ugyanazzal az algoritmussal prediktál
 - A legegyszerűbb módszer
 - Nincs predikció
- Nincs algoritmikus késleltetés
- Waveform kódoló



- ADPCM, Adaptív DPCM

- Prediktálja a következő minta értékét

- Az előző minták alapján

- Felhasznál információt a hang változási dinamikájáról

- A hibát kvantálja és kódolja

- Kevesebb bit szükséges

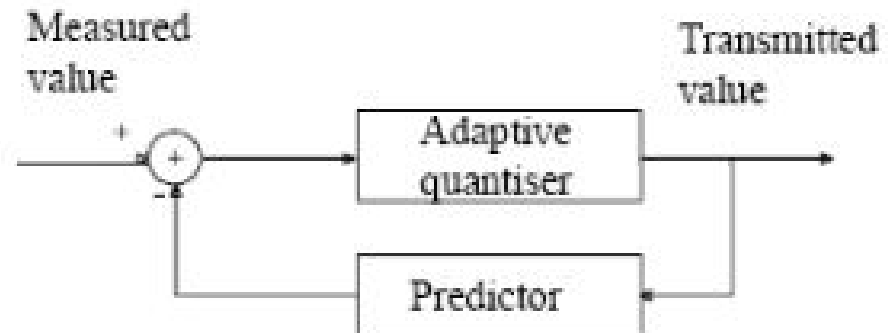
- G.721

- 32 kbps

- G.726

- A-law/mu-law PCM -> 16, 24, 32, 40 kbps

- A MOS körülbelül 4.0 - 32 kbps-en





- Hibrid kódoló

- A forrás modell és hullámforma közötti rést tölti be
- A legsikeresebb és legelterjedtebb
 - Nem egy egyszerű, beszéd/beszédszünet alapú
 - Az analóg jel mintáit előre meghatározott mintákkal hasonlítja
 - A legjobban találó minta kódját küldi el a vevőnek
 - MPE, Multi-Pulse Excited
 - RPE, Regular-Pulse Excited
 - CELP, Code-Excited Linear Predictive

G.728 LD-CELP



BME-TMIT

- CELP kódolók
 - Egy szűrő melynek karakterisztikái időben változnak
 - „codebook” – egy tár of akusztikus vektorokról
 - A vektor = egy szűrő leíró készlet
 - Küldés:
 - Szűrő paraméterek, nyereség, pointer a vektorra
- Low Delay CELP
 - Visszafele adaptív kódoló
 - Az előző mintákat használja a szűrő együtthatók meghatározására
 - Egyszerre 5 mintát használ
 - A késleltetés < 1 ms
 - Csak a vektor pointert küldi

G.723.1 ACELP



BME-TMIT

- 6.3 or 5.3 kbps
 - Minkettő kötelezően implementálandó
 - Bármikor átválthat egyikről a másikra
- A kódoló
 - Sávszélesség-limitált hang bemenet
 - 8 KHz mintavétel, 16-bit uniform PCM kvantálás
 - 240 mintát tartalmazó blokkokkal dolgozik
 - 7.5 ms előretekintést igényel
 - 37.5 ms algoritmikus és egyéb késleltetés
 - Nagy számításigényű



- Különböző műveleteket hajt végre a megfelelő szűrő együtthatók meghatározására
 - 5.3 kbps, Algebraic Code-Excited Linear Prediction
 - 6.3 kbps, Multi-pulse Maximum Likelihood Quantization
- Küldéskor:
 - Lineáris predikció együtthatói
 - nyereség
 - A „codebook” index
 - 24-byte keretek 6.3 kbps-on, 20-byte keretek 5.3 kbps-on



- G.723.1 Annex A
 - Silence Insertion Description (SID) keretek - 4 byte
- Az elküldött keret első bitjei:
 - 00 6.3kbps 24 byte/keret
 - 01 5.3kbps 20 byte/keret
 - 10 SID frame 4 byte/keret
- A MOS kb 3.8
 - Legalább 27.5 ms késleltetés!

G.729



BME-TMIT

- 8 kbps
- 10 ms-os kereteket használ, 80 minta 8 KHz mintavételezés
- 5 ms előrettekintés
 - 15 ms algoritmikus késleltetés
- Az eredmény egy 80-bit keret minden 10 ms hangra
- Komplex kódoló
 - G.729.A (Annex A), több ajánlás írja le
 - Ugyanaz a keret formátum
 - Encoder/decoder, G.729/G.729.A
 - Egy kicsit rosszabb minőség
- G.729 esetében a MOS = 4.0
- G.729A esetében a MOS = 3.7



- G.729.B
 - VAD, Voice Activity Detection
 - A bemenő jel paramétereit alapján meghatározva
 - Az aktuális + a két előző keret alapján
 - DTX, Discontinuous Transmission
 - Nem küld semmit vagy SID keretet küld
 - SID keret a megbecsült háttérzajt próbálja utánozni
 - CNG, Comfort Noise Generation
- G.729 Annex D
 - Kisebb sáv szélességű kiterjesztés
- G.729 Annex E
 - Nagyobb sáv szélességű kiterjesztés

- GSM Enhanced Full-Rate (EFR)
- A mobil telefonok által is használt kódoló
 - GSM 06.60
 - ACELP alapú kódoló
 - 12.2 kbps
 - 20 ms keretekkel dolgozik
 - Támogatja a nem-folyamatos küldési módot DTX (discontinuous transmission)
- Linux alapú implementációja elérhető

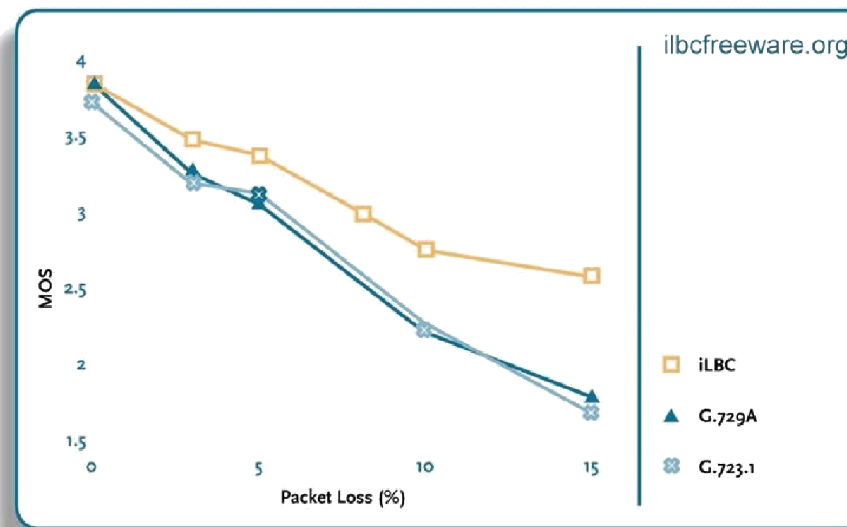
- Adaptive Multi-Rate (AMR) codec
 - 20 ms kódolási késleltetés
 - 8 különböző mód
 - 4.75 kbps-től 12.2 kbps-ig
 - 12.2 kbps = GSM EFR
 - 7.4 kbps = IS-641 (TDMA cellular system)
 - Bármikor válthat módot
 - Lehetővé teszi a DTX módot
 - The SID (Silence Descriptor) minden 8. keretben küldve, 5 byte hosszú
 - Sok 3G hálózat ezt a kódolót használja

Internet Low Bitrate Codec (iLBC)



BME-TMIT

- Ingyenes (nyílt) kódoló robusztus VoIP számára
- 13.33 kbit/s, 30 ms keretidő és 15.20 kbps 20 ms keretidővel
- A számításigénye megközelítőleg azonos a G.729A kódolóval



The tests were performed by Dynstat, Inc., an independent test laboratory.
Score system range: 1 = bad, 2 = poor, 3 = fair, 4 = good, 5 = excellent

DTMF kódok átvitele



BME-TMIT

- A hybrid kódolóknak emberi beszédre vannak optimalizálva
 - Egyéb jeleket is át kell vinni
 - Tónusok: fax, tárcsahang, foglalt-jelzés
 - DTMF jelek - kódok
- G.711 - OK
- G.723.1 és G.729 értelmezhetetlenné teheti
- A bejövő átjáró el kell kapja ezeket a jeleket
 - A tárcsahangot és a DTMF jeleket
 - Külső jelzési rendszer alkalmazása

DTMF kódok átvitele - 2



BME-TMIT

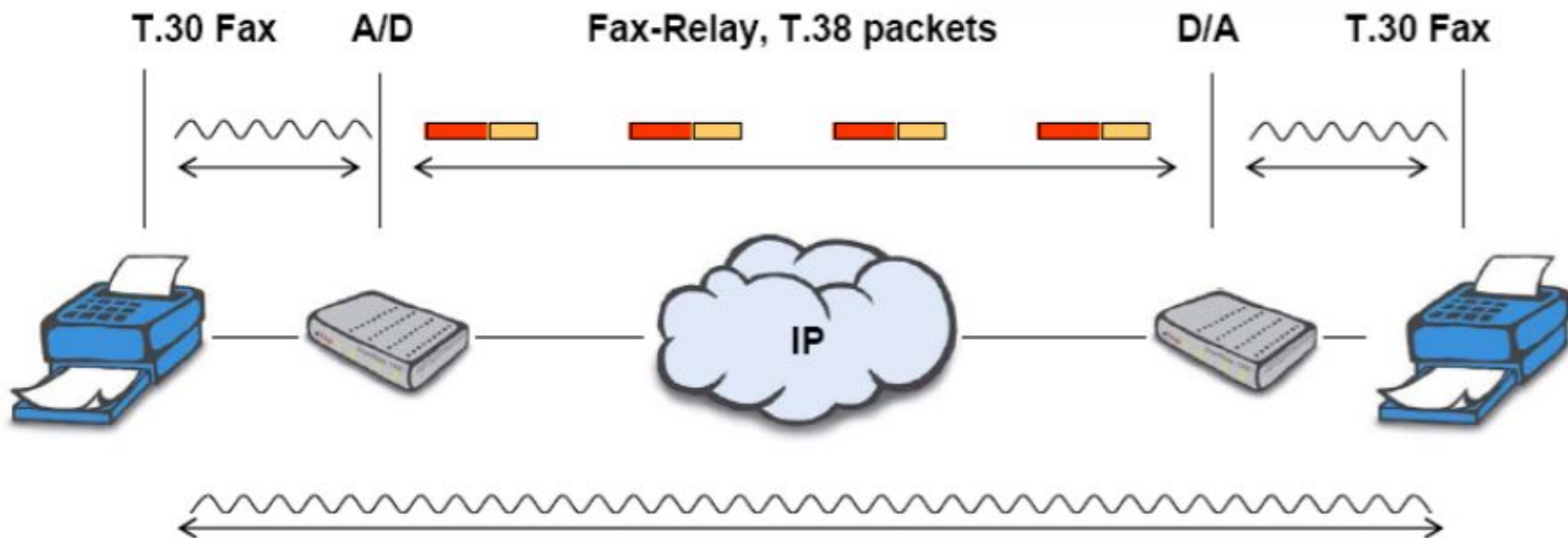
- A hívás elején könnyű
- Beszéd közben nehéz!
- Más kódolás szükséges
 - De ugyanazon a média folyamaton kell küldeni
 - Egy RTP csomagban küldve megjelölve
 - RFC 2198
 - RTP tartalom formátum leírás redundáns audió adat számára
 - Kétféle RTP formátumú csomagot küld

Fax átvitele



BME-TMIT

- **Fax - G711 hangként**
- G.711 kódolással átvihető normál hívásként
 - + kompatibilis minden átjáróval
 - Nagy sávszélességigény
 - Kevésbé robusztus
- **Fax relay, T.38**
- A fax hangok végződnek az első átjárónál és visszaállítódnak a túlóldalon
 - + kisebb sávszélességigény
 - + megbízhatóbb
 - Támogatás szükséges





VoIP-re jellemző kérdések, minőségbiztosítás

Kódolók sávszélességigénye



BME-TMIT

- Kódoló sávszélessége \neq VoIP hívás sávszélessége
- A sávszélesség a „dróton” jóval nagyobb, akár többszöröse!
- Számítása:
 - $1/\text{keretidő} * \text{csomagméret}$ másodpercenként
 - A csomagméret függ a hálózattól, Ethernet esetében:
 - 14 byte Ethernet fejléc +
 - 20 byte IP fejléc +
 - 8 byte UDP fejléc +
 - 12 byte RTP fejléc +
 - keret méret (pl. GSM EFR esetében 33)
 - Összesen 87 byte másodpercenként 50-szer =
34800bit/s

Sáv szélességigény - folytatás



BME-TMIT

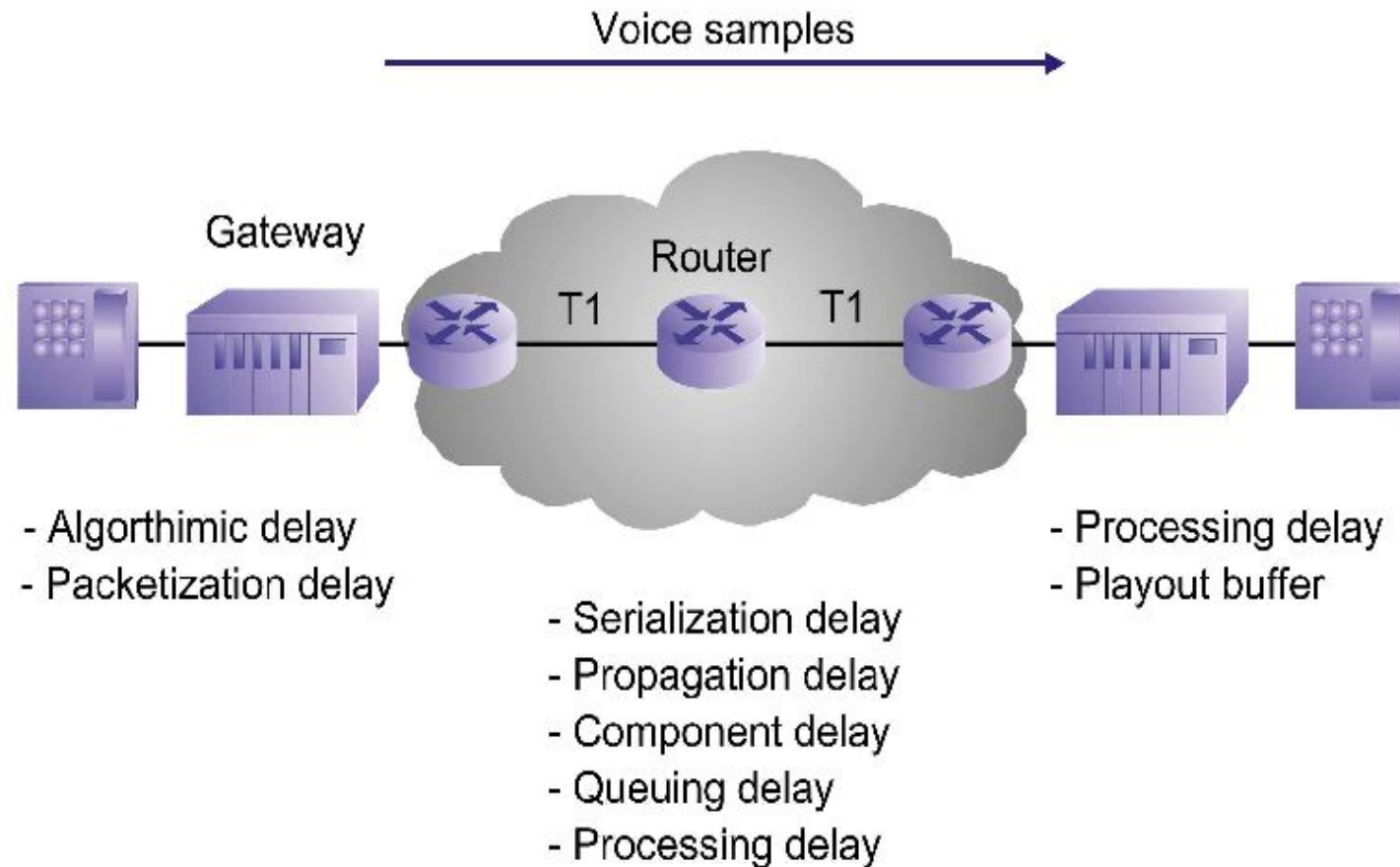
- Bizonyos kódolók esetében 1 keret csomagonként nem hatékony
- - nagy „overhead”
 - Bizonyos kódolók esetében 1 keret csomagonként nem hatékony
 - nagy „overhead”
 - Pl. G.729 – 10 ms keret = 10 byte, G.711...
 - Több keret/csomag hatékonyabb
 - Nagyobb a késleltetés...
- Általában 20 vagy 30 ms hang megy egy keretben

- Algoritmikus késleltetés
 - Kódoló függő
- Csomagolási késleltetés
- Soros késleltetés
- Terjedési késleltetés
- Hálózati komponensek késleltetése

A komponensek késleltetése



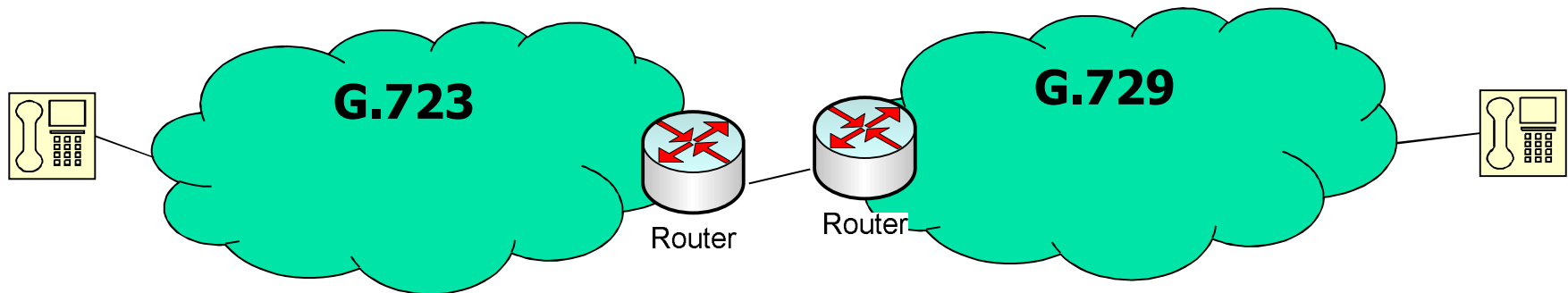
BME-TMIT



Tandem kódolás



BME-TMIT



- Transcoding
 - Dekódolás – újrakódolás más codec-el
 - Mindkét kódolás veszteséges ->a minőség romlik
- Tandem kódolás – megvizsgálandó hogy hogy viselkedik a két kódoló együtt

Kódolók - implementáció



BME-TMIT

- Általában optimalizált formában elérhető
 - Library (.lib)
- Állapot tárolás szükséges: kontextus mentés
- Használat:

- Inicializálás -> létrejön a kontextus

```
- void * E_IF_init (dtx);           // enkóder  
- void * D_IF_init (void);         // dekóder
```

- Encode/decode függvények hívása

```
- int E_IF_encode (Word16 mode, Word16 *speech, Uword8 *serial);  
- void D_IF_Decode (Uword8 *bits, Word16 *synth);
```

- Kontextus megszüntetése

```
- E_IF_exit ();           // enkóder  
- D_IF_exit ();          // dekóder
```

- Modern kódolók alkalmazásával a PCM forgalom 16x-osát lehet átvinni
 - T1/E1/J1 vonalkapcsolt hálózatokon VoIP technológiát alkalmazva lényegesen meg lehet növelni a hálózat kihasználtságát
- Van azért hátránya is
 - Gyengébb minőség (MOS)
 - Késleltetés és emiatt miatt visszhang
 - Megbízhatóság



Kérdések?

KÖSZÖNÖM A FIGYELMET!