

ABSTRACT

Dal Teorema di Shannon al VoIP: è questo un percorso che mostra come un risultato matematico astratto abbia permesso lo sviluppo di tecnologie che hanno modificato radicalmente gli odierni sistemi di comunicazione.

Oggetto dello studio è il segnale sonoro: questo viene analizzato da un punto di vista fisico ed espresso mediante funzioni matematiche che ne descrivono le equazioni d'onda che lo compongono. Attraverso il criterio fornito dal Teorema del Campionamento di Shannon lo stesso segnale viene digitalizzato: da funzione continua viene trasformato in somma di campioni estratti in precisi intervalli di tempo. In questo modo il segnale sonoro analogico iniziale diventa un “pacchetto di dati”, cioè una sequenza di bit che viene prima compressa e poi trasmessa secondo un rigido schema imposto dagli standard delle Telecomunicazioni, e sotto questa forma attraversa Internet: indipendentemente dalla distanza che deve percorrere arriva al destinatario, viene decodificata e riconvertita in segnale analogico. Questo è quanto avviene in una comunicazione telefonica VoIP.

L'aspetto fondamentale è che il segnale sonoro finale è identico a quello iniziale solo se il campionamento viene fatto seguendo il criterio indicato nel Teorema di Shannon:

un segnale, funzione continua del tempo, con banda W approssimativamente limitata ed approssimativamente limitato anche nel tempo (T), può essere ricostruito con un alto grado di accuratezza attraverso i suoi valori nei $2TW$ punti campione distanziati nel

tempo: $\frac{1}{2W}$.

PREMESSA

La storia della matematica è ricca di scoperte la cui enorme e rivoluzionaria utilità pratica si è rivelata solo decenni o addirittura secoli dopo.

Di certo Edmund Taylor Whittaker mentre nel 1915 pubblicava su “*On the Functions Which are Represented by the Expansions of the Interpolation Theory*” il suo studio sulle Funzioni Cardinali non immaginava che le future telecomunicazioni avrebbero avuto come base proprio quei suoi calcoli. La sua ricerca verteva infatti sulla possibilità di esprimere una funzione reale (non necessariamente continua) tramite una serie formale di valori della stessa e della funzione $\text{sinc}(x)$.

Claude Elwood Shannon quale matematico-ingegnere scoprì invece le potenzialità degli studi di Whittaker e li mise a frutto a partire dal 1940, quando costruì una completa teoria sul campionamento del segnale. Egli dimostrò che i segnali a banda limitata possono essere rappresentati tramite una serie, somma di valori discreti ricavati campionando i segnali stessi in precisi intervalli di tempo. Tali serie, sotto opportune ipotesi, convergono assolutamente alla funzione "segnale" originale.

Le applicazioni del Teorema di Shannon sono numerosissime in vari settori: il campionamento e la quantizzazione dell'audio e del video per la registrazione e la trasmissione (si pensi ai comunissimi CD e DVD e le programmazioni televisive digitali), le comunicazioni a distanza, utilizzando diversi metodi e canali trasmissivi, ad esempio la telefonia mobile (GSM, UMTS), e le comunicazioni via Internet.

Tra queste si trova il sistema VoIP, che nell'ultimo decennio ha cambiato le regole della telefonia fissa, ha abbattuto i costi della tradizionale PSTN convogliando su Internet il traffico voce, e sta oggi aprendo le porte alla sperimentazione della convergenza fisso-mobile.

INTRODUZIONE

Il VoIP

Voice over IP (*Voce tramite Protocollo Internet*), acronimo **VoIP**, è una tecnologia che rende possibile effettuare una conversazione telefonica sfruttando una connessione di rete che utilizza il protocollo IP, senza farla passare attraverso la rete telefonica tradizionale (**PSTN**: *Public Switched Telephon Network*). Quando uno degli utenti collegati parla, le informazioni vocali vengono codificate in forma digitale ed instradate sulla rete sotto forma di pacchetti di dati [8].

Schema della comunicazione

Tutto si basa su un principio matematico valido per le trasmissioni dei segnali a distanza: per trasmettere un generico segnale analogico lontano dal mittente è possibile utilizzare il formato digitale. Per prima cosa è necessario quindi digitalizzarlo con un **ADC** (*convertitore analogico-digitale*), trasmetterlo, e trasformarlo di nuovo in formato analogico con un **DAC** (*convertitore digitale-analogico*) per permettere al destinatario di utilizzarlo.

Il VoIP lavora proprio in questo modo, digitalizzano la voce in pacchetti, mandandoli in rete e riconvertendoli in voce una volta giunti a destinazione.

I vantaggi del formato digitale sono notevoli: si possono comprimere i dati, instradarli (utilissimo su Internet), convertirli nuovamente in un formato più consono al mezzo utilizzato e così via; il segnale digitale è inoltre meno influenzato dai disturbi rispetto a quello analogico.

Per la trasmissione usiamo le reti UDP/IP, che sono costituite dai pacchetti IP contenenti un'intestazione (per controllare la comunicazione) e dai dati.

Tutto questo funziona correttamente se il segnale (voce) che arriva a destinazione è identico a quello della sorgente. Qui interviene il **Teorema di Shannon**: la trasmissione avviene correttamente secondo lo schema seguente se il campionamento del segnale viene effettuato seguendo un preciso criterio.

Lo schema della comunicazione è il seguente [14]:

TRASMITTENTE	Convertitore A.D.	CANALE COMUNICATIVO	Invertitore D.A.	RICEVENTE
Manda il segnale continuo: suono, parole parlate	Il segnale viene trasformato in impulsi elettrici	Il media usato e attraversato dal segnale Es: Internet, linea telefonica	Processo inverso Riconversione in analogico	Arriva il segnale riconvertito in analogico

Tabella 1

Il VoIP presenta molti vantaggi immediati rispetto alla telefonia tradizionale:

- minore costo per chiamata, specialmente su lunghe distanze;
- minori costi delle infrastrutture: quando si è resa disponibile una rete IP nessun'altra infrastruttura è richiesta;
- si possono eliminare le centrali di commutazione ed economizzare sulla larghezza di banda occupata;
- utilizzo dello stesso canale per il trasporto di dati, video e voce;
- l'implementazione di future opzioni non richiederà la sostituzione dell'hardware.

Inoltre le conversazioni VoIP non devono necessariamente viaggiare su Internet, ma possono usare come mezzo trasmissivo una qualsiasi rete privata basata sul protocollo IP, per esempio una LAN all'interno di un edificio o di un gruppo di edifici. I protocolli che verranno analizzati in questa tesi e che vengono usati per codificare e trasmettere le conversazioni VoIP sono solitamente denominati *Voice over IP protocols*.

1. LA STORIA DEL TEOREMA

La teoria del campionamento così come la conosciamo ora ha circa una sessantina d'anni, anche se le basi su cui è fondata sono state tracciate nel passato da grandi matematici come Poisson, Borel, Hadamard.

Si concorda sul fatto che il primo matematico che scoprì il teorema del campionamento fu W.T. Whittaker nel 1915 nei suoi studi sulle Funzioni Cardinali, e fu in seguito diffuso dal suo secondogenito J.M. Whittaker, ma diversi altri matematici rivendicarono la paternità della scoperta, o comunque giunsero contemporaneamente allo stesso risultato, tra questi F.J.W. Whipple (5 anni prima di Whittaker), e un matematico giapponese, K. Ogura, nel 1920, che diede una dimostrazione semplice e rigorosa dello stesso risultato, usando il calcolo dei residui (analisi complessa) [14].

Il teorema di Whittaker afferma che se una funzione $f(t)$ non contiene frequenze più alte di $W/2$ cicli al secondo, è completamente determinata dai suoi valori $f(t_n)$ dove i punti $t_n = n/W$, $n=0, \pm 1, \pm 2, \dots$, sono distanziati da $1/W$ secondi ciascuno; tale funzione e può essere ricostruita dai suoi valori attraverso la formula [11] [12]:

$$f(t) = \sum_{n=-\infty}^{+\infty} f\left(\frac{n}{W}\right) \frac{\sin \pi(Wt-n)}{\pi(Wt-n)}. \quad (1.1)$$

Questo è un principio di base ampiamente usato in ingegneria: tutta l'informazione contenuta in un segnale è completamente ricavabile dai soli campioni (numerabili) ottenuti dai valori della funzione calcolata in momenti equidistanti tra loro. Oltretutto la conoscenza della banda di frequenza del segnale determina la minima frequenza alla quale il segnale deve essere campionato per essere ricostruito completamente. Questa minima banda è conosciuta come la frequenza di Nyquist, in riferimento a H. Nyquist che fu il primo a riconoscerne l'importanza in relazione alla telegrafia (1925).

Lo studio di Whittaker fu di tipo puramente matematico; egli non aveva in mente alcuna possibile applicazione, non usò mai termini come segnale o cicli o frequenze a banda limitata, ma termini matematici che equivalgono ad essi.

Fu **Claude E. Shannon** ad introdurre i risultati di Whittaker nella teoria dell'informazione qualche decennio dopo: i suoi scritti risalgono al 1940 sebbene le prime pubblicazioni si ebbero al termine della II Guerra Mondiale (precisamente negli

anni 1948-49). Con le sue due famose opere “*A Mathematical Theory of Communication*” e “*Communication in the Presence of Noise*”, egli ottenne parecchi riconoscimenti proprio per il suo contributo nella teoria della comunicazione, tant'è che viene oggi ricordato con l'appellativo di *padre della Teoria dell'Informazione*; fu Shannon a divulgare in questo contesto il lavoro fatto da Whittaker [14].

Profilo di Claude E. Shannon

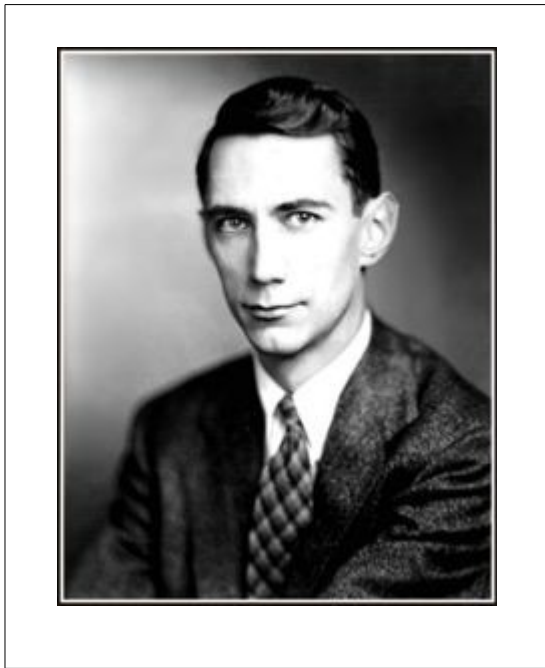


Fig. 1.1 Claude Elwood Shannon

Nacque nel Michigan il 30 Aprile 1916. Studiò matematica ed ingegneria e lavorò al MIT (Massachusetts Institute of Technology) e nei laboratori Bell; dimostrò vari legami tra l'algebra booleana e reti logiche elettroniche e gettò le basi per la costruzione dei circuiti digitali. Ottenne vari riconoscimenti tra cui il premio *Alfred Noble American Institute of American Engineers*. Morì il 24 Febbraio 2001 [4].

Nel 1948 Shannon pubblicò “*A Mathematical Theory of Communication*”, articolo suddiviso in due parti pubblicati uno in Luglio e l'altro in Ottobre sul Bell System Technical Journal. Questo lavoro forniva una risposta al problema di come si doveva codificare un'informazione affinché questa potesse essere trasmessa. Egli utilizzò il

calcolo di probabilità sviluppato da Norbert Wiener; diffuse il concetto di entropia dell'informazione intesa come una misura di incertezza in un messaggio e sostanzialmente introdusse la *Teoria dell'Informazione*.

Diversi mesi più tardi Shannon pubblicò un secondo articolo, “*Communication in the Presence of Noise*”, sul “*Proceedings of the Institute of Radio Engineers*”, il giornale istituito dal **IRE** agli inizi del XX secolo sul quale venivano pubblicate le scoperte nel campo delle comunicazioni radio e venivano approvati i nuovi standard. In questo secondo articolo Shannon rielaborò ed estese i concetti espressi nel precedente, affrontandoli da un punto di vista più tecnico che matematico [13].

Il libro “*A Mathematical Theory of Communication*”, di cui è co-autore Warren Weaver è una riscrittura dell'articolo di Shannon del 1948; grazie a Weaver divenne popolare poiché il suo contributo rese accessibile i contenuti del libro anche a chi non era uno specialista del settore.

Shannon fu anche autore nel 1949 di “*Communication Theory of Secrecy Systems*”, col quale diede un notevole contributo alla crittografia.

Shannon mostrò come oltre a insiemi di punti come quelli campionati dal segnale, anche le loro derivate o insiemi di campioni presi in istanti non equamente distribuiti possono essere usati per ricostruire il segnale a banda limitata.

Nella fine degli anni Cinquanta si venne a sapere che gli stessi risultati ottenuti da Shannon furono trovati da un ingegnere russo, Kotel'nikov, nel 1933, ma tali scoperte, note in quel tempo solo nella letteratura dell'Europa dell'Est, si diffusero nell'Ovest solo successivamente alle pubblicazioni di Shannon; più tardi infine, si seppe che un altro ingegnere giapponese, Someya, giunse ai medesimi risultati di Shannon sempre in contemporanea ad esso.

Viste le contorte origini del Teorema del Campionamento così come lo conosciamo oggi, spesso si usa chiamarlo, oltre che Teorema di Shannon, anche Teorema WSK dalle iniziali dei suoi autori Whittaker-Shannon-Kotel'nikov.

A partire dagli anni 60 si iniziarono a comprendere i vantaggi delle trasmissioni digitali rispetto a quelle analogiche: **il punto di partenza delle moderne telecomunicazioni è proprio la Teoria del Campionamento introdotta da Shannon.**

2. TEORIA DEI SEGNALI

2.1 Rappresentazione matematica del suono

Il suono si manifesta come una vibrazione prodotta da un corpo in oscillazione. Tale vibrazione si propaga nell'aria o in un altro mezzo e quando raggiunge l'orecchio dà il via ad un complesso meccanismo che crea la sensazione uditiva. Le molecole del mezzo che collega la fonte del suono all'orecchio, si mettono a vibrare ed eccitano le molecole ad esse vicine: si propaga in questo modo un'onda sonora che ha una certa velocità che dipende dal mezzo di propagazione (circa 344 m/s nell'aria ad una temperatura di 20°C e pressione atmosferica standard) [8].

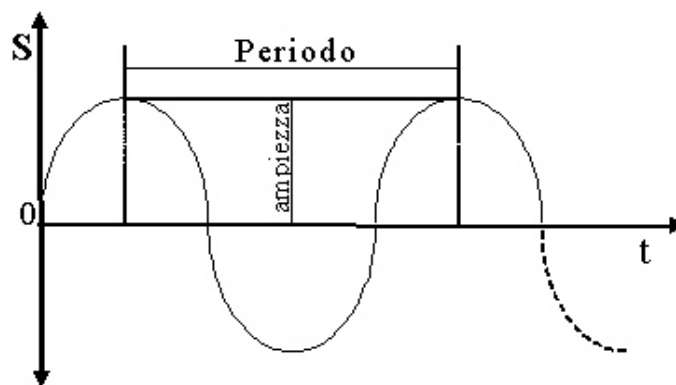


Fig. 2.1 - Rappresentazione schematica di un'onda sonora

Come tutte le onde, quelle sonore sono caratterizzate da una intensità, un timbro, una frequenza ed un'altezza (che nel caso specifico è correlata direttamente alla frequenza).

L'*intensità*, comunemente detta “volume sonoro”, è proporzionale alla *densità* dell'onda sonora. È la qualità sonora associata alla definizione intuitiva di *forza* del suono, ed è determinata dall'ampiezza della vibrazione e dalla distanza del punto di percezione da quello di emissione del suono, ovviamente a causa dell'attrito del mezzo di propagazione, la forza e quindi l'intensità diminuiscono man mano che ci si allontana dalla fonte sonora.

Il *timbro* è la qualità che, a parità di frequenza, distingue un suono da un altro. Dipende dalla forma dell'onda sonora, determinata dalla sovrapposizione delle onde sinusoidali caratterizzate dai suoni fondamentali e dai loro armonici. Dal punto di vista della produzione del suono, il timbro è determinato dalla natura (forma e composizione) della

sorgente del suono e dalla maniera in cui questa viene posta in oscillazione.

L'*altezza* è la qualità che fa distinguere un suono acuto da uno grave. Dipende in massima parte dalla frequenza ma anche dall'intensità. L'orecchio umano percepisce solo i suoni che vanno da **20 a 20.000 vibrazioni al secondo**. Al di sotto di questi valori abbiamo gli *infrasuoni*, al di sopra gli *ultrasuoni*. Strumenti come il sonar, e, nel mondo animale, i delfini ed i pipistrelli, percepiscono gli ultrasuoni, mentre gli elefanti percepiscono gli infrasuoni.

La pratica musicale copre una gamma di suoni, le cui fondamentali vanno dal do grave che ha circa 65 vibrazioni semplici al secondo al do acuto che ha 8276 vibrazioni semplici. La voce umana ha un registro ancora più limitato. Per calcolare l'altezza dei suoni, è stato scelto come punto di riferimento il **La3** (che corrisponde all'ottava centrale del pianoforte) che chiamiamo *diapason* o *corista*.

Il diapason ha la proprietà di emettere suoni puri, cioè suoni di una determinata frequenza; tale frequenza è stata stabilita nell'ambito di diversi congressi, così come spesso capita per l'esigenza di avere uno standard unico in tutto il mondo: nel 1885, al Congresso di Vienna, si stabilì che il La3 corrispondesse a 870 vibrazioni semplici che, a loro volta, corrispondevano a 435 vibrazioni doppie. Ora invece il valore di riferimento è 440 Mhz.

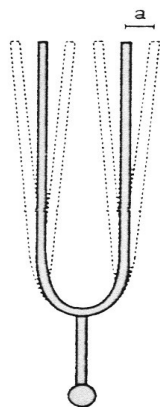


Fig. 2.2 - Schema di un diapason

Quando un diapason si mette a vibrare possiamo descrivere le sue vibrazioni con una

funzione periodica semplice come: $x(t) = a \cos \frac{2\pi}{\tau} t$ oppure $x(t) = a \sin \frac{2\pi}{\tau} t$

dove: $x(t)$ misura lo spostamento, al tempo t , di un corno del diapason dalla sua posizione di riposo; il numero positivo a misura l'ampiezza dell'oscillazione (intesa come spostamento massimo) e viene percepito dall'orecchio come intensità o volume di suono: $1/\tau$ è la frequenza tipica del diapason e viene percepita dall'orecchio umano come altezza del suono [5].

In generale i suoni emessi dagli strumenti musicali (per esempio una corda di un pianoforte) non sono puri, ma risultano dalla sovrapposizione di diverse armoniche.

Questo capita perché le corde con gli estremi fissati possono oscillare in vari modi, cioè con diverse frequenze: oltre che con la frequenza propria della nota relativa alla lunghezza della corda, queste frequenze (teoricamente infinite) sono multipli interi di quella fondamentale, quindi la corda di un do può vibrare anche con frequenza doppia ed emettere una nota che corrisponde al do dell'ottava superiore, o tripla, con una nota corrispondente al sol dell'ottava superiore, e così via.

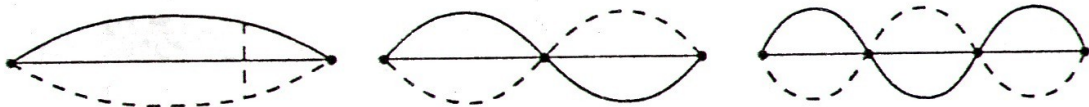


Fig. 2.3 Una corda con gli estremi fissati: esempi di sue possibili oscillazioni

La vibrazione effettiva di un dato punto della corda risulta dalla sovrapposizione delle vibrazioni corrispondenti ai singoli modi: essa è perciò descritta da una somma (anche questa quindi teoricamente infinita) di funzioni periodiche semplici del tipo:

$$a_n \cos \frac{2\pi n}{\tau} t \quad \text{oppure} \quad b_n \sin \frac{2\pi n}{\tau} t \quad (2.1)$$

Tale somma al tendere di n all'infinito è una serie di Fourier.

In pratica, però solo alcuni (pochi) termini della serie contribuiscono effettivamente alla vibrazione totale della corda.

Il nostro orecchio percepisce il risultato di queste vibrazioni, trasmesse nell'aria, come un suono regolare: la nota do, la nota mi, Il numero delle armoniche che si accompagnano all'armonica fondamentale determina il cosiddetto timbro del suono (la stessa nota emessa da una corda di violino o di pianoforte ha un timbro diverso, così

come hanno timbri diversi le voci maschili e femminili).

Quando più corde vengono percosse contemporaneamente, al nostro orecchio arriva la sovrapposizione delle vibrazioni di ciascuna di queste corde. La vibrazione totale è descritta perciò come somma di funzioni periodiche di diverso periodo: se i periodi stanno tra loro in particolari rapporti, si ha una sensazione gradevole (per esempio: accordo do-mi-sol) altrimenti si può percepire un suono sgradevole (do-re).

La scomposizione di un suono nelle proprie componenti sinusoidali fondamentali (*armoniche*) è detta **analisi in frequenza**. La frequenza vengono misurate in Hz, ovvero oscillazioni al secondo. In generale la somma $a_1 \cos x + b_1 \sin x$ è detta *armonica fondamentale*, mentre la: $a_n \cos nx + b_n \sin nx$ è detta *armonica n-esima*.

Se poi si sommano vibrazioni di frequenze qualsiasi (la funzione che rappresenta la vibrazione totale, essendo somma di funzioni con periodi qualsiasi, anche non multipli interi della fondamentale, potrebbe anche non essere più periodica) si ha una sensazione di suono non più regolare: è il *rumore*.

Quindi le armoniche di un suono sono esse stesse suoni con frequenze multiple intere del suono principale. Tanto più un suono è composto da diverse componenti, tanto più esso risulta complesso, ad es, il suono di un flauto dolce è composto dalla fondamentale e da pochissime altre armoniche, mentre il suono degli strumenti ad arco è composto da moltissime frequenze armoniche secondarie.

2.2. La Teoria dei segnali

Il problema della trasmissione del suono viene generalizzato dalla **teoria dei segnali**: è lo studio delle proprietà matematiche e statistiche dei segnali, siano essi suoni o immagini, definiti come funzioni reali o complesse in cui la variabile indipendente è solitamente il *tempo* [8]:

$$f : \mathbf{R} \rightarrow \mathbf{R} \quad \text{o} \quad f : \mathbf{R} \rightarrow \mathbf{C}$$

La legge di associazione non è in genere analitica (non è cioè esprimibile con formule matematiche finite: f è analitica nel punto x se se è sviluppabile secondo la *serie di Taylor* in un intorno di x ; è analitica se lo è in ogni punto del dominio), può essere nota (si parla di **segnali determinati o deterministici**) e quindi può essere possibile predire il valore del segnale in qualunque istante, ma può anche essere del tutto ignota (si parla in questo caso di **segnali stocastici o processi aleatori**, su cui è possibile fare soltanto delle stime statistiche). Un segnale può inoltre avere più dimensioni, ed in tal caso è descritto da una funzione a più variabili. Un segnale digitale è invece una sequenza di “numeri” cioè una successione: $\{f(t_n)\}; n=1, \dots, +\infty$

Un'altra applicazione della teoria dei segnali è quella di previsione di sistemi complessi non lineari (fenomeni meteorologici, andamento della borsa).

L'oggetto della teoria dei segnali è lo studio degli effetti della trasmissione dei segnali attraverso un canale di comunicazione e la degradazione che questo subisce nel passaggio, la misura del loro contenuto informativo e le possibili trasformazioni che questo può subire. Le alterazioni del segnale possono presentarsi come scariche di elettricità elettrostatica nelle trasmissioni radio o telefoniche, o come distorsioni delle figure e delle immagini nelle trasmissioni televisive. Ci sono due categorie principali in cui racchiudiamo le alterazioni del segnale: le **DISTORSIONI**, che sono operazioni sistematiche, avvengono in ogni trasmissione e possono essere corrette applicando l'operazione inversa (feedback negativo e controreazione), e il **DISTURBO**, che si presenta casualmente e non è quindi prevedibile né correggibile.

2.3. Trasmissioni senza distorsione

Consideriamo le trasmissioni senza distorsione. La figura finale è riprodotta identicamente a quella iniziale. Matematicamente possiamo descrivere il processo come una conversione lineare: $g(t) = Lf(t) = Af(t - t_0)$ **(2.2)**

dove: $f(t)$ è il segnale inviato, $g(t)$ è il segnale ricevuto, L è la trasformazione lineare indipendente dal tempo [14].

La f e la g per quanto visto nel paragrafo precedente è possibile scriverle come serie di Fourier.

Se nella relazione precedente scriviamo applichiamo la trasformazione di Fourier su entrambi i lati otteniamo:

$$G(\omega) = H(\omega)F(\omega) \quad \mathbf{(2.3)}$$

dove: F e G sono rispettivamente le trasformate di Fourier delle f e g , e $H(\omega) = Ae^{it_0\omega}$ è detta *funzione di trasferimento del sistema*, in breve, *funzione di sistema*, e la sua trasformazione di Fourier inversa $h(t)$ è detta la *risposta di impulso del sistema*.

3. II TEOREMA DEL CAMPIONAMENTO DI SHANNON

3.1 Premesse al teorema

Sia $f : \mathbb{R} \rightarrow \mathbb{R}$ un segnale. La sua energia è definita da: $E = \int_{-\infty}^{+\infty} |f(t)|^2 dt$.

Supponiamo che l'energia sia finita: $E < +\infty$; questo ha un'importante significato dal punto di vista fisico e matematico, infatti la finitezza di questo integrale implica che la funzione $f \in L^2(\mathbb{R})$, e viceversa $L^2(\mathbb{R})$ coincide con tutto lo spazio delle funzioni ad energia finita [14].

La potenza media di un segnale f su un intervallo (a, b) è definito da:

$\frac{1}{(b-a)} \int_a^b |f(t)|^2 dt$ e la potenza media su tutto \mathbb{R} è definita da:

$$\tilde{E} = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T |f(t)|^2 dt \quad (3.1)$$

Si dice che un segnale ha potenza finita se \tilde{E} è finito.

La trasformazione di Fourier

$$F(\omega) = \hat{f}(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f(t) e^{it\omega} dt \quad (3.2)$$

di un segnale f è detta **spettro di ampiezza del segnale** ed ha notevole importanza nell'analisi applicata all'ingegneria. Conoscere lo spettro di ampiezza del segnale permette agli ingegneri di riguardare il segnale come una somma di sinusoidi di differenti frequenze attraverso la sua trasformazione di Fourier inversa:

$$f(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} F(\omega) e^{-it\omega} d\omega \quad (3.3)$$

Possiamo riguardare $f(t)$ e $F(\omega)$ come due rappresentazioni dello stesso segnale, uno dipendente dal tempo e uno dipendente dalla frequenza. Qualche volta è utile

vedere un segnale dipendente contemporaneamente dal tempo e dalla frequenza, ma questo richiede tecniche matematiche che sono molto più complicate delle trasformazioni di Fourier, per cui ci limiteremo a valutarla in base ad un punto di vista per volta.

Si dice che un segnale è a banda limitata se l'*ampiezza dello spettro* (quindi la sua trasformata di Fourier) si annulla al di fuori di un intervallo tipo $(-W, W)$ e il più piccolo W è detto la larghezza di banda del segnale. Tale segnale può essere scritto come:

$$f(t) = \frac{1}{\sqrt{2\pi}} \int_{-W}^W F(\omega) e^{-it\omega} d\omega \quad (3.4)$$

con $F \in L^2(-W, W)$

Prendiamo in considerazione quei segnali che hanno banda limitata.

Abbiamo molti esempi in natura di segnali a banda limitata: la voce umana ha frequenze non più alte di 8000 Hertz (cicli al secondo); un segnale audio, compresa la musica di un'orchestra convenzionale, ha una banda che va da 20 a 20.000 Hz (totale 20 kHz, trascurando il limite inferiore). Una trasmittente radio FM ha una banda di 50 kHz, anche se lo spettro di frequenza prodotto dalla modulazione ha armoniche anche molto al di là di questo limite, mentre una comunicazione telefonica analogica occupa le frequenze che vanno da 300 a 3400 Hz, quindi ha una banda di 3100 Hz (cioè 3400 – 300).

Analogamente un segnale si dice essere limitato nel tempo se $f(t)$ si annulla al di fuori di un intervallo tipo $(-T, T)$. È noto che se un segnale è a banda limitata, allora la sua rappresentazione rispetto al tempo $f(t)$ prosegue indefinitivamente sebbene diventi sempre più piccolo. È impossibile avere un segnale che sia contemporaneamente limitato nella banda e nel tempo, a meno che non sia identicamente nullo.

Si può però avere un segnale sia “quasi” limitato nel tempo e nella banda. Esistono tanti esempi riconducibili a questo caso. Precisamente, detti T e W i valori limite approssimativi, cioè tali che al di fuori dagli intervalli limitati da questi valori il segnale sia molto piccolo, si ha che questo segnale sarà caratterizzato dalla relazione:

$2TW \geq 1$. Questo fatto è noto come *principio di incertezza* nella teoria del campionamento [6][7]. Shannon affermò che ***un segnale può essere determinato con un alto grado di accuratezza attraverso i suoi valori nei $2TW$ punti campione***

distanziati nel tempo: $\frac{1}{2W}$.

L'intervallo di frequenze è strettamente correlato alla quantità di informazione che può fluire attraverso un canale, come vedremo verrà specificato nel teorema del campionamento.

3.2. Interpretazione geometrica del segnale

Una delle idee più importanti da sottolineare nello studio di Shannon è quella relativa alla sua **interpretazione geometrica** del sistema di comunicazione e della codifica-decodifica dei segnali e dei messaggi. Egli infatti rappresentava il segnale come un punto nello spazio Euclideo N -dimensionale, dove $N \simeq 2WT$, mentre l'energia del segnale continuo veniva messa in relazione con la norma Euclidea nello spazio vettoriale reale \mathbb{R}^N mediante l'uguaglianza di Parseval [13] .

Anche il messaggio sorgente era rappresentato come punto in uno spazio con dimensione $2DT$, dove D era – secondo la terminologia usata da Shannon – la dimensione effettiva dello spazio in cui il messaggio veniva generato.

L'adozione della terminologia geometrica apportò notevoli semplificazioni nei calcoli, inoltre fu un elemento fondamentale per la comprensione dei problemi e per la progettazione di codici, compresi quelli più sofisticati. La terminologia geometrica permise negli anni successivi agli studi di Shannon la diffusione e lo sviluppo della teoria della comunicazione.

3.3. Il Teorema

Il **teorema del campionamento** definisce un criterio per quantizzare un segnale analogico evitando che nella sua riconversione esso presenti distorsioni.

Una volta che un segnale con larghezza di banda finita è noto, la frequenza minima con cui si effettua il campionamento deve essere almeno il doppio della sua massima frequenza.

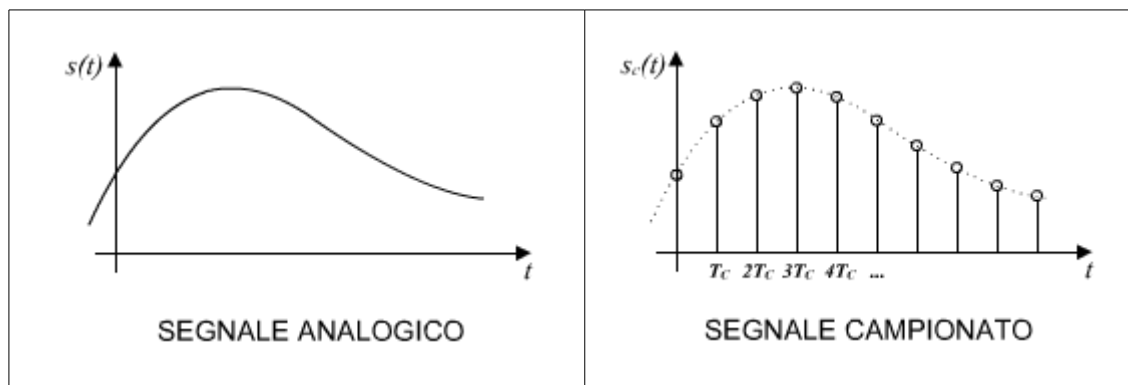


Fig. 3.1 – Rappresentazione di un segnale prima e dopo il campionamento

Il campionamento è il primo passo che si compie nel processo di conversione analogico-digitale di un segnale, e consiste nel prelievo di campioni da un segnale analogico e continuo nel tempo ogni Δt secondi.

Δt è l'intervallo di campionamento, mentre $F_s = \frac{1}{\Delta t}$ è la frequenza di campionamento.

Ciò che si ottiene in questo modo è una successione di valori appartenenti ad un segnale continuo ma rappresentati in un tempo discreto.

Il passo successivo sarà quello di codificare il segnale così quantizzato per renderlo accessibile a qualsiasi elaboratore digitale.

Quindi, il teorema del campionamento determina alcuni criteri per la progettazione e costruzione degli apparati elettronici di campionamento e conversione analogico-digitale: un campionatore che lavora a frequenza F_s , potrà campionare un segnale a

banda limitata da $\frac{F_s}{2}$.

Dato che generalmente un segnale analogico non è limitato in frequenza, dovrà essere prima filtrato in modo tale da eliminare le componenti della sua frequenza maggiori di

$\frac{F_s}{2}$, e poi campionato.

In elettronica la parola “filtro” indica un circuito o un sistema che ha un meccanismo di selezione delle frequenze [14]; generalmente i filtri sono caratterizzati come *filtro passa-basso*, *filtro passa-alto*, *passa-banda* e *blocca-banda*. Le funzioni di trasferimento del sistema corrispondenti sono:

$$\text{FILTRO PASSA-BASSO} \quad H(\omega) = \begin{cases} A e^{it_0\omega} & \text{per } |\omega| \leq \omega_1 \\ 0 & \text{altrove} \end{cases} \quad (3.5)$$

$$\text{FILTRO PASSA-ALTO} \quad H(\omega) = \begin{cases} A e^{it_0\omega} & \text{per } |\omega| \geq \omega_2 \\ 0 & \text{altrove} \end{cases} \quad (3.6)$$

$$\text{FILTRO PASSA-BANDA} \quad H(\omega) = \begin{cases} A e^{it_0\omega} & \text{per } \omega_2 \leq |\omega| \leq \omega_1 \\ 0 & \text{altrove} \end{cases} \quad (3.7)$$

$$\text{FILTRO BLOCCA-BANDA} \quad H(\omega) = \begin{cases} 0 & \text{per } \omega_2 \leq |\omega| \leq \omega_1 \\ A e^{it_0\omega} & \text{altrove} \end{cases} \quad (3.8)$$

dove ω_1 e ω_2 sono le *frequenze di taglio*.

Nel processare un segnale $f(t)$ operiamo su di esso in modo tale da cambiarne la figura, configurazione e proprietà, oppure per estrarne qualche utile informazione.

In tutti i casi l'operazione deve essere invertibile (deve esserlo quindi la funzione).

A volte solo certi dati estratti dal segnale vengono trasmessi ed usati dal ricevente per ricostruire il segnale. Qui entra in gioco la teoria dei segnali.

Campionando un segnale da un insieme di punti discreto, stiamo applicando una trasformazione lineare per convertire $f(t)$ in un segnale digitale.

Enunciato:

Un segnale $f(t)$ a banda limitata da f_M può essere univocamente ricostruito dai suoi campioni $f(n\Delta t)$ ($n \in \mathbf{N}$) presi a frequenza $F_s = \frac{1}{\Delta t}$, se $F_s \geq 2f_M$.

Dimostrazione:

Sia $F(\omega)$ la trasformata di Fourier di $f(t)$. Poiché $f(t)$ ha come limite di banda f_M , risulta $F(\omega) = 0$ per $|\omega| > 2\pi f_M$. Sia $W = \frac{F_s}{2}$, allora per ipotesi:

$$W \geq f_M \rightarrow F(\omega) = 0; \forall |\omega| > 2\pi W$$

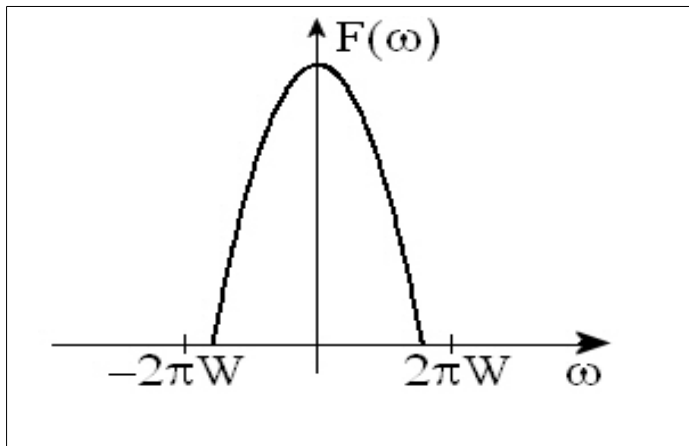


Fig. 3.2: La funzione $F(\omega)$

Sia $Q(\omega)$ la funzione periodica di periodo $4\pi W$ che coincide con $F(\omega)$ nell'intervallo $[-2\pi W, 2\pi W]$.

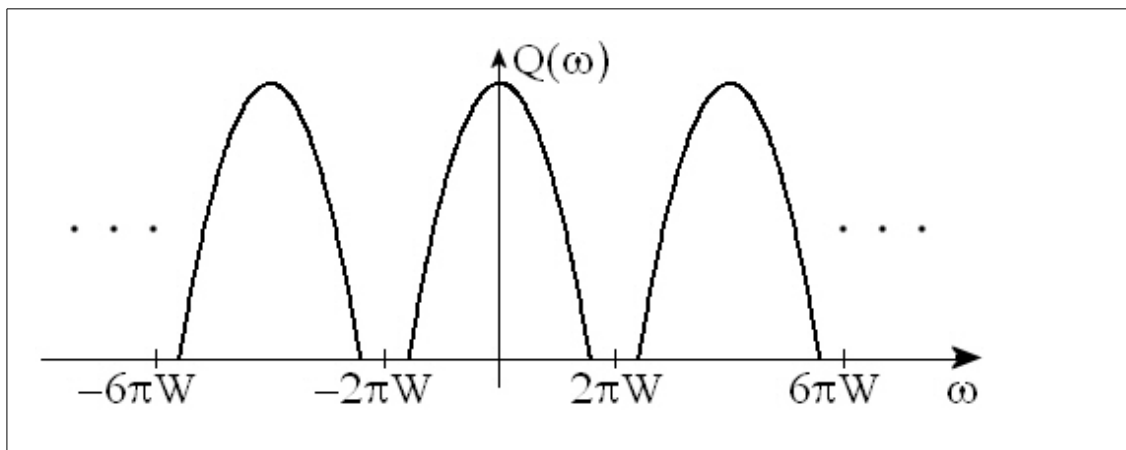


Fig. 3.3 Il segnale esteso $Q(\omega)$: è la trasformata di Fourier del segnale campionato $f(n\Delta t)$

Come si nota è periodica di periodo $4\pi W$ e coincide con $F(\omega)$ in $[-2\pi W, 2\pi W]$.

Il suo sviluppo in serie di Fourier sarà:

$$Q(\omega) = \sum_{n=-\infty}^{+\infty} c_n e^{i \frac{n}{2W} \omega} \text{ dove: } c_n = \frac{1}{4\pi W} \int_{-2\pi W}^{2\pi W} Q(\omega) e^{-i \frac{n}{2W} \omega} d\omega \quad (3.9)$$

Poiché $Q(\omega) = F(\omega)$ in $[-2\pi W, 2\pi W]$ possiamo porre:

$$c_n = \frac{1}{4\pi W} \int_{-2\pi W}^{2\pi W} F(\omega) e^{i \frac{n}{2W} \omega} d\omega \quad (3.10)$$

Si osservi ora che $f(t)$ è l'antitrasformata di Fourier di $F(\omega)$, cioè:

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega) e^{i\omega t} d\omega = \frac{1}{2\pi} \int_{-2\pi W}^{2\pi W} F(\omega) e^{i\omega t} d\omega \quad (3.11)$$

Dalle (3.10) e (3.11) si ottiene:

$$c_n = \frac{1}{2W} f\left(-\frac{n}{2W}\right) = \Delta t f(-n \Delta t) \quad \text{Ricordiamo che: } W = \frac{F_s}{2} = \frac{1}{2\Delta t}$$

Si definisce la funzione caratteristica:

$$\chi_W(\omega) = \begin{cases} 1 & \text{se } |\omega| \leq 2\pi W \\ 0 & \text{se } |\omega| > 2\pi W \end{cases} \quad (3.12)$$

allora:

$$\begin{aligned} F(\omega) &= Q(\omega) * \chi_W(\omega) = \sum_{n=-\infty}^{\infty} e^{i \frac{n}{2W} \omega} \chi_W(\omega) = \\ &= \Delta t \sum_{n=-\infty}^{\infty} f(-n \Delta t) e^{i \Delta t \omega} \chi_W(\omega) \end{aligned} \quad (3.13)$$

$$\begin{aligned} f(t) &= F^{-1}[F(\omega)] = \\ &= \Delta t \sum_{k=-\infty}^{+\infty} f(k \Delta t) \frac{\sin\left(\frac{\pi}{\Delta t} t - k \pi\right)}{\pi(t - \Delta t)} \end{aligned} \quad (3.14)$$

La (3.13) e la (3.14) mostrano che $F(\omega)$, e quindi anche la sua antitrasformata $f(t)$,

possono essere ricostruite sulla base della conoscenza di $f(n\Delta t)$, come volevasi dimostrare [10].

3.4. Problematiche

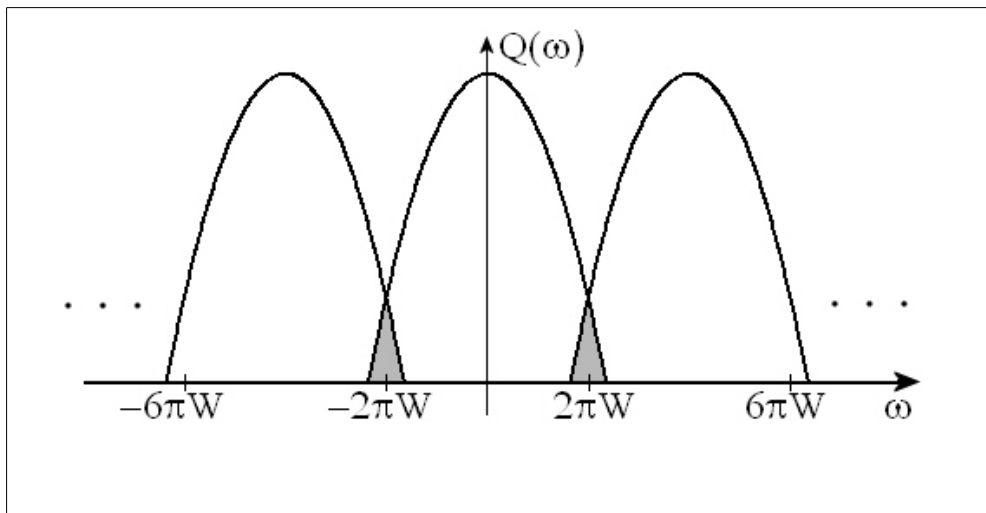


Fig. 3.4: Se $F(\omega)$ ha componenti in frequenza maggiori di $\frac{F_s}{2}$ allora le ripetizioni periodiche di

$Q(\omega)$ si sovrappongono ed il segnale ricostruito (2.18) risulta distorto.

Lo spettro di un segnale campionato è uguale allo spettro del segnale originale ripetuto periodicamente in frequenza (vedi fig. 3.2 e 3.3) . Il periodo di questa ripetizione è uguale alla metà della frequenza di campionamento F_s , quindi se la frequenza

massima del segnale originale supera $\frac{F_s}{2}$ (Fig. 3.4) le ripetizioni nello spettro del segnale campionato si sovrappongono. A causa di questa sovrapposizione risulta impossibile ricostruire esattamente il segnale originale da quello campionato; tale ricostruzione risulterà distorta [10]. Tale distorsione è chiamata effetto **aliasing** (vedi paragrafo successivo).

3.5 Aliasing

Durante la conversione analogico-digitale il segnale viene sottoposto a due distinte discretizzazioni: la discretizzazione dei tempi, detta campionamento, e la discretizzazione delle ampiezze, detta quantizzazione [10]. Nella conversione digitale-analogica il segnale digitale viene trasformato nuovamente in un segnale analogico operando una interpolazione. L'**aliasing** (dal latino alias, *altrove*) è il fenomeno per il

quale due segnali analogici diversi possono diventare indistinguibili una volta campionati; questo costituisce un serio problema che si riflette direttamente sull'uscita del sistema in esame, alterandone la veridicità. L'aliasing può verificarsi sia nel tempo (aliasing temporale) che nello spazio (aliasing spaziale).

Esempio:

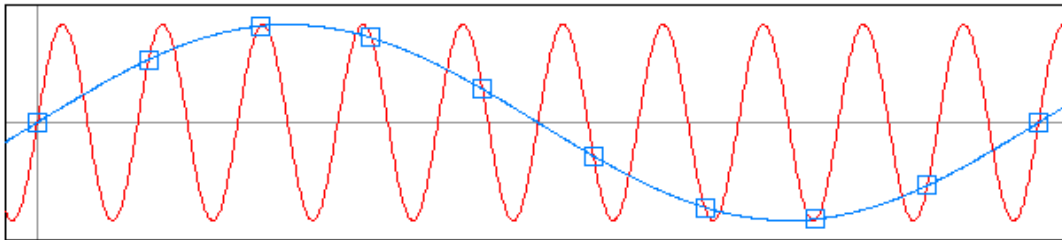


Fig. 3.5 - Esempio di aliasing dovuto ad un campionamento a frequenza insufficiente

Campionando un segnale sinusoidale con una certa frequenza l'insieme dei punti che si acquisisce non permette di identificare univocamente una sola sinusoide. Dato che, tramite l'analisi di Fourier, ogni segnale continuo può essere visto come sovrapposizione di seni e coseni, risulta importante in teoria dei segnali riuscire a limitare questa ambiguità.

Per far questo ogni apparato di conversione analogico-digitale ha un filtro anti-alias che agisce prima del campionamento: il ruolo di tale filtro è quello di eliminare dal segnale in ingresso le componenti di frequenza maggiori della metà della frequenza di

campionamento dell'apparato $\frac{F_s}{2}$. Lo schema è quello di un filtro passa-basso (3.5).

Ci sono soluzioni alternative a questo problema, cioè se non si vuole perdere la gamma di frequenze che altrimenti dovremmo tagliare, è possibile utilizzare uno strumento di conversione Analogico/Digitale più veloce, o utilizzare tecniche di sotto-campionamento.

La seconda opzione è realizzabile quando le frequenze di interesse sono racchiuse in un

range $\Delta F = F_{max} - F_{min} < \frac{F_s}{2}$, anche se sia F_{max} che F_{min} superano $\frac{F_s}{2}$

(cioè stiamo considerando la differenza e non più i valori delle frequenze). In questo caso tuttavia il limite imposto dal teorema del campionamento non è più sufficiente a garantire un campionamento corretto.

4. APPLICAZIONE: IL SISTEMA VoIP

4.1. Vista d'insieme di una connessione VoIP

Una comunicazione VoIP avviene attraverso i seguenti passaggi [1]:

1. La prima cosa necessaria è la conversione del segnale voce in digitale. Per farlo viene utilizzato un ADC (convertitore analogico-digitale). In questa fase si attua il campionamento e la quantizzazione.
2. Una volta digitalizzato il segnale, si deve comprimere l'insieme di bit per economizzare il più possibile sulle risorse impiegate; questa compressione deve seguire però determinati protocolli per far non perdere la qualità dei dati.
3. I pacchetti voce ottenuti vanno inseriti in pacchetti di dati standard: tutto ciò va fatto in tempo reale, andrà quindi utilizzato un protocollo real-time (solitamente RTP su UDP/IP)
4. Affinché i dati arrivino a destinazione occorrerà un protocollo di segnalazione per chiamare gli utenti: tra i vari protocolli che svolgono questo compito ci sono H.323 e SIP
5. Quando i dati arrivano a destinazione, saranno sotto forma di “pacchetti di dati”: è necessario quindi disassemblarli, estrarli, convertirli in segnali vocali analogici (mediante un convertitore DAC, cioè digitale-analogico). I segnali vanno quindi mandati alla scheda audio del PC o alla cornetta del telefono, a seconda che si stia utilizzando per la comunicazione un software ed un PC, o un apparato ATA e un tradizionale telefono.
6. Tutto questo deve essere effettuato in tempo reale. Ovviamente una conversazione telefonica non può avere eccessivi ritardi. La valutazione della comunicazione è detta QoS, *Qualità del Servizio*.

Architettura base

Si può schematizzare quanto detto nel modo seguente:

VOCE	ADC	Algoritmo di compressione	Assemblaggio	RTP in UDP/IP
------	-----	---------------------------	--------------	---------------



RTP in UDP/IP	Disassemblaggio	Algoritmo di decompressione	DAC	VOCE
---------------	-----------------	-----------------------------	-----	------

4.2. Conversione analogica digitale

Una prima rappresentazione digitale del segnale analogico è la **Pulse-code modulation (PCM)**: il segnale campionato viene trasformato in formato digitale e quindi diventa una sequenza di bit (di solito si usa il sistema binario) [10].

La quantizzazione viene fatta da un hardware, chiamato **ADC** (*convertitore analogico digitale*).

La quantizzazione può essere **lineare**, quando tutti gli intervalli sono uguali, oppure **non lineare** quando gli intervalli sono diversi in relazione all'ampiezza. In questi casi ogni quantizzazione viene conservata, indipendentemente dai valori precedenti o successivi, pertanto la dimensione dello *stream* (flusso) digitale è dato semplicemente dal prodotto della frequenza alla quale avviene il campionamento per la dimensione in bit della parola.

Esistono altre modalità di quantizzazione, principalmente utilizzate in telefonia: **adattativa**, quando gli intervalli di campionamento vengono variati in relazione alla dinamica del segnale (con lente variazioni i campioni vengono eseguiti a frequenze più basse) o **differenziale**, attraverso la codifica della sola differenza tra un campione ed il precedente. Questi metodi consentono una discreta riduzione dello *stream* digitale ("compressione") e sono particolarmente indicate in presenza di segnali la cui dinamica sia ragionevolmente prevedibile, ad esempio la voce umana.

Ogni quantizzazione introduce un degrado, chiamato *rumore di quantizzazione* che è in stretta relazione con la dimensione in bit della parola digitale e con la frequenza di campionamento: il rumore può essere ridotto aumentando i bit oppure aumentando la frequenza di campionamento.

Il degrado provocato dalla quantizzazione è irreversibile e può essere ridotto rispettando il Teorema Shannon.

Oggi pressoché ogni scheda audio permette di convertire in 16 bit = (2 byte) una banda di 22050 Hz (per il campionamento della quale abbiamo bisogno di 44100 Hz per il teorema del Shannon) ottenendo una velocità di 2 byte * 44100 (campioni al secondo) = 88200 Byte/s, 176.4 kByte/s per i flussi dati stereo.

Le attuali linee telefoniche esistenti, che sono i canali di comunicazione usati per le trasmissioni, mediamente non hanno una banda così elevata (176.4 kByte/s): è

necessario quindi effettuare particolari codifiche.

4.2.1 Rapporto tra banda digitale e banda analogica

La banda di un segnale analogico si intende la differenza tra la sua massima e minima frequenza.

Per banda di un canale di trasmissione analogico (cavo coassiale, doppio telefonico, ponte radio, ecc.) si intende la sua massima velocità di trasmissione, cioè la massima quantità di informazione che esso può trasmettere nell'unità di tempo. La banda passante è determinata dall'intervallo di frequenze che tale canale è in grado di trasmettere efficacemente (differenza tra le due frequenze, massima e minima); la quantità di informazione che il canale è capace di portare, invece, dipende non solo dalla banda passante, ma anche dalla gamma dinamica che il canale è in grado di garantire.

Nelle comunicazioni digitali la banda si misura direttamente in bit al secondo (più comunemente si usano i suoi multipli: kbit/s, Mbit/s ecc.).

L'equivalenza fra i due tipi di banda è data dall'operazione di campionamento necessaria per trasformare un segnale analogico in uno digitale: per farlo si deve stabilire la frequenza di campionamento e il numero di bit di ciascun campione, in modo tale da definire una banda digitale "equivalente" a quella analogica. Il teorema del campionamento dirà che, per poterlo rendere correttamente, un segnale analogico deve essere campionato almeno ad una frequenza doppia della sua *banda passante*; è necessario quindi calcolare il numero di bit per campione. Per farlo si controlla la dinamica del segnale e la si confronta con la dinamica di campionamento; dato che questa è di 6 dB per ogni bit per campione utilizzato, si adotta un numero di bit per campione che garantisca una dinamica almeno pari a quella originale.

La dinamica di un segnale è data dal **rapporto segnale/rumore**, spesso abbreviato con la sigla inglese **SNR** (*Signal to Noise Ratio*) o **S/N** (presente anche nell'uso italiano), è una grandezza numerica che mette in relazione la potenza del segnale utile rispetto a quella del rumore in un qualsiasi sistema di acquisizione, elaborazione o trasmissione dell'informazione [10].

Il rapporto segnale/rumore è un numero puro dato dal rapporto fra due grandezze

omogenee, che esprime quanto il segnale sia più potente del rumore nel sistema considerato. È espresso dalla relazione:

$$SNR = \frac{P_{segnale}}{P_{rumore}} \quad \text{con} \quad 0 \leq SNR \leq +\infty \quad (4.1)$$

dove $P_{segnale}$ è la potenza del segnale utile e P_{rumore} la potenza totale del rumore presente nel sistema, grandezze queste solitamente espresse in Watt o dBm.

È un parametro di merito molto importante per il dimensionamento dei sistemi di telecomunicazioni. Più è basso l'SNR, infatti, e più sarà difficoltosa la ricezione del segnale.

Nel caso di trasmissioni analogiche una diminuzione di SNR determina un deperimento graduale della qualità del segnale ricevuto (si pensi ad esempio ad una radio FM che riceve male o ad un televisore in cui compaiono audio o video disturbati: tipici casi di SNR basso); per le trasmissioni analogiche tuttavia è l'utente finale a stabilire una soglia di fruibilità del sistema.

Nel caso di trasmissioni numeriche, invece, esiste una soglia minima di SNR sotto la quale il sistema non è in grado di funzionare (si pensi alla televisione digitale satellitare: o si vede bene o non si vede del tutto), che tuttavia grazie alle moderne tecniche di modulazione e di protezione dei dati tramite codici è piuttosto bassa, ed è comunque al di sotto di quella che servirebbe per ottenere prestazioni analoghe su sistemi analogici.

La soglia minima di SNR è determinata dalla tecnologia dell'apparato ricevente; in fase di progetto di un sistema di telecomunicazioni il primo obiettivo è quindi quello di far pervenire al ricevitore un SNR sufficientemente elevato.

Così, per esempio, una conversazione telefonica che ha una banda di 3100 Hz e un rapporto segnale/rumore di circa 45 dB può venire digitalizzata con 8 bit per campione ($8 \text{ bit} * 6 \text{ dB} = 48 \text{ dB}$ di dinamica) e una frequenza di 6200 Hz ottenendo una banda digitale di 48,5 kbit/s (tutte le compagnie telefoniche del mondo allocano 4 kHz per ogni canale telefonico analogico: perciò quando si trattò di digitalizzare la rete telefonica analogica, la frequenza di campionamento unanimemente scelta fu di 8 kHz, originando quindi una "banda telefonica" mondiale standard di $8 * 8000 = 64 \text{ kbit/s}$).

Questo risultato sembrerebbe essere in contraddizione con l'esistenza dei modem analogici V90 a 56K: come è possibile che questi possano trasmettere 56 kbit/s attraverso un canale che ne porta solo 48,5? Per capirlo, bisogna prima notare che il collegamento di questi modem è asimmetrico: ricevono dalla centrale telefonica a 56K ma trasmettono a 33,6 kbit/s.

Il "trucco" sta nello sfruttare il convertitore analogico/digitale della centrale telefonica, pilotandolo direttamente dal circuito modulatore e abbassando il rumore di conversione: questo permette di aumentare il rapporto segnale-rumore (ma solo in trasmissione) e quindi la dinamica disponibile.

4.2.2 Voice Over IP su connessione analogica e ISDN

Le tecnologie di compressione della voce dedicano una banda che varia dai 4 kbit/sec. agli 82 kbit/s nei formati di compressione meno efficienti. Ad esempio il GSM codifica la voce con algoritmo RPE-LTP (Regular Pulse Excitation - Long Term Prediction) con LPC (Linear Predictive Coding) loop e campionamento a 13 kbit/s.

Come visto nel capitolo precedente, la voce umana richiede una larghezza di banda di almeno 40 kHz. Per il teorema di Shannon per poterla ricostruire correttamente da una serie di valori digitali discreti dobbiamo prenderli con una frequenza di campionamento pari o superiore al doppio della massima frequenza contenuta nel segnale in input. Ad un kilohertz di segnale analogico corrisponde un kbit/s di segnale digitale: per avere un buon campionamento occorre però il doppio di banda: **40kbit/s**.

Tuttavia, con i formati di compressione diventa possibile ridurre drasticamente la banda richiesta, la compressione interviene dopo la digitalizzazione della voce e prima del suo invio.

Un segnale analogico ha una qualità in generale maggiore di uno digitale e proprio questa qualità è l'obiettivo da raggiungere quando si effettua la conversione.

Linea PSTN

Le connessioni su linea telefonica analogica raggiungono velocità di 56kbit/s solamente per il download di pagine Internet. Per il download di altri file o il downstream della

voce la velocità massima è 33.6 kbit/s. Anche l'upstream del 56kbit/s è 33.6 kbit/s.

Una chiamata via Internet richiede una banda maggiore della frequenza di campionamento della voce sia in upstream che in downstream, essendo la comunicazione telefonica bidirezionale (full duplex).

La linea telefonica analogica utilizza le frequenze comprese tra 300 Hz e 3.4 kHz, mentre le frequenze superiori vengono filtrate con l'impiego di un condensatore o da un filtro passa-basso in centrale prima che il segnale venga campionato in digitale. Infatti i segnali analogici vengono trasportati sulle reti primarie come flusso di dati digitali dopo un campionamento che viene effettuato utilizzando una frequenza di campionamento a 8 kHz.

Se ciò non avvenisse si creerebbe una situazione in cui l'ordine dei pacchetti di dati (che sono la conversione della voce) non verrebbe rispettato, e allora potrebbe capitare che un interlocutore ascolti in sequenze di suoni e parole senza senso, poiché non nella giusta sequenza di invio.

Mentre per dati quali bit di pagine Internet o di file scaricati, il modem è in grado di fare controlli, interpolare (ossia ricostruire) pacchetti e sequenze di bit danneggiati e nel caso questo non fosse possibile, può richiederne il rinvio, non esistono controlli che rimedino ai difetti di trasmissione della voce; non è possibile in sostanza migliorare la qualità della comunicazione.

Con un adeguato protocollo di comunicazione si può campionare la voce in modo tale da avere un pacchetto di dati voce compresso intorno ai 12-13 kbit/s.: è possibile con un pacchetto-voce trasmesso per secondo avere pacchetti di *circa* 1,5 Kbyte (che sono 12000 bit appunto; 1 byte=8 bit). In questo modo la trasmissione resta al di sotto della soglia critica che crea problemi con la connessione Internet.

Linea ISDN

Con linea ISDN si dispone di una connessione simmetrica su due linee: la velocità in upstream è la stessa che si ha in downstream. Con una larghezza di banda di 64 kbit/s, ovvero impegnando anche una sola linea, è possibile stabilire una chiamata Voice Over IP anche con formati di compressione della voce meno efficienti che campionano a 50-60 kbit/s (con conseguente valore di indice di qualità della riproduzione). Con l'utilizzo

dei due canali ISDN è possibile addirittura gestire due chiamate Voice Over IP contemporaneamente.

In America, gli operatori telefonici utilizzano il Voice Over Ip fino all'abitazione dell'utente; dunque, anche l'ultimo miglio è digitalizzato per consentire la comunicazione. Il vantaggio di un'estensione delle connessioni ISDN per la digitalizzazione dell'ultimo miglio è meno utilizzato dalle compagnie europee.

4.2.3 Connessioni Internet veloci

Per garantire servizi come il VoIP, ma anche l'e-learning, il commercio elettronico, il video on demand (in generale i servizi BBTTH: BroadBand To The Home), si è giunti alla diffusione delle linee xDSL: HDSL, SHDSL, ADSL, ... tali da garantire trasmissioni in larga banda attraverso il doppino in rame usato tradizionalmente per le linee telefoniche. Tra i vari tipi di xDSL i più diffusi sono [3]:

HDSL

High data rate DSL garantisce un collegamento simmetrico ad alta velocità. Un collegamento si dice **simmetrico** quando la velocità di emissione e la velocità di ricezione dei dati sono uguali. Le velocità raggiunte in questo caso sono entrambe dell'ordine dei 2Mbps. Il limite maggiore è che richiede due o tre linee per funzionare. Viene usata per reti private o per collegamenti dedicati, intranet aziendali, per elevati volumi di messaggistica o videoconferenza.

SHDSL

Single line High data rate DSL è una semplificazione di HDSL e usa una sola linea telefonica fornendo un collegamento simmetrico. In genere viene in applicazioni dove è richiesto un flusso simmetrico (es. LAN o server remoti).

ADSL

Asymmetric DSL è invece ottimale per tutte quelle applicazioni che richiedono una capacità di download molto più grande di quella di upload. Applicazioni come la navigazione, il video e l'audio on-demand, ecc. sono tra le più usate nelle utenze domestiche. È proprio la differenza di velocità che definisce l'asimmetria di questa soluzione. Inoltre ADSL usa una parte della banda della linea telefonica che non si

sovrappone con quella vocale. Il che consente la coesistenza del servizio telefonico analogico e del servizio dati. Le velocità raggiunte sono 8Mbps in ricezione e 640Kbps in emissione. La distanza massima che si può coprire è di circa 3-4Km.

ADSL 2

La tecnologia ADSL2 incrementa la velocità di ricezione dati fino a 12Mbps contro gli 8Mbps dell'ADSL, e quella in emissione fino ad 1Mbps contro i 640Kbps della precedente. Inoltre ne estende il raggio d'azione.

ADSL 2+

Con **ADSL2+**, infine, si raddoppia la frequenza massima utilizzata per la trasmissione dei dati, da 1.1 a 2.2 MHz, questo consente di poter modulare più informazioni nella stessa unità di tempo, quindi di portare un flusso dati di 25Mbps su cavo telefonico. In questo caso bisogna fare ancor più attenzione alle distanze. Infatti una buona resa di questa tecnologia è assicurata per soli 1,5-2 Km dalla centrale.

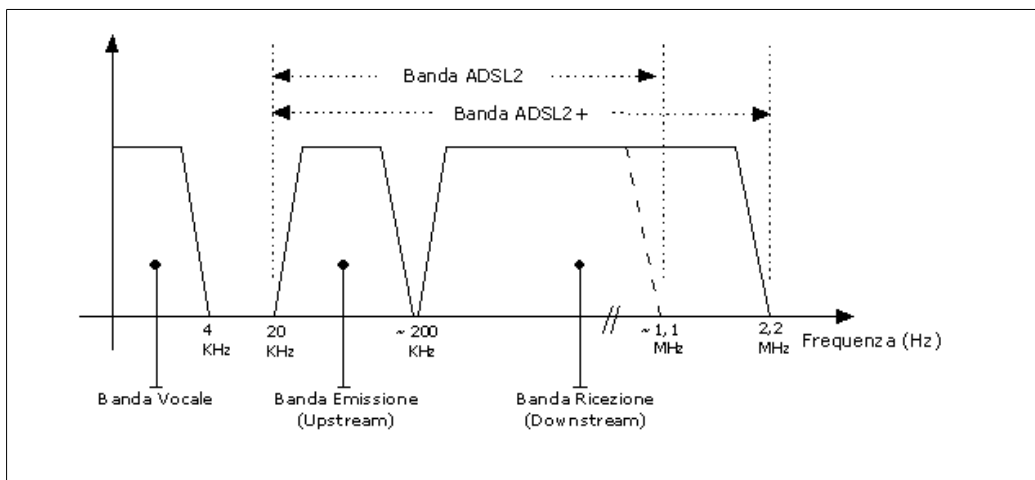


Fig. 4.1 – Schema delle frequenze utilizzate dalle linee PSTN + ADSL

Il doppino telefonico: il mezzo di trasmissione

Nelle reti telefoniche pubbliche PSTN chiamiamo **local loop** il collegamento tra le utenze (abitazioni e sedi aziendali) e la centrale telefonica più vicina. La scelta del mezzo di trasmissione cadde sul doppino intrecciato (*twisted pair*): un cavo composto da due fili di rame con guaina di plastica intrecciati tra loro. Tale scelta fu condizionata

dal costo contenuto del materiale e dal fatto che tale mezzo fosse adeguato alla trasmissione vocale, poiché questa non necessitava di una banda particolarmente elevata.

La diffusione delle trasmissioni digitali a banda larga trovano oggi il **collo di bottiglia** proprio su questo tratto terminale. Infatti le comunicazioni tra centrali sono spesso realizzate in fibra ottica e su questo mezzo i dati possono transitare ad altissima velocità. La maggior parte delle case e degli uffici in Italia non è raggiunta dalla fibra ottica ma ha il doppino come unico mezzo possibile per la connessione.

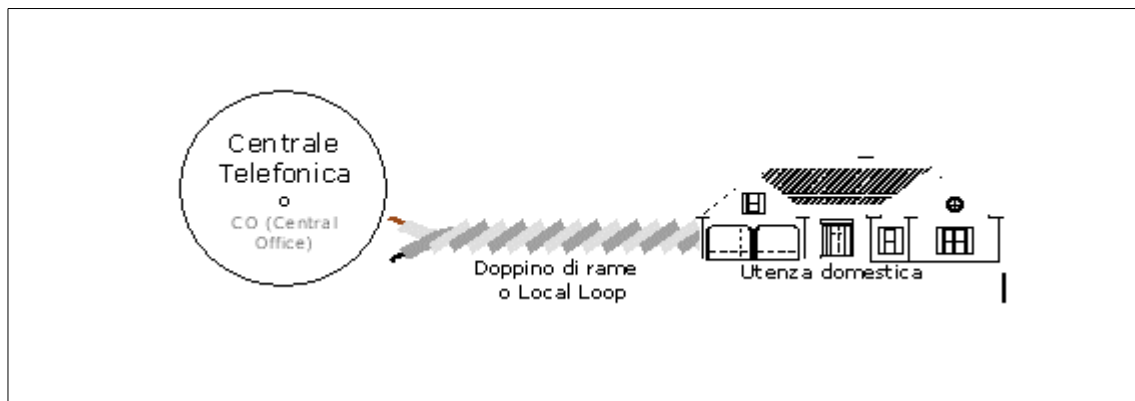


Fig. 4.2 – Rappresentazione del Local Loop

La caratteristica centrale da considerare nell'uso di un mezzo di trasmissione è l'**ampiezza di banda** ovvero l'intervallo di frequenze che esso consente di trasmettere. Il segnale vocale trasmesso sul local loop copre l'intervallo da 0 a 4KHz, che come visto in base al teorema di Shannon è sufficiente per il segnale vocale campionato. La banda disponibile sul doppino è più ampia ma come fattore limitante interviene la lunghezza del cavo: più un terminale è distante dalla propria centrale, più il cavo è lungo e il segnale subisce un deterioramento.

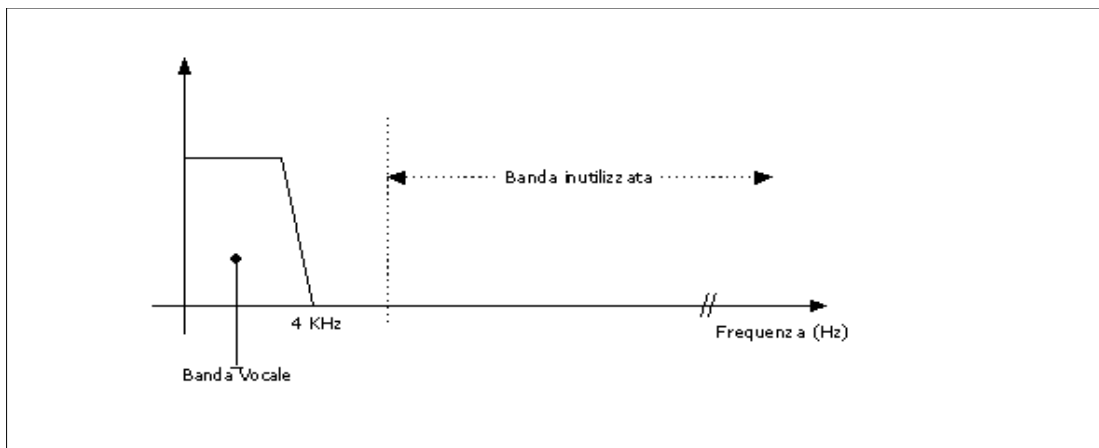


Fig. 4.3 – Utilizzo della banda nel trasporto

Il **segnale si deteriora** per vari motivi: per esempio per effetto della diafonia (interferenza tra più linee poste una vicino all'altra) oppure per interferenza dei segnali radio AM, in ogni caso più si accorciano le distanze, migliore è la resa. Ci sono tecniche per assicurare il segnale vocale anche a lunghe distanze dell'ordine di poche decine di Km, ma per poter usare al meglio l'ADSL, che sfrutta più a fondo la banda del doppino, non ci si può allontanare per più di 3,5Km circa, oltre i quali non arriva una banda sufficiente alla trasmissione per cui non è possibile far transitare il VoIP.

Una conseguenza di quanto detto è che non tutti i local loop sono uguali. È stato sviluppato allora un sistema di modulazione che si adatta alle caratteristiche della linea, è detto **DMT** (*Discrete Multitone Modulation*) ed è una modulazione pensata per adeguare il carico di informazione alle condizioni della linea grazie ad un algoritmo chiamato "bit loading" che agisce nel seguente modo: esegue un rapido test sul canale all'inizio della connessione per misurarne la capacità, in base al responso verrà poi dimensionata la velocità di trasferimento.

La suddivisione delle frequenze

Come dicevamo, ADSL è nata per sfruttare a fondo la banda del local loop. Infatti sullo stesso mezzo fisico (parliamo sempre del doppino), che in telefonia trasportava solo il segnale vocale, si trovano a viaggiare **tre flussi di segnale** separati:

- canale Voce da 0 a 4KHz
- canale ADSL di emissione verso la rete da 25 a 200KHz
- canale ricezione dalla rete da 200KHz a 1,1GHz circa.

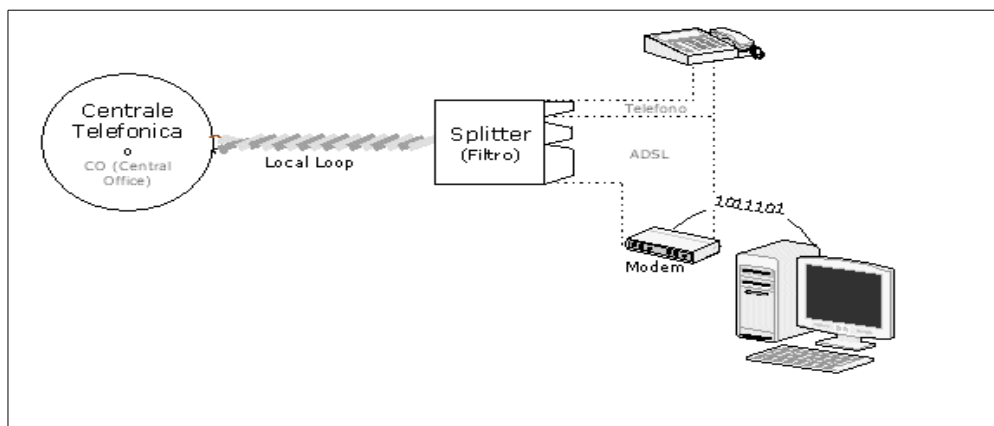


Fig. 4.4 – Suddivisione dei segnali trasportati

La banda vocale deve rimanere intatta altrimenti non sarebbe possibile utilizzare il contemporaneamente sullo stesso cavo telefonia PSTN ed Internet. La suddivisione viene effettuata da circuiti a filtri passivi detti **splitter**. Un filtro è un circuito che taglia alcune frequenze nello spettro e ne fa “passare” altre (rif pag. 21). Si dice passivo quando per funzionare non ha bisogno di alimentazione. In questo caso è utile che siano passivi perché in mancanza di corrente viene comunque garantito il servizio di fonia.

Da uno splitter o da un filtro **si diramano 2 linee**. Una porta il segnale telefonico analogico (in banda base a 4 KHz) e l'altra porta il segnale digitale modulato. Alla prima viene collegato il telefono, alla seconda invece un modem capace di trasformare il segnale digitale modulato nel segnale digitale per il computer, o un apparato *ATA* che ritrasforma in analogico il segnale VoIP. Sullo stesso doppino quindi possiamo far transitare un segnale voce analogico tradizionale, la linea ADSL e su questa ancora un altro segnale voce (uno o più se la banda totale lo consente) stavolta digitalizzato e riconvertito in analogico al suo arrivo a destinazione.

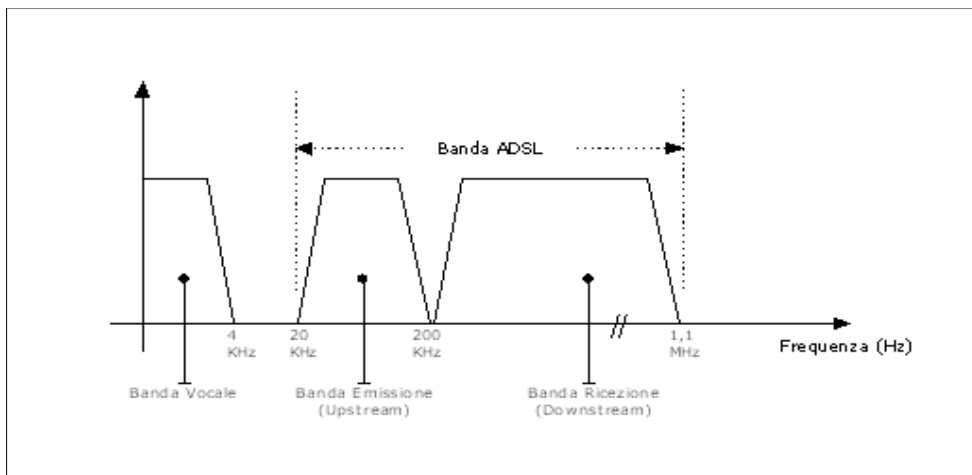


Fig. 4.5 – Suddivisione delle frequenze su linee ADSL

Gli **ATA** (*Adattatore Telefono Analogico*) sono apparecchi hardware che permettono di interfacciare qualsiasi telefono analogico alla rete VoIP convertendo il segnale analogico in formato compatibile. Di solito gli adattatori dispongono di un ingresso di rete per il collegamento alla LAN interna e di un collegamento RJ11 in uscita verso il telefono tradizionale.

E' possibile interfacciare ad un ATA qualsiasi apparato telefonico: cordless, telefono o FAX, e convertire in pochi istanti tutti i telefoni tradizionali in telefoni IP, quindi ovviare alla necessità di effettuare/ricevere chiamate tramite il PC.

Telefoni Ip

Sono terminali esteriormente molto simili ai tradizionali apparecchi telefonici ma che collegati ad una rete LAN permettono di effettuare e ricevere telefonate senza la presenza di un computer acceso. Nel loro hardware l'ATA è incorporato.

Il mercato offre ormai numerosi modelli, con prezzi variabili a seconda delle funzioni, che generalmente sono collegabile direttamente a una presa di rete ethernet RJ 45.

Dispositivi all-in-one

Questa tipologia di apparecchi integra le funzioni di modem/router Adsl, switch ethernet (in alcuni casi anche l'access point wireless), firewall e adattatore per telefoni analogici (ATA), ed infine la possibilità di collegamento alla linea telefonica tradizionale. Grazie alla possibilità di configurare le regole di instradamento in fase di configurazione, i

dispositivi telefonici possono così essere indirizzati sulla linea PSTN (o ISDN laddove presente) o su quella VoIP in base al numero chiamato. Con un unico apparecchio si può cioè effettuare la connessione in Internet (cablata e wireless), chiamare e ricevere chiamate dalla linea PSTN (o ISDN) e tramite l'ADSL dalla linea VoIP.

In questo modo sono abilitate le chiamate d'emergenza e verso i numeri speciali (ad esempio quelli con prefisso 199 e 848) non effettuabili al momento su linee VoIP, che possono invece transitare sulla linea PSTN (o ISDN).

Caratteristiche minime della linea per ottimizzare l'uso del VoIP

Per scegliere adeguatamente un servizio ADSL sul quale far transitare il traffico VoIP è necessario valutare alcune caratteristiche specifiche del servizio ADSL.

Solitamente i provider indicano nelle proprie offerte l'**MCR** (*Minimum Cell Rate*) cioè la velocità minima garantita dalla connessione, sotto il quale il provider garantisce di non scendere mai. La si può esprimere in Kbps o KBps per cui una banda di 10 Kbps equivale a circa 1,25 kBps (8 Kbit = 1 KByte).

Ad esempio una linea un max di 640 Kbps e un MCR di 10 Kbps, al massimo avrà una velocità di 640 Kbps e non scenderà mai sotto i 10Kbps.

Il calcolo dell'MCR dipende dal numero di utenti connessi nello stesso momento sul canale di appartenenza, solitamente un canale di classe 1 con una portata di 1 Mbps di banda complessiva viene assegnato a 100 utenti, se fossero tutti connessi nello stesso istante, 1 Mbps diviso 100 è uguale 10 Kbps, in questo modo la minima banda viene garantita.

Più spesso viene pubblicizzata una linea solo per il suo **PCR** (*peak cell rate*) cioè la velocità di picco che può raggiungere la connessione, ma questo ovviamente non dà garanzie sulla qualità del servizio.

Altra modalità di configurazione indicata raramente dai provider ma di notevole importanza è se la linea è **Fast** o **Interleaved**; la prima è una modalità del collegamento ADSL particolarmente efficace per le applicazioni online di giochi, streaming audio/video e VoIP, e consiste nel garantire una latenza (tempo medio utilizzato per elaborare e trasmettere un dato) molto bassa, circa 15-30 millisecondi, durante lo scambio dei dati. La seconda, che invece è solitamente fornita dai provider più diffusi,

ha una latenza di circa 60-100 millisecondi. Questo incrementa la risposta del sistema anche durante la semplice navigazione in rete. Il maggior rendimento della modalità Fast è ottenuto grazie all'eliminazione del controllo d'errore nell'invio dei pacchetti di dati verso il server di connessione remota (controllo che invece è tipico dell'Interleaved). Molti provider addebitano all'utente un importo aggiuntivo per l'utilizzo della modalità Fast.

Infine un'altra caratteristica è il tipo di **indirizzo IP** fornito dal provider.

Gli indirizzi IP, ovvero Internet Protocol, sono dei codici formati da quattro gruppi di numeri che vanno da 0 a 255, tipo ad esempio: 192.168.178.20, separati da punti che permettono di identificare e contattare tutti i computer presenti nella Rete.

Ogni computer connesso ad Internet è contraddistinto da un indirizzo IP unico. Mentre si naviga, a prescindere dal tipo di connessione, al computer connesso viene assegnato un numero IP dal provider che ha permesso il collegamento.

Il provider possiede un range di numeri IP che assegna ai suoi utenti, in modo statico se è sempre lo stesso per tutta la durata del contratto, dinamico se viene assegnato in modo temporaneo dall'inizio della singola connessione e fino alla disconnessione.

Alla successiva ri-connessione verrà assegnato un altro indirizzo che in quel momento si è reso disponibile.

L'IP statico identifica univocamente l'utente a cui viene assegnato, non può quindi essere utilizzato per altre connessioni e rappresenta quindi un costo maggiore per il provider. Molti provider per questa ragione lo forniscono dietro il pagamento di un canone aggiuntivo o lo includono solo nelle offerte di fascia più alta.

L'utilità di avere una connessione con IP statico risiede nella possibilità di poter installare applicativi che richiedono un indirizzo fisso sul computer collegato alla linea, ad esempio un server web o FTP al proprio domicilio, probabilmente meno performante rispetto ad un servizio hosting, ma comunque possibile e oggi scelto da moltissime aziende che vogliono tenere sotto diretto controllo la posta e vogliono fornire ai propri clienti e ai propri partner aziendali servizi di download.

Per un utenza privata che come esigenza ha quella di una normale navigazione in rete è comunque sufficiente un IP dinamico.

4.3. I CODEC

Un CODEC è un programma o un dispositivo che si occupa di **CO**dicificare e/o **DE**Codificare digitalmente un segnale (tipicamente audio o video) perché possa essere salvato su un supporto di memorizzazione o richiamato per la sua lettura [10].

E' lo stesso tipo di programma utilizzato su PC o integrato in un componente hardware dedicato (ad es. nei lettori CD o DVD casalinghi o in alcune schede video/audio per PC) per la riproduzione di file audio.

Se si utilizza il PC per la comunicazione VoIP è necessario avere un software come Skype o Asterisk, e un **CODEC** (codificatore) di compressione della voce, che ovviamente deve essere presente nei PC di entrambi gli utenti interessati alla comunicazione.

Esistono vari tipi di CODEC, differenti tra loro per il tipo di segnale su cui devono operare e per l'algoritmo di codifica/compressione in essi implementato.

La **compressione audio** è una tecnica che permette di ridurre (anche di molto) le dimensioni di un file audio o la banda passante richiesta per una trasmissione audio.

Un file è una sequenza di cifre binarie (bit) utilizzata come veicolo di informazione. Comprimere significa ridurre il numero delle cifre che costituiscono la sequenza mantenendo l'informazione in un intorno dell'informazione originaria (ossia facendo in modo che la nuova informazione approssimi quella precedente).

I motivi che rendono necessaria o comunque utile la compressione sono:

- occupare minor spazio in fase di immagazzinamento
- impiegare minor tempo in fase di trasferimento dati.

Il costo della compressione è invece rappresentato dall'aumento dei tempi di lettura/scrittura legati rispettivamente ai tempi di decompressione/compressione. Nel caso di file audio si ha un costo anche in termini di qualità dell'audio.

Per realizzare tale compressione si fa ricorso alla riduzione della precisione dei colori dei singoli pixel (CODEC video) o delle frequenze da riprodurre (in alcuni CODEC audio vengono soppresse le frequenze non udibili dall'orecchio umano), alla eliminazione delle ridondanze o alla scrittura delle sole differenze (CODEC video) rispetto ad una immagine di riferimento.

Esistono due tipi di compressione:

- con perdita (*lossy*): quando l'informazione contenuta nel file compresso è minore di quella contenuta nel file di origine
- senza perdita (*lossless*): quando l'informazione contenuta nel file compresso è identica a quella contenuta nel file di origine

La prima permette compressioni maggiori a scapito della qualità sonora.

4.3.1. Compressione senza perdita

Usando un algoritmo di compressione senza perdita, dal risultato della compressione si può riottenere tutta l'informazione originaria. In questo caso la riduzione massima generalmente ottenibile, utilizzando algoritmi studiati appositamente per l'audio è all'incirca del 60%, ma solo con alcuni tipi di suono. Si possono utilizzare gli stessi algoritmi generali di compressione (come per esempio ZIP o Gzip) ma i risultati in termine di riduzione sono inferiori.

Esempio: **FLAC (Free LossLess Audio Codec)**, la cui compressione non rimuove informazioni dal flusso audio, ed è quindi adatto sia all'ascolto normale che per l'archiviazione. Il formato FLAC attualmente ha un buon supporto da parte di vari software audio, e raggiunge compressioni dell'ordine del 30-50% (gli algoritmi lossless tradizionali raggiungono soltanto un 10-20%).

4.3.2. Compressione con perdita

Dal risultato della compressione audio con perdita non si può più ottenere un suono identico all'originale ma la riduzione ottenibile è molto spinta: finché ci si mantiene con rapporti di compressione di 10 a 1, il risultato è quasi indistinguibile dall'originale, se si va oltre lo si fa a discapito della qualità.

Gli studi di psicoacustica (lo studio della percezione soggettiva dei suoni) hanno permesso di accertare che l'uomo non è sensibile allo stesso modo a tutte le frequenze da lui udite, e che un suono ad alta intensità ne maschera uno con frequenza vicina ma intensità più bassa. Sfruttando questi studi, si può pensare di eliminare l'informazione

che non verrebbe comunque percepita ed ottenere quindi un buon rapporto di compressione.

Ad esempio possiamo elencare i seguenti formati di compressione:

- mp3 (MPEG-1 Layer III): è stato introdotto negli anni '80 ed è il formato audio compresso più popolare. Essendo il più datato, è anche il meno efficiente e spesso il peggiore in termini di qualità.
- Windows Media Audio (WMA): è molto diffuso sui sistemi Windows.
- Ogg Vorbis è un CODEC più efficiente dell'mp3 ed è open source (ossia liberamente distribuibile e modificabile)
- AAC è stato reso popolare dalla Apple. Apple's iTunes Music Store fa uso di file compressi con 128Kbps CBR AAC e lo standard video MPEG4 raccomanda l'uso dell'AAC audio nei prossimi apparecchi e software.
- Dolby Digital (ac3) può comprimere fino a 6 canali audio, di cui 5 a piena larghezza di banda ed uno per gli effetti a bassa frequenza (LFE), fino a 384 kbit/s. Viene utilizzato nei DVD e nel sistema americano ATSC DTV.
- MPC o Musepack è un formato open source con una qualità maggiore dell'mp3 a parità di bitrate (vedi oltre).

In aggiunta, il CODEC oltre a comprimere la sequenza di dati, qualche volta cancella l'eco.

Un altro modo per ridurre la larghezza di banda impiegata è l'uso di un *soppressore di silenzio*. E' usato nel VoIP ed è uno dei suoi tanti vantaggi; consiste nell'interruzione della pacchettizzazione e nella trasmissione dei dati quando un utente non sta parlando, in altri termini il meccanismo di trasmissione viene sospeso quando da un lato della comunicazione c'è silenzio.

4.3.3. Algoritmi di compressione

Tra i formati di digitalizzazione e compressione più utilizzati nel VoIP si trovano [9]:

Quantizzazione differenziale (Differential PCM)

Generalmente il livello di due campioni consecutivi ha generalmente una differenza molto piccola. Nella quantizzazione differenziale, anziché trattare ciascun campione separatamente, viene quantificata e codificata la *differenza tra un campione e il suo precedente*. Così come il numero degli intervalli di quantizzazione necessario per quantificare la differenza tra due campioni consecutivi è minore di quella necessaria per quantificare un campione isolato, allora la quantizzazione differenziale permette di ridurre la frequenza di trasmissione essendo questa proporzionale al numero degli intervalli.

Dato che la codifica digitale è una sequenza di cifre (0 e 1), e che di solito in telefonia si usano 256 intervalli di quantizzazione per rappresentare tutti i possibili valori del campione (ad esempio nel CODEC G.711 A law e μ law), sono necessari 8 bits per rappresentare tutti gli intervalli ($2^8 = 256$). Altri CODEC che usano ***ADPCM*** o la ***Quantizzazione DELTA*** usano meno intervalli e quindi hanno bisogno di meno bit per codificare il campione.

Quantizzazione differenziale DELTA e ADPCM (DELTA Adattivo PCM)

Se si aumenta la frequenza di campionamento in una quantizzazione differenziale, i campioni consecutivi avranno pochissima differenza tra due valori. Quindi un singolo intervallo può essere usato per quantificare la differenza tra due campioni consecutivi.

Per rappresentare un intervallo è sufficiente impiegare un singolo bit; la velocità di trasmissione (vedi oltre: bit-rate) è uguale alla velocità di campionamento. Questo tipo di quantizzazione è conosciuta come “***DELTA***”.

Qui il livello di variazione in uscita è unico. In altri tipi di quantizzazione la variazione non è fissata e dipende dalla variazione dei segnali in ingresso. Per esempio ***ADPCM (DELTA Adattivo PCM)*** si basa sull'adattare la scala della quantizzazione in modo dinamico, in base alla piccola o grande differenza del segnale in ingresso.

Parametri che caratterizzano i CODEC

BitRate – E' la frequenza con la quale i bit sono trasmessi su un percorso di comunicazione. Di solito è espresso in kilobit al secondo (Kbps)

I file multimediali sono per loro natura connessi al tempo che scorre. In altri termini ad ogni secondo è associato un certo contenuto informativo e quindi una certa sotto-sequenza di cifre binarie. Il numero di cifre binarie che compongono queste sotto-sequenze è detto *bitrate*. In altre parole il *bitrate* è il numero di cifre binarie impiegate per immagazzinare un secondo di informazione. Questo può essere costante per tutta la durata del file o variare all'interno di esso. Ad esempio i cd musicali vengono campionati (registrati) ad una frequenza pari a 44.100 Hz. Da ciò si evince che ogni secondo si hanno 44.100 valori registrati dall'ipotetico microfono che vanno poi moltiplicati per i 2 canali del suono stereo che vanno a loro volta moltiplicati per 2 poiché la registrazione avviene a 16 bit (pari appunto a 2 byte). Quindi avremo:

$44.100 \times 2 \times 2 \times 60 \text{ (secondi)} = \sim 10 \text{ MB ogni minuto}$

La compressione, diminuendo la lunghezza globale del file, diminuirà di conseguenza la lunghezza media delle sotto-sequenze ossia diminuirà il bitrate medio. Il bitrate medio diventa dunque in questi casi l'indice dell'entità della compressione. Ad esempio se il file di origine possedesse un bitrate di 1411 Kbit/s e il file compresso possedesse un bitrate medio di 320 Kbit/s, allora avremmo ridotto di un fattore pari a circa 4.5.

Sampling Rate – (S.A.) il numero di campioni presi al secondo quando si digitalizza il suono. La qualità della riproduzione digitale migliora quando aumenta il numero di campioni nell'unità di tempo.

Frame size – il tempo che intercorre tra l'invio di due pacchetti successivi.

MOS - (*Mean Opinion Score*). E' una misura soggettiva della qualità del suono, cui si attribuisce un valore che va da 1 a 5.

Alcuni protocolli di compressione:

Cod.	ST.	Descrizione	BitR kb/s	S.A. kHz	Frame size ms	Remarks	MOS
G.711	ITU-T	Pulse code modulation (PCM)	64	8	Campionamento	μ -law (US, Japan) and A-law (Europe) companding	4.1
G.721	ITU-T	Adaptive differential pulse code modulation (ADPCM)	32	8	Campionamento	Ora descritto in G.726; obsoleto.	
G.722	ITU-T	7 kHz codifica audio entro 64 kbit/s	64	16	Campionamento	Subband-codec che divide una banda da 16 kHz in due sottobande, ognuna usa ADPCM	
G.722.1	ITU-T	Codifica a 24 e 32 kbit/s per operazioni a mano libera in sistemi con bassa perdita frame	24/32	16	20		
G.723	ITU-T	Estensione della raccomandazione G.721 ADPCM da 24 e 40 kbit/s per circuiti digitali	24/40	8	Campionamento	Superato da G.726; obsoleto ed è completamente diverso dal G.723.1	
G.723.1	ITU-T	Codificatore di doppia frequenza per le comunicazioni che trasmettono tra 5.3 e 6.3 kbit/s	5.6/6.3	8	30	Parte di H.324 video conferenza. Codifica la velocità di altri segnali audio in strutture usando l'analisi lineare predittiva-di sintesi. Per le alte frequenze si usa il codificatore Multipulse (MP-MLQ) per le basse frequenze il codificatore è il codificatore di previsione Algebrico Lineare (ACELP).	3.8-3.9
G.726	ITU-T	40, 32, 24, 16 kbit/s (ADPCM)	16/24/32/40	8	Campionamento	ADPCM; sostituisce G.721 e G.723.	3.85
G.728	ITU-T	Velocità 16 kbit/s basso ritardo previsione lineare	16	8	2.5	CELP.	3.6
G.729	ITU-T	Velocità di decodifica at 8 kbit/s usa strutture algebriche e CS-ACELP	8	8	10	Basso ritardo (15 ms)	3.92
GSM 06.10	ETSI	Impulso regolare previsione a lungo termine (RPE-LTP)	13	8	22.5	Usato per i cellulari GSM	
DVI	IMA	DVI4 usa un codice con modulazione di impulso variabile (ADPCM)	32	Varia bile	Campionamento		

ITU sta per **Unione Internazionale delle Telecomunicazioni**: è un'organizzazione internazionale che si occupa di definire gli standard nelle telecomunicazioni e nell'uso delle onde radio. È stata fondata il 17 maggio 1865 da 20 membri con il nome di *International Telegraph Union*, cambia con il nome attuale nel 1932. Dal 1947 è una delle agenzie specializzate delle Nazioni Unite e l'attuale sede è a Ginevra.

La ITU è divisa in tre settori:

- **ITU-R** (CCIR, *Comité Consultatif International des radiocommunications*, fino al 1993): regola le radiocomunicazioni
- **ITU-T** (CCITT, *Comité Consultatif International Téléphonique et Télégraphique*, fino al 1993): regola le telecomunicazioni
- **ITU-D** (BDT, *Bureau de Développement des Télécommunications*, fino al 1993): tenta di promuovere l'accesso alle telecomunicazioni nei paesi in via di sviluppo.

Tra gli altri organismi che si occupano di definire gli standard si trovano l'**IMA**, acronimo di Interactive Multimedia Association che ha tra l'altro definito il DVI e l'ETSI, che ha definito il GSM 6.10.

Uno standard ampiamente usato nel VoIP è ITU-T G.711, descritto nel prossimo paragrafo.

4.3.4. Lo Standard ITU-T G.711

Segue lo schema [9]:

- La banda della voce è 4 kHz, quindi abbiamo campioni alla velocità di 8 kHz (dal teorema del campionamento).
- Rappresentiamo ogni campione con 8 bit (quindi con 256 valori possibili).
- La velocità compressiva è $8000 \text{ Hz} * 8 \text{ bit} = 64 \text{ kbit/s}$, la tipica velocità di un canale telefonico digitale.
- Nelle applicazioni reali si utilizzano le varianti μ -law (Nord America) e a-law (Europa) che codificano il segnale analogico non linearmente ma su scala logaritmica sfruttando 12 o 13 bits invece di 8 bits e la differenza sta nel metodo di acquisire i campioni.

$$T=1/8000= 0.000125 \text{ s} = 125 \mu\text{s}$$

Quindi due campioni consecutivi di uno stesso segnale sono separati da 125μs. Questo è chiamato intervallo di campionamento.

L'A-law è formato da 13 segmenti dritti; segue la legge:

$$y = \begin{cases} \frac{Ax}{1 + \ln(A)} & \text{per } 0 \leq x \leq \frac{1}{A} \\ 1 + \frac{\ln(Ax)}{1 + \ln(A)} & \text{per } \frac{1}{A} \leq x \leq 1 \end{cases} \quad (4.1)$$

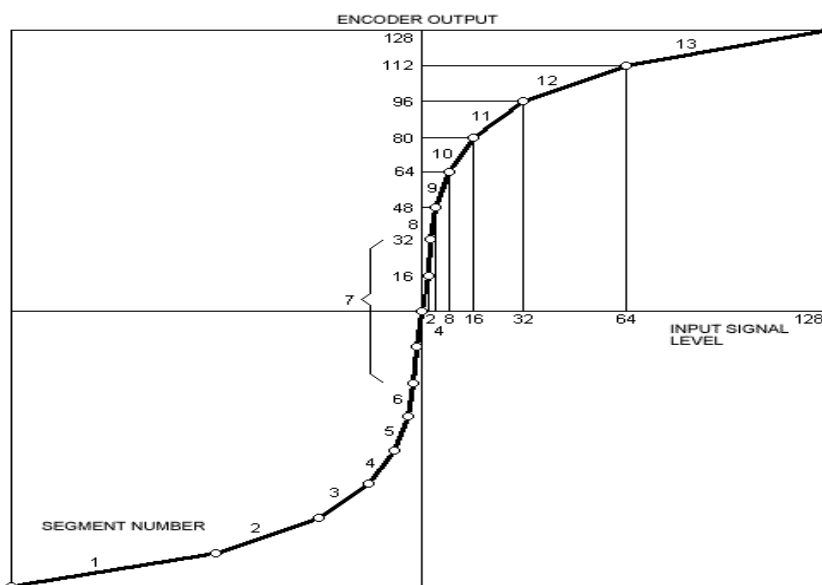
Il parametro A assume il valore di 87,6, x e y rappresentano i segnali di ingresso e uscita della compressione.

La compressione μ-Law segue invece la seguente legge:

$$y = \frac{\ln(1 + \mu x)}{\ln(1 + \mu)} \quad \text{per } 0 \leq x \leq 1 \quad (4.2)$$

dove il coefficiente è: $\mu=255$

Nella seguente immagine la codifica A-law è rappresentato graficamente:



4.4. RTP Protocollo di trasporto Real Time

È il protocollo che permette la distribuzione di servizi che necessitano di trasferimento in tempo reale [10].

È stato sviluppato da un gruppo di ricerca noto come **Audio-Video Transport Working Group**, facente capo alla IETF (Internet Engineering Task Force). Il corrispondente RFC (documento che riporta informazioni o specifiche riguardanti nuove ricerche, innovazioni e metodologie dell'ambito informatico o più nello specifico, di Internet. Attraverso l'Internet Society gli ingegneri o gli esperti informatici possono pubblicare dei memorandum, sotto forma di RFC, per esporre nuove idee o semplicemente delle informazioni che una volta vagliati dall'IETF possono diventare degli *standard Internet*) è stato pubblicato nel 1996.

RTP doveva inizialmente essere un protocollo multicast (distribuzione simultanea di informazione verso un gruppo di destinazioni), ma viene più spesso impiegato in applicazioni unicast (distribuzione verso un'unica destinazione). È basato sul protocollo UDP e viene usato in congiunzione con RTCP (RTP Control Protocol) che controlla la qualità del servizio e trasporta le informazioni riguardo ai partecipanti ad una sessione. RTCP è sufficiente per sessioni “loosely controlled”, in cui cioè non c'è un reale controllo dei partecipanti e set-up della sessione, e non è necessario che tutti i requisiti di controllo siano soddisfatti. Per far questo RTP può essere coadiuvato da un protocollo apposito per la gestione delle sessioni (come SIP o H.322). RTP rappresenta una delle tecnologie fondamentali nell'industria della telefonia su IP.

Partiamo dall'analisi dell'UDP:

L'**User Datagram Protocol (UDP)** è uno dei principali protocolli della suite di protocolli Internet. UDP è un protocollo di trasporto a pacchetto. È usato di solito in combinazione con il protocollo IP.

A differenza del TCP, l'UDP non gestisce il riordino dei pacchetti né la ritrasmissione di quelli persi. Per questa ragione risulta essere un protocollo di rete abbastanza inaffidabile, ma d'altro canto molto rapido ed efficiente per le applicazioni “leggere” o “time-sensitive”. Dato che le applicazioni in tempo reale spesso richiedono un ritmo minimo di spedizione, per cui non deve essere eccessivamente lenta la trasmissione dei pacchetti ed è tollerabile qualche piccola perdita di dati, allora per la trasmissione di

informazioni audio o video si preferisce utilizzare il protocollo UDP che segue queste caratteristiche, a discapito del più affidabile ma più lento modello di servizio TCP.

Per quanto detto UDP fornisce soltanto i servizi basilari del livello di trasporto, ovvero:

- moltiplicazione delle connessioni, ottenuta attraverso il meccanismo delle porte;
- verifica degli errori mediante una checksum, inserita in un campo dell'intestazione del pacchetto;

oltre questo TCP garantisce anche il trasferimento affidabile dei dati, il controllo del flusso e il controllo della congestione.

UDP è un protocollo *privo di stato*: non mantiene cioè lo stato della connessione dunque rispetto a TCP ha informazioni in meno da memorizzare. Un server dedicato ad una particolare applicazione che sceglie UDP come protocollo di trasporto può supportare molti più client attivi.

4.4.1. Struttura di un datagramma UDP

Un datagramma UDP è composto da 5 campi di cui due sono opzionali. I campi sono:

Porta Sorgente	Porta Destinazione
Lunghezza	Checksum
Dati	

Dove:

Porta Sorgente contiene la porta sorgente da cui proviene il datagramma (Lunghezza: 16 bit)

Porta Destinazione contiene la porta destinazione a cui deve essere spedito il datagramma (Lunghezza: 16 bit) .

Lunghezza contiene la lunghezza totale del datagramma UDP (testata+dati) (Lunghezza: 16 bit)

Checksum contiene il codice di controllo del datagramma (testata+dati). L'algoritmo di calcolo è definito nell'RFC del protocollo (Lunghezza: 16 bit)

Dati contiene i dati del datagramma (Lunghezza: quanto occorre per il dato, in teoria può essere infinita)

4.4.2. Analisi del protocollo RTP

Una volta che si hanno i dati, codificati in digitale e compressi dal CODEC, dobbiamo incapsularli ed inviarli in rete affinché raggiungano il destinatario. Seguiamo la seguente struttura [1]:

VoIP pacchetti di dati

RTP

UDP

IP

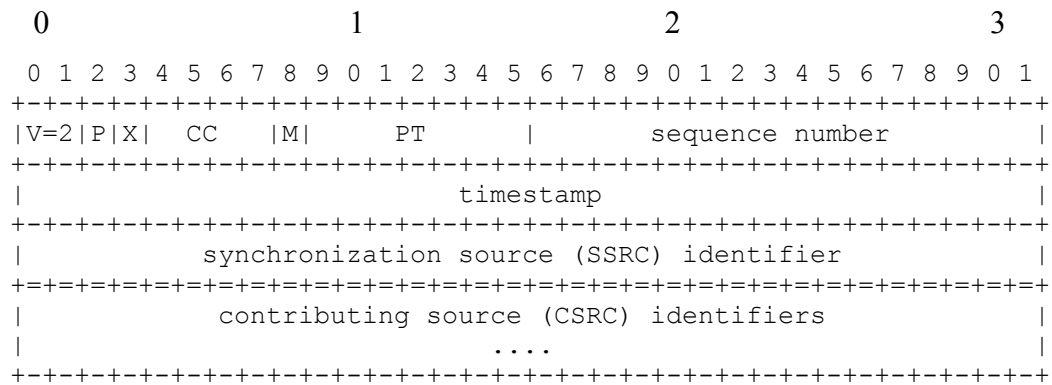
I,II layers

I Pacchetti di voce risiedono in pacchetti RTP (Protocollo di trasporto Real-Time) che a loro volta giacciono su pacchetti UDP-IP.

Con l'UDP non possiamo implicitamente controllare l'ordine di arrivo dei pacchetti o quanto impiegano ad arrivare (concetto di datagramma): queste informazioni sono molto importanti per ottenere una buona qualità audio della voce (per poter distinguere le parole) e una buona qualità di conversazione (la facilità con cui si segue un discorso).

Il protocollo RTP risolve allora il problema permettendo al ricevente di ordinare i pacchetti in base all'ordine di arrivo (metodo FIFO di recupero dati da uno stack : first in first out) e di tralasciare gli eventuali pacchetti persi: affinché una conversazione sia comprensibile abbiamo realmente bisogno di ogni singolo pacchetto di voce, purché però la ricezione dei pacchetti avvenga mediante un flusso continuo, e ovviamente, venga conservato il corretto ordine di invio.

Lo schema del RTP è il seguente



Dove:

- V identifica la versione di RTP
- P indica il padding, un byte di riempimento.
- X indica la presenza dell'intestazione estesa.
- Il campo CC rappresenta il numero di identificatori CSRC che seguono l'header a lunghezza fissa. Per esempio il campo CSRC viene utilizzato in caso di conferenza.
- M é un bit per marcare.
- PT indica il tipo di PayLoad (carico utile dei dati)

I numeri di sequenza (sequence number) che troviamo nel protocollo RTP permettono all'utente che riceve i dati di ricostruire la sequenza dei pacchetti del mittente.

4.5. Protocolli di segnalazione

Una volta che i pacchetti di dati sono incanalati nella rete, è necessario instradarli correttamente in modo tale da far raggiungere la corretta destinazione: hanno questo compito i protocolli per la codifica della segnalazione della conversazione (ricostruzione del frame audio, sincronizzazione, ecc.) che utilizzano alternativamente dei protocolli tra i quali hanno maggiore diffusione [2]:

- H.323 della ITU
- SIP (*Session Initiation Protocol*) della IETF

4.5.1. Il protocollo di segnalazione H.323

H.323 è uno standard creato nel 1996 dallo ITU-T Study Group 16. La sua creazione fu indotta dalla crescita della comunicazione multimediale su reti a commutazione di pacchetto.

Lo standard fu progettato specificamente con il seguente proposito:

- Basarsi sugli standard già esistenti, includendo H.323 e il protocollo di trasporto in tempo reale (RTP).
- Incorporare alcuni dei vantaggi che le reti di commutazione di pacchetti offrono per trasportare dati in tempo reale.
- Trovare una soluzione al problema che individuano le reti di commutazione di pacchetti per l'invio di dati in tempo reale.

Chi ha progettato H.323 sapeva che le caratteristiche della comunicazione sono differenti a seconda del luogo, degli utenti, delle compagnie telefoniche e del tempo. Visti questi fattori, i progettisti di H.323 lo programmarono in modo tale che i produttori degli impianti potessero aggiungere le proprie specificità al protocollo e potessero definire altre strutture di standard che permettessero ai dispositivi di acquisire nuove caratteristiche o capacità.

La prima versione di H.323 non garantiva la QoS. La comparsa e il successo del VoIP spinsero alla revisione di H.323 nel 1998.

Attualmente ci troviamo nella sesta versione (2006).

Fra le modifiche introdotte spiccano:

- fax-over-packet networks
- gatekeeper-gatekeeper communications
- fase-connection mechanisms

H.323 fornisce dettagliatamente le specifiche per definire un'infrastruttura per la trasmissione di dati multimediali, quali audio e video su una rete a commutazione di pacchetto quale ad esempio Internet; permette quindi ai seguenti elementi di dialogare tra loro:

TERMINALI

Ovvero client connessi alla rete, devono essere in grado di codificare e decodificare il parlato ed il video compresso con i CODEC (ad es. quelli precisati nelle specifiche ITU-T); così potranno dialogare con altri Terminali o con i Gateways (vedi il punto successivo)

GATEWAYS

O meglio *Voicegateways*, sono gli apparati posti sulla rete che fornisce comunicazione bidirezionale e in tempo reale tra un Terminale H.323 posto sulla LAN e un altro Terminale anch'esso posto in un segmento della rete; nel gateway possono essere presenti più interfacce alle quali vengono collegati dispositivi telefonici: (telefoni analogici o ISDN), in tali casi è il gateway a fungere da terminale per quel telefono;

GATEKEEPERS

Si occupano del controllo dello stato delle chiamate in corso e dell'utilizzo della banda, abilitano o disabilitano i terminali e i gateways interessati alla comunicazione; operano inoltre una conversione nome-indirizzo, per utilizzare i nomi al posto degli indirizzi IP. Controllano una zona della rete dove sono presenti terminali, telefoni e gateways.

Esempio di funzionamento del H.323

Tutti i terminali al momento di effettuare una chiamata, devono contattare il gatekeeper, mandano un pacchetto UDP broadcast, al quale il gatekeeper risponde inviando il suo indirizzo IP, a questo punto il terminale manda un pacchetto di informazioni chiamato RAS (registrazione-accettazione-stato) con il quale chiedono la registrazione. Quando questa viene accettata (il che equivale all'autenticazione del terminale) il terminale manda la richiesta per riservare la banda necessaria alla telefonata. Il gatekeeper accetta se la banda disponibile è sufficiente o rifiuta se ci sono già un quantitativo di telefonate in corso che saturano la banda: così viene garantita la qualità delle telefonate; a questo punto viene inoltrata la chiamata, se il terminale remoto accetta la chiamata il gatekeeper avverte il terminale e lascia il controllo della chiamata, e i due terminali in gioco iniziano a negoziare attraverso il canale di controllo della chiamata sul CODEC da usare per la compressione. Se i due terminali trovano un accordo vengono utilizzati due canali RTP e uno RTCP per il controllo delle congestioni e nel caso di trasmissioni video per la sincronizzazione tra audio e video. Per segnalare il termine della comunicazione si utilizza infine uno specifico canale (precisamente il Q.931).

Tra le applicazioni che utilizzano il protocollo H.323 spicca Microsoft NetMeeting e Net2Phone.

4.5.2. Il protocollo di segnalazione SIP

Nasce dall'esigenza di semplificare l'H.323, poiché quest'ultimo, essendo costruito in modo tale da mantenere la compatibilità con molti protocolli già esistenti, impone delle raccomandazioni atte a far funzionare il VoIP che sono molto complesse da seguire. Risulta per questo motivo un sistema chiuso che mal si integra con molte applicazioni.

SIP è stato creato dallo IETF ed è rappresentato da un canale per lo scambio di brevi messaggi in formato **testo** in grado di gestire la segnalazione della chiamata ed il suo controllo. Non spiega invece come devono essere realizzati i gateways per la connessione dei dispositivi SIP alle altre apparecchiature. Viene ridotto il numero di porte per le varie segnalazioni e così viene ridotto il tempo necessario per instradare la chiamata. Inoltre il fatto che le segnalazioni vengano fatte attraverso messaggi di testo rende possibile realizzare implementazioni del protocollo in maniera più semplice e controllare che il sistema funzioni correttamente.

E' un protocollo in fase di evoluzione, poiché si è adeguato alle modifiche che nell'utilizzo si sono rese necessarie; questo aspetto però crea delle incompatibilità tra diverse versioni del protocollo, con conseguenti difficoltà alla comunicazione tra molti terminali.

Utilizzo di SIP

I messaggi di testo delle segnalazioni utilizzano altri protocolli Internet. Un client SIP è infatti identificabile da una URL del tipo: sip:user@domani.com: è possibile effettuare una chiamata da pagina web semplicemente registrando nel browser un'applicazione in grado di gestire il protocollo SIP, “cliccando” sul link.

Il protocollo SIP permette di gestire il trasferimento di chiamata nel caso di occupato o di mancata risposta, di identificare chiamante e chiamato; può gestire conferenze multimediali e distribuire contenuti multimediali.

Per svolgere questi servizi SIP deve essere in grado di creare, modificare e chiudere sessioni con uno o più partecipanti, all'interno delle quali avvengono le applicazioni gestibili dal SIP.

SIP funziona nel seguente modo:

I chiamanti ed i chiamati sono identificati con un univoco indirizzo; quando viene fatta una chiamata, il chiamante prima individua il server del chiamato, poi gli invia una richiesta, un **invito**. Tale richiesta può essere inviata direttamente al destinatario o al proxy (il proxy è un programma che si interpone tra il server e il client, e fa sostanzialmente da intermediario delle comunicazioni fra i due) che poi inoltrerà la richiesta al destinatario.

La struttura della richiesta è la seguente:

COMANDO	URI RICHIESTA	VERSIONE SIP
---------	---------------	--------------

Dove:

COMANDO è il tipo di messaggio che si intende inviare alla risorsa rappresentata dalla URI possono essere *Invite, Ack, Option, Cancel, Register, Bye*.

INVITE avvia una chiamata

ACK conferma un messaggio ricevuto

OPTION caratteristiche supportate dall'inviante

CANCEL cancella le ricerche o lo stato di ringing

REGISTER registra presso un server

BYE termina o trasferisce la chiamata

URI (Uniform Resource Identifier) o più specificamente una **URL (Uniform Resource Location)** rappresenta l'utente o il servizio a cui la richiesta viene inviata.

VERSIONE SIP è quella usata dal mittente.

Il formato di un messaggio di risposta è:

VERSIONE SIP	CODICE DI STATO	INDICAZIONE REGIONE
--------------	-----------------	---------------------

VERSIONE SIP è la versione usata dal ricevente

CODICE DI STATO è il risultato del tentativo di interpretare e soddisfare la richiesta; è composto da 3 cifre decimali che indicano:

- **1xx** in ricerca, squillo, accodato
- **2xx** successo
- **3xx** forwarding (re-inoltro)
- **4xx** errori lato client
- **5xx** errori lato server
- **6xx** occupato, rifiutato, non disponibile

INDICAZIONE REGIONE: una descrizione testuale del codice di stato

Entità

Le **entità** che organizzano e strutturano insiemi di client qui sono:

SIP USER AGENT

E' un end-point (terminale) e può fungere da client o da server; i due ruoli sono dinamici, nel senso che nel corso di una sessione un client può fungere da server e viceversa. Quando funge da client, dà inizio alla transazione originando richieste. Quando funge da server, *ascolta* le richieste e le soddisfa, se possibile. E' in sostanza una macchina a stati, che si evolve in funzione dei messaggi SIP, e registra informazioni rilevanti del dialogo. Il dialogo ha inizio quando si risponde positivamente al messaggio di *Invite* e termina con un messaggio di *Bye*.

REGISTRAR SERVER

E' un server che accetta richieste del tipo REGISTER e quindi inserisce le informazioni che riceve nelle richieste del servizio di locazione del dominio che gestisce. Ad esempio per il dominio *domain.com* potrebbe esistere un *registrar* con opportune regole per la registrazione dei suoi membri, come il controllo di *login* e *password*.

SIP offre inoltre un servizio di **ricerca utenti**, se un utente vuole iniziare una sessione

con un altro utente, il *registrar* deve sapere dove può trovare l'host che gestisce l'utente chiamato. La ricerca degli utenti è compiuta da elementi di reti come proxy server e server per la redirectione (*redirect*) che sono responsabili, dopo la ricezione di una richiesta, di determinare dove essa va inviata, basandosi sulla conoscenza che hanno della locazione dell'utente. Per fare ciò i vari dispositivi supportanti SIP consultano un servizio astratto chiamato servizio di locazione il quale fornisce il legame **utente-indirizzo IP** o il riferimento ad altre *URI*, in modo analogo al funzionamento dei sistemi **DNS** (*Domanin Name System*) di Internet.

PROXY

E' un'entità intermedia che agisce sia da server che da client allo scopo di fare richieste per conto di altri client. Un server proxy fondamentale ritrasmette in maniera intelligente le richieste, inviando ogni richiesta ad un'entità più vicina all'utente destinatario del messaggio. I proxy possono essere utilizzati anche per fissare determinati criteri protezione come, ad esempio, l'obbligo per un determinato utente di autenticarsi prima di effettuare una chiamata.

Un proxy interpreta e se necessario riscrive parti specifiche di una richiesta prima di inoltrarla.

Risulta particolarmente utile legare un proxy ad un particolare dominio e proprio per questo motivo la maggior parte dei dispositivi con supporto SIP è in grado di tradurre i nomi dei domini attraverso l'accesso a server DNS.

REDIRECT SERVER:

E' un server "user agent" che genera responsi di tipo 3xx (forwarding, cioè redirectione, re-inoltro). Indicano di contattare un differente insieme URI. Questo consente ad un server di dare informazioni utili allo UA e di far cadere la nuova connessione aperta con il client. Si può far in modo che dopo un certo numero di connessioni un proxy, raggiunto il carico massimo, re-inoltri le ulteriori richieste verso un altro proxy.

Gestione della chiamata

Quando si genera una chiamata verrà generato dal UA un messaggio di tipo *Invite* verso il proxy in cui è registrato. Se l'indirizzo non è riconosciuto dal proxy (poiché non è un UA registrato presso di lui) viene inviato in un proxy di livello superiore finché non viene riconosciuto. A questo punto il destinatario può contattare direttamente il mittente (grazie all'indirizzo contenuto nei campi del messaggio SIP) o accettare la chiamata inviando un messaggio al proxy; una volta giunta la conferma al mittente la conversazione può iniziare .

4.5.3. Confronto tra H.323 e SIP

H.323 è un protocollo più stabile e collaudato, è reso compatibile con le precedenti versioni ogni volta che vengono apportate modifiche; prevede una codifica binaria che è più compressa rispetto a quella ASCII dei testi di SIP; d'altro canto è molto più costoso di SIP, che appare invece più flessibile e permette modifiche più rapide e semplici per gli sviluppatori, sebbene allo stato attuale non sia ancora completo e quindi supportato da ogni terminale.

H.323 utilizza il protocollo TCP per la segnalazione e il controllo della chiamata, le richieste RAS invece utilizzano l'UDP.

SIP invece utilizza solo il protocollo UDP.

H.323 utilizza l'architettura telefonica esistente (quindi i numeri tradizionali +39 078132148) mentre SIP utilizza indirizzi Internet esistenti (le destinazioni delle chiamate sono URL) quindi un numero deve essere trasformato nel nome di un host invertendo i numeri, separandoli con punti e aggiungendo il suffisso e164.arpa (8.4.1.2.3.1.8.7.0.9.3.e164.arpa).

Sia SIP che H.323 si basano su protocolli già esistenti per lo scambio dei frame dati veri e propri e cioè RTP e RTCP.

La scelta sull'uno o l'altro viene fatta in base alle esigenze, valutando gli aspetti menzionati.

4.6. QUALITA' DEL SERVIZIO

Il termine Qualità di Servizio o più semplicemente **QoS** (dall'inglese *Quality of Service*) è usato per indicare i parametri rilevati per valutare la qualità del servizio offerto dalla rete (ad esempio perdita di pacchetti, ritardo), o gli strumenti necessari ad ottenere una qualità di servizio desiderata [10].

La qualità del servizio è in genere proporzionale alle risorse impegnate per realizzare e gestire la rete, e inversamente correlata con il traffico in transito nella rete stessa.

Il traffico in transito nella rete e i tempi di intervento per i malfunzionamenti seguono un andamento stocastico, pertanto per una trattazione completa del problema è necessario ricorrere all'uso di variabili casuali.

Spesso nella stipula di un contratto VoIP vengono presi accordi sul livello di qualità del servizio (SLA), e se l'azienda che eroga il servizio non rispetta tale accordo è tenuta a pagare all'utente delle penali.

Nel campo della telefonia i parametri rilevanti sono:

- livello di rumore sul circuito
- livello sonoro
- probabilità di trovare una linea libera per iniziare una comunicazione
- probabilità di interruzione indesiderata di una comunicazione
- disponibilità del servizio
- durata media e massima dei disservizi

Come visto le telefonate in VoIP prevedono sostanzialmente la trasmissione di pacchetti di dati. Partiamo quindi dall'analisi dei pacchetti di dati per valutare la QoS di una telefonata.

4.6.1. Pacchetti di dati

In una rete a commutazione di pacchetto l'informazione da trasmettere viene suddivisa in pacchetti di dimensione abbastanza piccola; ad ognuno di essi viene aggiunta un'intestazione che contiene tutta l'informazione necessaria affinché il pacchetto possa venire inoltrato alla sua destinazione finale, e la sua posizione all'interno dell'informazione totale che viene trasferita. I pacchetti vengono inviati individualmente attraverso la rete e vengono poi riassemblati nella loro forma originale quando arrivano a destinazione.

L'intera capacità trasmissiva disponibile viene impegnata per la trasmissione di ciascun pacchetto. Se vi sono più pacchetti da trasmettere contemporaneamente, questi vengono memorizzati in una coda, e subiscono così un ritardo di accomodamento o rischiano di essere del tutto scartati in caso di esaurimento della memoria disponibile per la coda.

4.6.2. Commutatori

La **commutazione di pacchetto** è una tecnica di accesso multiplo (detto anche *multiplazione*) suddivisa nel tempo e utilizzata per condividere un canale di comunicazione tra più stazioni in modo casuale, generalmente impiegata per realizzare reti di calcolatori. Si distingue dalla commutazione di circuito, che è tipicamente usata nelle comunicazioni telefoniche.

Tali tecniche non comportano l'attivazione di una linea di comunicazione dedicata fra un elaboratore ed un altro, ma consentono lo svolgimento simultaneo di più comunicazioni fra computer, massimizzando così l'utilizzazione dei mezzi trasmissivi impiegati.

Quando un nodo intermedio detto **commutatore di pacchetto**, generalmente un router o uno switch, riceve un pacchetto, esso decide quale è il percorso migliore che il pacchetto può prendere per raggiungere la sua destinazione. Questa strada può cambiare da pacchetto a pacchetto dipendentemente dalle condizioni della rete, per cui pacchetti appartenenti ad uno stesso messaggio possono intraprendere anche percorsi distinti.

Mentre in una rete a commutazione di circuito la capacità del canale trasmissivo è interamente dedicata ad una specifica comunicazione, la commutazione di pacchetto si rivela molto più efficiente nonostante la maggior quantità di dati inviata, in quanto i canali fisici sono utilizzati solo per il tempo strettamente necessario e vengono liberati

non appena termina una singola trasmissione di dati. Inoltre, poiché ogni pacchetto porta con sé la sua identificazione, una rete può trasportare nello stesso tempo pacchetti provenienti da sorgenti differenti.

La commutazione di pacchetto permette quindi a più utenti di inviare informazioni attraverso la rete in modo efficiente e simultaneo, risparmiando tempo e costi mediante la condivisione di uno stesso canale trasmissivo (cavo elettrico, etere, fibra ottica, ecc.).

Storicamente la commutazione di pacchetto poneva qualche problema quando era necessaria una disponibilità garantita di banda o nelle trasmissioni *real time* (si pensi a una trasmissione video, dove le immagini arrivano con un flusso costante). Al giorno d'oggi è però possibile aggiungere una "priorità" ai pacchetti per garantire che un numero sufficiente di essi venga inviato, a scapito di altri pacchetti che non abbiano un'urgenza specifica – come, ad esempio, un file da trasferire.

4.6.3. Ritardi

Un pacchetto che attraversa una rete subisce un ritardo dovuto in parte alle caratteristiche del percorso che sta seguendo ed in parte allo stato di carico della rete.

RITARDO DI TRASMISSIONE

Il pacchetto di dati deve essere trasmesso, la velocità con cui viaggia è quella determinata dalla velocità della linea di trasmissione. Quindi considerando la lunghezza del pacchetto stesso si avrà un tempo di trasmissione dato dal rapporto tra la lunghezza del pacchetto e velocità della linea. Molti commutatori utilizzano la tecnica **store-and-forward**, ovvero ricevono tutto il pacchetto, verificano i checksum (codici di controllo) per controllare se ci sono stati errori di trasmissione, e poi lo accodano per la ritrasmissione. Pertanto i tempi di trasmissione dei commutatori attraversati devono essere sommati tra loro, e si genera così il ritardo di trasmissione.

Nel dettaglio:

- **ritardo di propagazione** - il tempo necessario al segnale fisico per propagarsi lungo una linea di trasmissione. Nei mezzi trasmissivi come rame o fibra ottica, i

segnali si propagano alla velocità della luce calcolata nello specifico mezzo trasmissivo, quindi minore di 200.000 km/s. Quindi il ritardo di trasmissione è stimabile in 5 μ s/km.

- **ritardo di coda o accomodamento** (*queuing delay*) – è dovuto al fatto che i pacchetti in uscita non sempre vengono trasmessi immediatamente, quando ad esempio la linea in uscita è occupata da altri pacchetti in corso di trasmissione. In questo caso, il pacchetto viene salvato in una memoria temporanea del commutatore detta **coda**, per venir trasmesso appena possibile. Il tempo atteso dal pacchetto nella coda è appunto detto ritardo di coda, e non è possibile sapere con certezza quanto ritardo di coda subirà un pacchetto. E' possibile solo farne una stima che si basa su considerazioni statistiche: si parla in tal senso di **teoria delle code**.

È anche possibile che i pacchetti inviati tra due stessi host in momenti diversi subiscano ritardi differenti. La variazione del ritardo subito dai pacchetti è detta **jitter** (letteralmente *nervosismo*), cioè l'errore nella base dei tempi quando un campione digitale è convertito in segnale analogico.

Il **jitter** è una brusca ed indesiderata variazione di una o più caratteristiche di un segnale, come l'intervallo, l'ampiezza, la frequenza o la fase di due impulsi successivi. E' un fenomeno che ha molta importanza nella progettazione dei bus per comunicazioni (per esempio l'USB) tra PC e periferiche.

Se addirittura la rete è molto congestionata e la coda non riesce a contenere tutti i pacchetti, il commutatore è costretto a scartarli. Per questa ragione, una rete a pacchetto non può generalmente garantire che tutti i pacchetti inviati arrivino a destinazione.

4.6.4. Multiplazione e teoria delle code

Lo studio di entrambi i sistemi di multiplazione sfrutta ampiamente la teoria delle code.

Ad esempio:

- nella commutazione di circuito si valuta la probabilità che una telefonata non possa essere servita a causa della congestione di fascio di linee o di un commutatore. I serventi sono le linee disponibili in un fascio, i clienti sono rappresentati dalle telefonate, i tempi di servizio dalla durata delle telefonate.

- nella commutazione di pacchetto si studia la probabilità che un pacchetto possa essere perso perché la coda di trasmissione è piena, e le caratteristiche statistiche del ritardo di accodamento a cui andrà incontro. I server sono le linee di trasmissione, i clienti sono i pacchetti, i tempi di servizio sono quelli necessari per la trasmissione dei pacchetti.

Il comportamento statistico delle code in una rete determina le sue caratteristiche di QoS.

4.6.5. Multiplazione inversa

Come dice il nome, la **multiplazione inversa** è il meccanismo per cui la capacità di *più* collegamenti viene aggregata per costruire un canale trasmissivo di capacità superiore. Questo comporta normalmente l'uso di protocolli per distribuire i dati da trasmettere sui collegamenti disponibili, per ricomporre i dati ricevuti sui collegamenti aggregati nell'ordine corretto, per gestire i guasti e il ritorno in servizio di parte dei collegamenti.

4.6.6. Consegna fuori ordine

(out-of-order) – su alcune reti, è possibile che una sequenza di pacchetti inviati da un nodo ad un altro venga consegnata in un ordine diverso da quello originale. Questo accade tipicamente perché i pacchetti vengono instradati su percorsi diversi. Per risolvere questo problema è necessario che i protocolli di trasporto riordinino i pacchetti fuori ordine una volta giunti a destinazione, e comporta ulteriori ritardi nella ricostruzione del flusso di dati a livello applicativo.

La probabilità che un pacchetto arrivi fuori ordine è un altro parametro rilevante per la valutazione della QoS.

4.6.7. Errori di trasmissione

Un pacchetto può essere consegnato a destinazione, ma non essere identico a quello inviato. Molte reti riconoscono la gran parte degli errori di trasmissione, e alcune sono anche in grado di correggere tali errori.

Anche in questo caso si considera la percentuale di pacchetti errati in relazione alla QoS. Normalmente i protocolli di trasporto riconoscono un pacchetto errato e ne richiedono la ritrasmissione (come se questo fosse stato perso), ma a volte capita che l'errore raggiunga l'applicazione finale.

Quando è stata creata Internet non era stata percepita la necessità di QoS per le applicazioni. Infatti l'intera Internet segue la filosofia del best effort, cioè il sistema garantisce di fare tutto il possibile per portare a termine un'operazione, ma non garantisce affatto che l'operazione verrà compiuta, né in che modo.

Ci sono fondamentalmente due modi per fornire garanzie sulla Qualità di Servizio.

4.6.8. Overprovisioning

Il primo metodo, detto **overprovisioning**, consiste nel fornire risorse in abbondanza, abbastanza da soddisfare la domanda di *picco* attesa, con un sostanziale margine di sicurezza. Una soluzione semplice, ma in pratica sia troppo costosa, e non è applicabile se la domanda di picco cresce più velocemente di quando predetto, poiché disporre nuove risorse richiede tempo.

4.6.9. Priorità

L'alternativa è amministrare la banda disponibile, facendo in modo che i pacchetti a cui deve essere garantita una certa QoS ottengano un trattamento privilegiato. Per ottenere questo, bisogna risolvere due problemi:

- identificare i pacchetti che devono ricevere un trattamento privilegiato.
- applicare a questi pacchetti una disciplina di coda che garantisca le prestazioni necessarie, applicata sulle porte di uscita dei router.

Classificazione

I metodi strutturati per identificare il traffico da privilegiare sono:

- “*integrated service*”, basato sulle prenotazioni: prima di iniziare una sessione

che ha dei requisiti di QoS, l'applicazione deve "chiedere" alla rete se può garantire le prestazioni necessarie. La rete valuta se dispone delle risorse adeguate, e in caso positivo accetta la prenotazione.

- “*differentiated services*”: prevede che gli utenti della rete stipulino a priori un contratto che definisce la quantità massima di traffico "privilegiato" che possono generare, e marchino il traffico a cui dare priorità utilizzando il campo Type of Service dell'header IP (rif. Protocollo RSVP, qui di seguito). Le prenotazioni quindi sono "statiche".

Soprattutto nelle reti di piccole dimensioni, è possibile utilizzare metodi più semplici, che prevedono di identificare manualmente sui router il traffico a cui dare priorità (ad esempio si può scegliere di privilegiare una telefonata a discapito di un'altra applicazione Internet).

RSVP

Nell'ambiente VoIP vengono utilizzati altri protocolli, come l'RSVP, che può gestire la Qualità del Servizio (QoS).

RSVP è un protocollo di segnalazione che permette di riservare una certa quantità di banda e di latenza massima in ogni nodo (router) di rete attraversato che lo supporta.

Le applicazioni VoIP richiedono un flusso di dati in tempo reale, ma il protocollo di rete non può *garantire* tale flusso (può solo *effettuare il massimo sforzo* per cercare di conseguirlo). E' quindi necessario introdurre degli artifici e delle politiche di schedulazione che possano gestire i pacchetti in OGNI router che attraversano.

Si introduce allora il campo TOS (Type Of Service) nel protocollo IP per descrivere il tipo di servizio: valori bassi indicano bassa urgenza, mentre valori via via sempre più alti segnalano urgenza crescente.

Discipline di coda

In un router che non applichi politiche di qualità di servizio, i pacchetti vengono trasmessi sulle porte in uscita nell'ordine in cui sono arrivati. Una disciplina di coda consiste invece nel gestire per ciascuna porta diverse code in uscita, in cui i pacchetti

vengono classificati e viene stabilito in quale ordine verranno prelevati i pacchetti dalle varie code.

Esempi di disciplina di coda:

- **priorità stretta:** le code sono ordinate per priorità. Ogni volta che si deve trasmettere un pacchetto, lo si preleva dalla coda a priorità più alta che avrà un pacchetto pronto da inviare. In questo modo, una applicazione di priorità superiore può monopolizzare l'intera banda disponibile, a danno di quelle di priorità inferiore.
- **weighted round robin:** viene prelevato a turno un pacchetto da ciascuna coda. In questo modo, si garantisce che tutte le classi di applicazioni potranno trasmettere. Il "weighted" significa che a ciascuna coda può essere attribuito un peso, ovvero una frazione della banda disponibile, e i pacchetti vengono prelevati in modo da garantire questa banda disponibile. Se una classe di traffico in un certo momento non utilizza la banda allocata, questa è utilizzabile dalle altre (si parla di *bandwidth borrowing*: banda prestata).
- Discipline di coda più avanzate, come *Hierarchical Packet Fair Queueing (HPFQ)* e *Hierarchical Fair Service Curve (H-FSC)*, permettono di esprimere per ciascuna coda sia un requisito sulla banda che uno sul ritardo. Al momento, sono disponibili solo su router software, basati su BSD o linux.

Altri strumenti utilizzati per amministrare la banda disponibile:

- **RED (Random Early Detection):** quando si approssima la congestione, la rete scarta arbitrariamente una piccola percentuale del traffico. Questo viene interpretato da UDP come una indicazione di congestione, per cui viene abbassata la quantità di traffico inviata. Un caso particolare di questa tecnica chiamato **WRED (Weighted Random Early Detection)** permette di distinguere il flusso di traffico dal quale iniziare a scartare i pacchetti in presenza di congestione. Con il WRED è possibile definire delle soglie di utilizzo del link che, una volta raggiunte provocano lo scarto di pacchetti appartenenti a specifiche classi di traffico. Così al raggiungimento della prima soglia verranno scartati solo pacchetti di flussi poco importanti, mentre al raggiungimento di soglie di utilizzo via via più alte verranno scartati anche pacchetti appartenenti a

flussi di traffico più importanti. Il "weighted" significa che la classe di traffico che sperimenterà il maggior numero di pacchetti perduti sarà quella associata alla soglia più bassa.

- **rate limiting:** una classe di traffico può essere limitata in modo che non utilizzi più di una certa banda.

Il mercato non ha ancora favorito la nascita di servizi QoS. Alcuni credono che una rete semplice che offra sufficiente banda per la maggior parte delle applicazioni e per la maggior parte del tempo, sia già economicamente la migliore soluzione possibile, mostrando poco interesse a supportare applicazioni non-standard capaci di QoS.

La rete Internet ha già accordi complessi tra i provider, e sembra che ci sia poco entusiasmo nel supportare il QoS attraverso connessioni che interessano reti appartenenti a provider diversi, o sugli accordi circa le politiche che dovrebbero essere sostenute al fine di poterle supportare.

Ci sono comunque provider che puntano sulla QoS e contemporaneamente si sforzano di mantenere i prezzi del servizio sempre più bassi (attualmente esistono servizi per i quali si paga solo un canone per la linea ADSL e su questo vengono comprese illimitate chiamate col VoIP verso le direttive nazionali).

Inoltre il mercato sta puntando verso la diffusione delle linee solo dati (senza cioè la parte analogica vista nel par. 3.2.3).

Dai dati di **AGCOM** (*Autorità per le Garanzie nelle Comunicazioni*) aggiornati al 27/03/07 emerge una buona prospettiva di diffusione del mercato VoIP: si passerà dall'attuale 5% al 10% nel 2009, e al 15% nel 2011 [8].

CONCLUSIONI

Come sempre accade nel mondo scientifico, si può affermare che anche negli sviluppi della tecnologia delle comunicazioni, e più in generale dell'informatica, il ruolo della matematica è basilare. Partendo da un teorema matematico si riesce a campionare la voce, digitalizzarla, comprimerla, trasportarla come sequenza di bit, farla arrivare a destinazione e ritrasformarla in modo tale da essere percepita come la voce iniziale. Senza il punto di partenza descritto dal teorema di Shannon questo non sarebbe possibile.

Ciò che si deduce anche in questo caso è che il compito della matematica continua oggi ad essere quello di leggere il linguaggio del Mondo reale, tradurlo in uno schema universalmente comprensibile, quindi elaborarlo e renderlo accessibile in modo tale che le altre scienze e la conseguente ricerca tecnologica lo possa mettere in pratica.

BIBLIOGRAFIA

- [1] Roberto Arcomano “*VoIP Howto*” v1.8 Agosto 2002 – pp. 4-9
- [2] Diego Gosmar, Giuseppe Innamorato, Dimitri e Stefano Osler, “*Asterisk e dintorni. La guida italiana al VoIP Open Source*” - (I edizione – Gennaio 2006) -- pp 45-62
- [3] Andrea Marzilli, “*La tecnologia ADSL: panoramica su xDSL*” art. su <http://voip.html.it>
- [4] MIT – “*MIT Professor Claude Shannon dies; was founder of digital communications*” - Articolo del 27 Febbraio 2001 - <http://web.mit.edu/newsoffice/2001/shannon.html>
- [5] C.D. Pagani-S. Salsa “*Analisi Matematica – Vol.2*” -- pp. 178-179
- [6] C.E. Shannon, “*A mathematical theory of Communication*” Bell System Technical Journal, Vol. 27 – pp 379-423 Luglio 1948; Vol. 27, pp. 623-656, Ottobre 1948.
- [7] C.E. Shannon, “*Communication in the presence of noise*”, IRE, Vol. 37, pp. 10-21, Gennaio 1949
- [8] Loredana Vajano – AGCOM “*I servizi IP: analisi di un mercato in evoluzione*” 27/03/07 – pp. 26-27
- [9] VoIPForo – <http://www.en.voipforo.com> - s.e.
- [10] Wikipedia – l'enciclopedia libera - <http://www.wikipedia.org> - s.e.
- [11] E.T. Whittaker, “*On the functione Which are Represented by the Expansions of the Interpolation Theory*”, Proc. Roy. Soc. Edinburg, 1915, pp. 181-194
- [12] J.M. Whittaker, “*Interpolatory function theory*” Cambridge Tracts of Mathematics and Matjematical Phisics, no. 33, Cambridge Univ. Press, 1935, cap. IV.
- [13] Aaron D. Wyner and Shlomo Shamai (Fellow, IEEE), *Introduction to “Communication in the Presence of Noise” by C.E. Shannon* -- pp. 442-446
- [14] Ahmed I. Zayed (Department of Mathematics, University of Central Florida, Orlando, Florida), *Advances in Shannon's Sampling Theory* -- pp. 1-17