

Reti di Telecomunicazioni

Studio delle prestazioni

Le reti locali si differenziano da quelle geografiche e dai sistemi multiprocessore dalla velocità di trasmissione (*data rate*) R usata e dalla distanza d del tratto di comunicazione: $R \times d$.

Sia V la velocità di propagazione del mezzo trasmissivo e L la lunghezza del frame dati.

Uno dei parametri più importanti è il seguente (Figura 1):

$$a = \frac{Rd}{VL} = \frac{\text{data - path - length}}{\text{frame - length}} = \frac{\text{propagation - time}}{\text{transmission - time}}$$

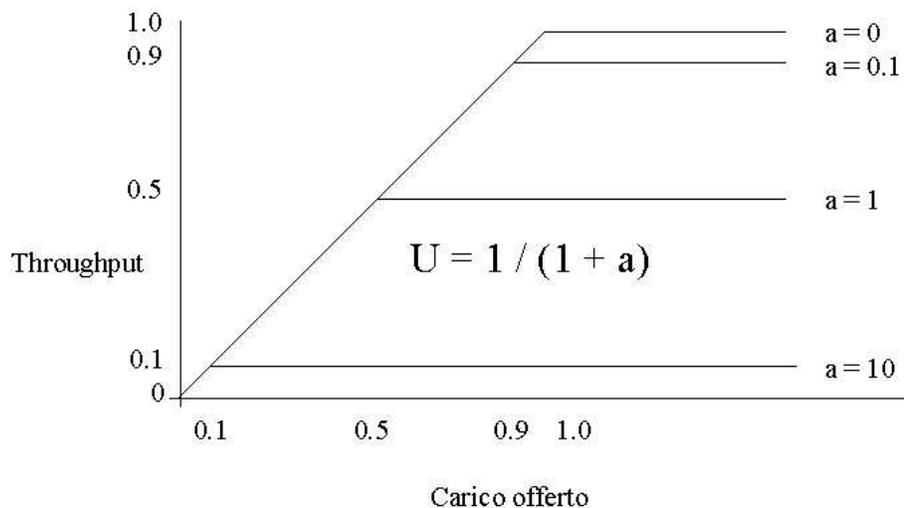


Figure 1: Effetto di a sul throughput.

Tecniche di accesso multiplo

Con accesso multiplo si intende la possibilità da parte di più utenti di condividere delle risorse fisiche per trasferire le informazioni. Il simultaneo servizio agli utenti del sistema di telecomunicazioni è reso possibile dalla suddivisione di una risorsa per esempio radio in canali elementari oppure limitando il tempo massimo in cui un terminale può trasmettere di continuo. Col termine **accesso multiplo** si indica perciò il fatto che alla risorsa condivisa possano accedere tutti gli utenti e non esistono canali dedicati permanentemente a singoli utenti ma essi sono assegnati dal sistema dietro richiesta degli utenti stessi. Possono esserci pure dei canali di controllo tramite i quali gli utenti effettuano le richieste di servizio ed il sistema in seguito assegna ad ognuno di essi risorse di comunicazione disponibili. Le due principali classi di protocolli di accesso sono:

- protocolli di accesso a contesa
- protocolli di accesso senza contesa

Protocolli di accesso a contesa

I protocolli di **accesso a contesa** fanno uso di metodi statistici per trasmettere l'informazione in pacchetti attraverso un accesso multiplo a divisione di tempo. Gli utenti devono contendersi l'uso degli slot temporali per effettuare le trasmissioni. Slot assegnati a utenti momentaneamente inattivi (periodi di silenzio durante la comunicazione) sono offerti ad altri utenti, i quali diventano attivi e hanno il permesso di *contendersi*, con una certa probabilità, l'uso degli slot liberi. Se è un solo utente a trasmettere su un certo slot allora la trasmissione va a buon fine, altrimenti, se più utenti si contendono lo stesso slot, i segnali ricevuti non possono essere separati quindi si ha una situazione di conflitto (detta *collisione*), che deve essere risolta attraverso l'applicazione di protocolli per la ritrasmissione dei pacchetti. La probabilità di collisione deve essere mantenuta al di sotto di una certa soglia (parametro fondamentale di questi protocolli di accesso) per evitare una degradazione della qualità percepibile del collegamento.

Il vantaggio dei protocolli di accesso a contesa è che permettono di sfruttare al meglio le risorse soprattutto nelle situazioni di traffico a “picchi” o traffico discontinuo, dove gli utenti hanno lunghi periodi in cui ascoltano o pensano seguiti da brevi momenti di traffico intenso.

Non sono comunque adatti per le applicazioni real-time a causa della loro penalità in termini di ritardo (a causa delle possibili collisioni).

Il protocollo ALOHA

Nel 1971 l'Università delle isole Hawaii iniziò ad usare il sistema ALOHA. Si faceva uso di un satellite come centro stella per interconnettere stazioni distribuite sulle diverse isole dell'arcipelago.

ALOHA è un protocollo ad accesso casuale, quindi del tipo a contesa. Il suo funzionamento era molto semplice:

1. *Transmission mode*: gli utenti possono trasmettere in qualunque momento;
2. *Listening mode*: dopo aver trasmesso, un utente si pone in attesa di un riscontro (ACK) dal ricevitore; in caso di collisione con altre stazioni si può ottenere un riscontro negativo (NAK);
3. *Retransmission mode*: con la ricezione di un NAK, un messaggio viene ritrasmesso ma dopo aver aspettato un ritardo casuale;
4. *Timeout mode*: se, dopo una trasmissione, non si riceve alcun riscontro entro un tempo massimo, si procede alla ritrasmissione del messaggio.

Il protocollo ALOHA: prestazioni

A causa delle collisioni, alcuni pacchetti verranno respinti. si definisce perciò frequenza totale di arrivo:

$$\lambda_t = \lambda + \lambda_r$$

come somma della frequenza dei pacchetti trasmessi con successo con quelli che hanno colliso e vengono ritrasmessi.

Sia di b bit la lunghezza di ogni pacchetto, da cui si definisce il *throughput* $\rho' = b\lambda$ sul canale in bit/s. Si definisce pure il *traffico totale* $G' = b\lambda_t$.

Se la capacità del canale è R bit/s, si definisce *throughput normalizzato* ρ :

$$\rho = \frac{b\lambda}{R}$$

e *traffico totale normalizzato* G :

$$G = \frac{b\lambda_t}{R}$$

Si noti che $0 \leq \rho \leq 1$ mentre $0 \leq G \leq \infty$.

Definendo inoltre il tempo di trasmissione di un pacchetto $\tau = b/R$, si ha che:

$$\rho = \lambda\tau$$

$$G = \lambda_t\tau$$

Un utente trasmette con successo il suo pacchetto se nessun altro utente inizia una trasmissione nei successivi τ secondi o non l'aveva iniziata nei precedenti τ s. Questo definisce la *finestra di contesa* che è pari a 2τ s.

Il processo degli arrivi al satellite considerando tutti gli utenti indipendenti tra loro può essere modellato dalla distribuzione di Poisson:

$$P(k) = \frac{(\lambda\tau)^k}{k!} e^{-\lambda\tau}$$

dove λ è la frequenza di arrivo media. Questa espressione viene usata per ricavare la probabilità che nessun altro utente trasmetta in nell'intervallo di tempo 2τ . La probabilità di successo P_s sarà quindi:

$$P_s = Prob(k = 0) = e^{-2\lambda_t\tau}$$

Ora, per definizione, $P_s = \frac{\lambda}{\lambda_t}$ da cui si ottiene:

$$\lambda = \lambda_t e^{-2\lambda_t\tau}$$

$$\rho = G e^{-2G}$$

che lega il throughput normalizzato ρ con il traffico totale normalizzato G .

Il massimo valore di ρ è pari a $\frac{1}{2e} = 0.18$ in corrispondenza a $G = 0.5$.

Il protocollo slotted ALOHA

Il puro ALOHA può essere migliorato mediante un minimo coordinamento tra le stazioni.

Con slotted ALOHA il satellite invia una sequenza di impulsi di sincronizzazione in modo broadcast a tutte le stazioni. La lunghezza dei pacchetti è ancora costante e ora devono essere inviati all'interno di un time slot dato dagli impulsi di sincronizzazione. Questo, di fatto, dimezza la finestra di congestione, ora pari a τ . Di conseguenza si ottiene:

$$\rho = Ge^{-G}$$

Ora il massimo valore di ρ è pari a $\frac{1}{e} = 0.36$ in corrispondenza a $G = 1$ con un raddoppio delle prestazioni.

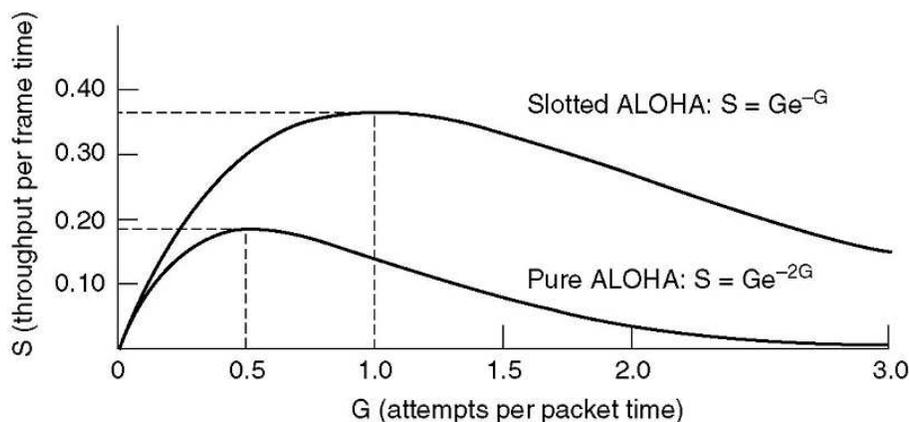


Figure 2: Throughput in ALOHA.

Carrier Sense Multiple Access (CSMA)

Topologia originaria: bus bidirezionale

Esistono tre possibilità:

- *non persistent* CSMA:
 1. se il mezzo è libero trasmetti, altrimenti vai al passo 2;
 2. aspetta un intervallo di tempo aleatorio (ritardo di ritrasmissione) ricavato mediante una distribuzione di probabilità e torna al passo 1;
- *1-persistent* CSMA:
 1. se il mezzo è libero trasmetti, altrimenti vai al passo 2;
 2. continua ad ascoltare il canale finché diventa libero poi trasmetti subito;
- *p-persistent* CSMA
 1. se il mezzo è libero trasmetti subito con probabilità p e con ritardo unitario (pari al ritardo di propagazione massimo) con probabilità $(1 - p)$;
 2. continua ad ascoltare il canale finché diventa libero poi torna al passo 1;
 3. se la trasmissione è ritardata di un'unità di tempo, torna al passo 1.

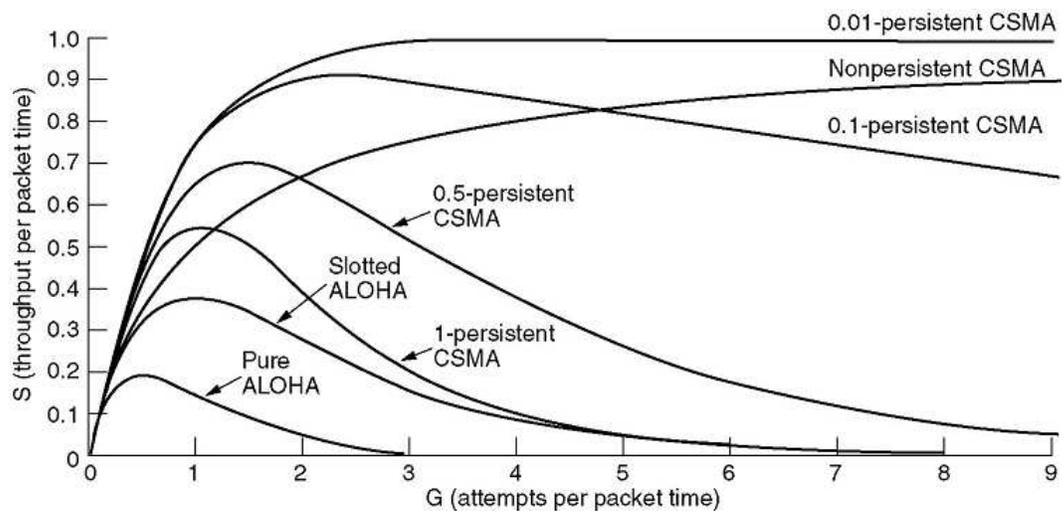


Figure 3: Confronto di prestazioni tra diverse tecniche di accesso casuale.

Carrier Sense Multiple Access with Collision Detection (CSMA/CD)

- Topologia originaria: bus bidirezionale;
- Bit rate: 10 Mbit/s;

Funzionamento in 3 passi:

- Carrier sensing: ogni stazione ascolta il bus e trasmette solo se il bus è libero;
- Multiple access: una volta iniziata la trasmissione una collisione può avvenire a causa del ritardo di propagazione;
- Collision detection: per rilevare una collisione la stazione trasmittente sta in ascolto sul mezzo; in caso di collisione, si ferma subito ed invia una particolare sequenza di bits (*jamming*) per informare tutte le altre stazioni dell'avvenuta collisione. In caso di collisione, la stazione ritenterà la trasmissione dopo un ritardo casuale determinato dall'algoritmo di back-off e la collisione non è un errore di trasmissione ma è il modo per gestire l'accesso multiplo.

CSMA/CD: throughput

Si supponga che il bus sia organizzato in time slot di durata doppia rispetto al ritardo di propagazione end-to-end. Il time slot rappresenta quindi il tempo massimo dall'inizio della trasmissione per rilevare una collisione.

Si abbiano N stazioni tutte uguali. ognuna può trasmettere in un time slot con probabilità p .

Vi sono due tipi di intervalli di tempo. Il primo è l'intervallo di trasmissione di durata $1/2a$ slot, assumendo unitario il tempo di trasmissione di un frame. Il secondo è l'intervallo di contesa, dato da una sequenza di slot con o collisione o nessuna trasmissione. Il throughput sarà dato dalla percentuale di tempo speso per gli intervalli di trasmissione.

Si calcola ora la probabilità A di avere una sola stazione in trasmissione in un certo time slot.

$$A = \binom{N}{1} p^1 (1-p)^{(N-1)} = Np(1-p)^{(N-1)}$$

che assume il massimo per $p = 1/N$:

$$A = \left(1 - \frac{1}{N}\right)^{(N-1)}$$

Si determina ora la lunghezza media in slot dell'intervallo di contesa w :

$$E[w] = \sum_{i=1}^{\infty} i \cdot Pr\{i \text{ slot con coll.} - o - no - tx\} =$$

$$E[w] = \sum_{i=1}^{\infty} i(1-A)^i A = \frac{1-A}{A}$$

L'utilizzazione massima sarà quindi:

$$S = \frac{1/2a}{1/2a + \frac{1-A}{A}} = \frac{1}{1 + 2a\frac{1-A}{A}}$$

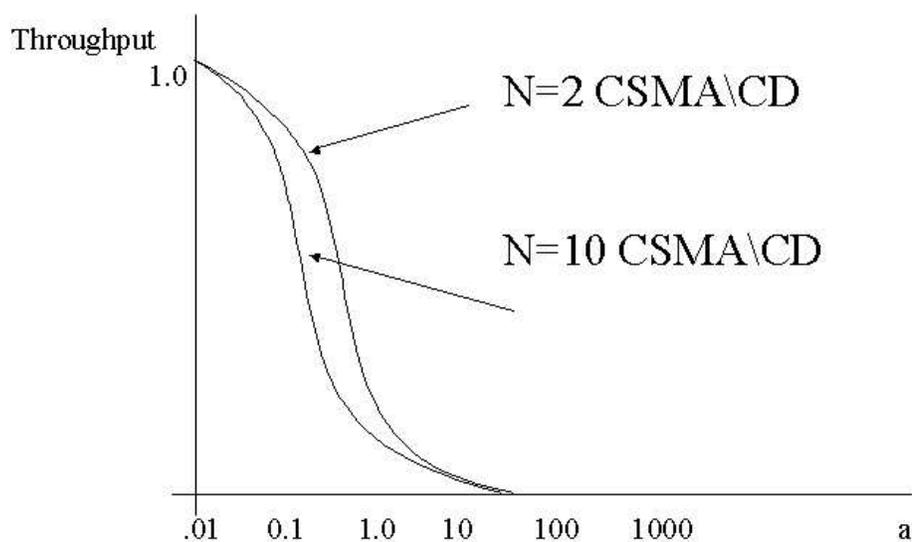


Figure 4: Throughput in funzione di a .

Protocolli di accesso senza contesa

Nei protocolli di **accesso senza contesa** viene definito un meccanismo di assegnazione del diritto a trasmettere che evita il fenomeno della collisione.

Ad esempio, nei sistemi radiomobili, le stazioni base assegnano i canali in modo deterministico agli utenti, quindi, una volta che a un utente è stato permesso di trasmettere su un certo canale, esso sarà l'unico ad avere il diritto di utilizzarlo. Questo ha come conseguenza che al ricevitore è possibile separare i segnali ricevuti su diversi canali. Se il numero di canali richiesti è eccessivo allora il servizio deve essere negato (situazione di blocco) a qualche terminale.

Le principali tecniche di accesso di questo tipo sono tre:

1. basata su *token*
2. FDMA (Frequency Division Multiple Access)
3. TDMA (Time Division Multiple Access)
4. CDMA (Code Division Multiple Access)

Token ring

Rete locale con topologia ad anello. L'accesso multiplo avviene mediante la gestione del *token*, pacchetto di controllo che transita in continuazione sull'anello per assegnare il diritto a trasmettere.

Una stazione deve impossessarsi del token per trasmettere, allorché può inviare uno o più pacchetti in funzione della loro lunghezza e del THT (Timer Holding Token).

Si assuma ancora unitario il tempo di trasmissione del frame ed il tempo di propagazione sia uguale ad a . Vi sono N stazioni attive.

Si ricava che il throughput S:

$$S = \frac{1}{1 + a/N}, \quad a < 1$$

$$S = \frac{1}{a(1 + 1/N)}, \quad a > 1$$

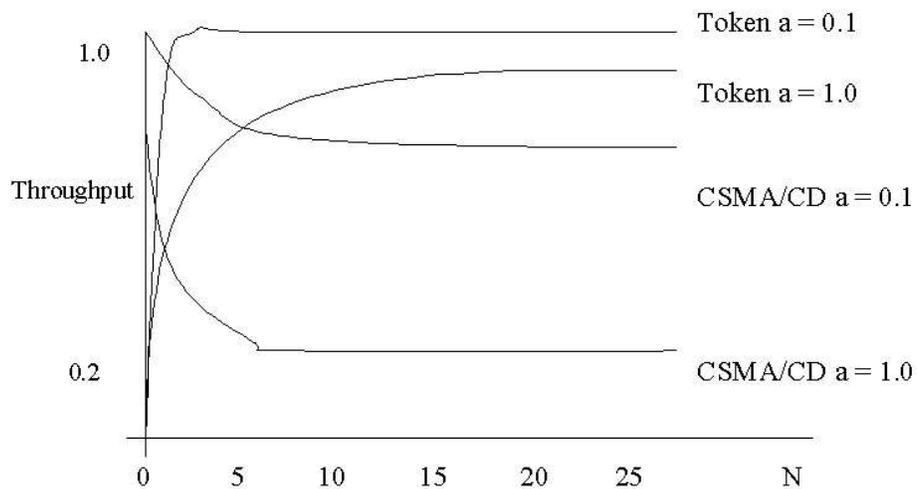


Figure 5: Throughput comparato CSMA/CD Token ring.

FDMA

(Frequency Division Multiple Access)

La tecnica di accesso multiplo a divisione di frequenza FDMA è stata una delle prime utilizzate, in particolare, nei sistemi radiomobili di prima generazione. La banda a disposizione del sistema BW viene suddivisa in un numero, diciamo n , di sottobande non sovrapposte in frequenza di larghezza BW/n (parametro che viene detto *passo di canalizzazione*) ognuna delle quali rappresenta un canale di servizio per gli utenti. Per realizzare la bidirezionalità della comunicazione con la tecnica FDD (Frequency Division Duplex) un gruppo delle n sottobande contigue è destinato per la comunicazione in salita, dalla stazione mobile (MS) alla stazione base (BS), mentre un analogo gruppo per il collegamento in discesa BS-MS. I due gruppi ovviamente non si intersecano sull'asse delle frequenze come si vede in Figura 6.

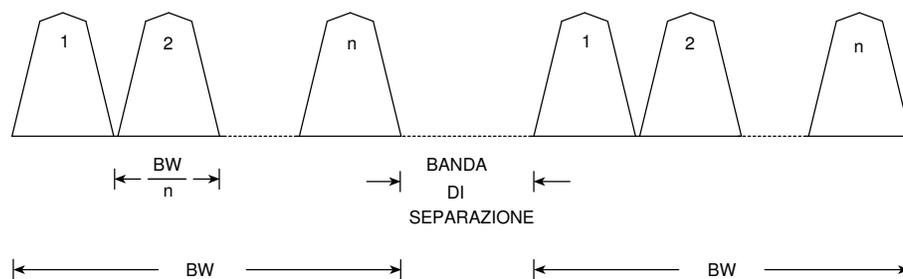


Figure 6: Divisione della banda con la tecnica FDMA/FDD.

Ad ogni stazione mobile viene assegnata una sottobanda in entrambe le bande FDD, che potrà utilizzare per tutto il tempo della durata della comunicazione. Il problema principale di questa tecnica è che bisogna cercare di garantire una sufficiente separazione dei canali, ossia evitare fenomeni di *aliasing*, fenomeni di sovrapposizione di sottobande adiacenti nel dominio delle frequenze. Il fenomeno di aliasing comporta la generazione di interferenza fra canali adiacenti, detta *interferenza intercanale*.

Il numero n di canali che è possibile ottenere dalla banda BW è limitato; infatti se B_{utente} è la banda del segnale modulato da trasmettere, allora il passo di canalizzazione non solo non deve essere inferiore a tale valore, ma deve anche essere sovradimensionato poiché altrimenti si verificherebbe appunto l'aliasing. L'efficienza di utilizzo della banda decresce all'aumentare del numero n di sottobande a causa di questa separazione diversa da zero fra canali adiacenti.

L'FDMA può essere impiegato anche per trasmissioni TDD (Time Division Duplex). La comunicazione è basata sul concetto di frame temporale: ogni singolo canale radio diventa bidirezionale, nella prima metà del frame si trasmette in una direzione (es. dal MS al BS) mentre nell'altra metà, nella direzione opposta (es. dal BS al MS). Un vantaggio della tecnica TDD è che, a differenza della FDD, non richiede l'uso esclusivo di due bande, difficili in genere da trovare nello spettro.

TDMA (Time Division Multiple Access)

La tecnica di accesso multiplo a divisione di tempo TDMA prevede che ogni utente acceda a tutta la banda BW del sistema per un intervallo di tempo determinato e, in generale, su base periodica. Nelle tecnica TDMA è quindi sufficiente la trasmissione di una sola portante, dove ogni utente guadagna l'accesso ad essa (può modularla) per un tempo pari a $1/N$ del tempo totale. La trasmissione avviene in intervalli di tempo di durata fissa $DS = T$, detti *slot*, e di periodo $DF = T_f$, frame o trame, come si vede in Figura 7.

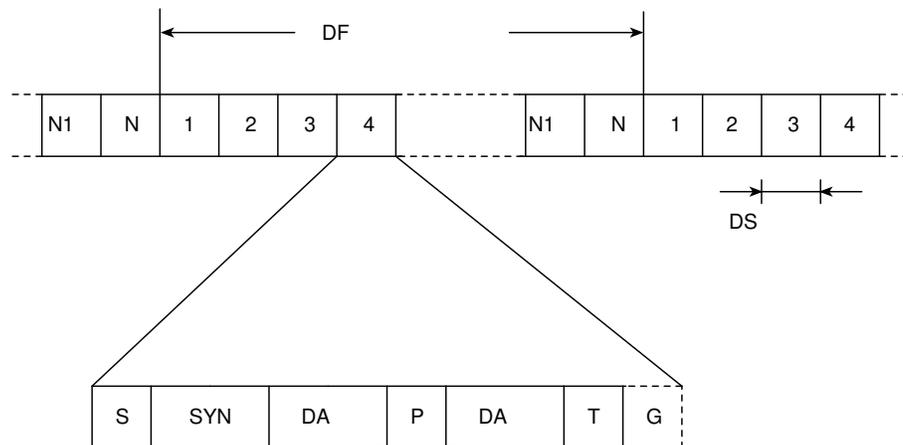


Figure 7: Struttura del frame nella tecnica TDMA.

Un sistema di questo tipo ha una struttura a *frames* di durata fissa di T_f secondi. Un canale elementare è definito come uno slot che si ripete in modo periodico all'interno dei frames. Definito il numero N di slot per frame è determinato anche il numero massimo N di utenti nel sistema. I dati trasmessi nel singolo time-slot sono detti *burst*. Il burst è composto da un insieme di bit, fra cui il bit di start (S) e di end (T) e spesso l'intestazione (*header*) comprende anche bit di sincronismo (SYN). Una sequenza di bit di controllo (P) può essere inserita qualora si voglia stimare la qualità del collegamento. Per facilitare la sincronizzazione viene inserita fra i pacchetti una sequenza di bit di guardia (G); in questo modo è possibile anche rimediare ai diversi ritardi di propagazione tra le MS lontane dalle BS rispetto quelle più vicine. Non esiste il problema dell'aliasing come nella FDMA, in quanto si trasmette su una unica portante, è però necessaria una ottima sincronizzazione per rispettare la struttura a slot temporali. Nel TDMA/FDD sono implementate due bande di frequenza, una per il collegamento in discesa, l'altra per il collegamento in salita. La comunicazione verso la stazione base è semplificata: alla BS è sufficiente un solo ricevitore per servire i propri utenti dal momento che questi trasmettono uno alla volta in istanti ben distinti determinati dalla posizione dello slot nel frame che stanno utilizzando.

La tecnica TDMA è molto flessibile e si adatta bene alle variazioni di traffico: se un utente vuole trasmettere una quantità di informazione superiore alla minima permessa con uno slot, la sua richiesta di servizio può essere facilmente soddisfatta, a parità di bit rate, aumentando l'intervallo di tempo in cui accede al canale, quindi assegnandoli un numero adeguato di slot per frame. È opportuno a questo punto la seguente puntualizzazione: la bit rate B_r di trasmissione è uguale per tutti i trasmettitori ed è determinata dalla banda BW del canale radio di cui si dispone. Assegnare più slot del frame ad un singolo utente non modifica la bit rate di trasmissione, che resta sempre la stessa, ma modifica la quantità di informazione (in bit) trasmessa nel tempo T_f durata del frame. Si tratta cioè di una sorta di “*bit rate media*” (la comunicazione ha una durata pari ad un numero intero di frames).

FDMA vs. TDMA: equivalenza di bit rate

Si supponga che entrambi i sistemi supportino un totale di R bit/s.

In un sistema FDMA la banda totale viene divisa in M bande di frequenza disgiunte, in modo che ognuna possa recare R/M bit/s.

In un sistema TDMA il frame è dato da M distinti time slots. Ogni time slot quindi occupa $1/M$ del tempo e consente la trasmissione a R bit/s.

In entrambi i casi una qualunque delle M sorgenti trasmette alla velocità media di R/M bit/s.

Si supponga inoltre che l'informazione trasmessa sia a pacchetto con dimensione b bit.

In FDMA, pacchetto di b bit sono trasmessi in T su ognuno dei diversi canali. Quindi la bit rate totale richiesta R_{FD} vale:

$$R_{FD} = M \frac{b}{T}$$

In TDMA, i b bit sono trasmessi in T/M s da ogni sorgente. Quindi la R_{TD} richiesta sarà:

$$R_{TD} = \frac{b}{T/M}$$

Siccome queste espressioni devono fornire lo stesso valore, si ha che:

$$R_{FD} = R_{TD} = R = \frac{Mb}{T}$$

da cui si deduce che entrambi i sistemi richiedono la stessa velocità di trasmissione di R bit/s.

FDMA vs. TDMA: ritardo medio

Si supponga che i canali di comunicazione, bande o time slot, siano sempre in uso. Il ritardo di un pacchetto D è dato da due contributi: $D = w + \tau$, dove w è il tempo di attesa medio prima della trasmissione e τ è il tempo di trasmissione del pacchetto.

In FDMA ogni pacchetto viene inviato in un tempo pari a T secondi, da cui $\tau_{FD} = T$.

In TDMA, ogni pacchetto viene inviato in slot di T/M s. In questo caso $\tau_{TD} = T/M = b/R$.

In FDMA, $w_{FD} = 0$.

In TDMA, un utente occupa un generico time slot nella trama di T s. Perciò il pacchetto S_{mk} , k -esimo pacchetto della sorgente m comincerà ad essere trasmesso $(m - 1)T/M$ s dopo la sua generazione. Quindi:

$$\begin{aligned} w_{TD} &= \frac{1}{M} \sum_{m=1}^M (m - 1) \frac{T}{M} = \\ &= \frac{T}{M^2} \sum_{n=0}^{M-1} n = \frac{T}{M^2} \frac{(M - 1)M}{2} = \\ &= \frac{T}{2} \left(1 - \frac{1}{M}\right) \end{aligned}$$

Riepilogando:

$$D_{FD} = T$$

$$D_{TD} = \frac{T}{2} \left(1 - \frac{1}{M}\right) + \frac{T}{M} = D_{FD} - \frac{T}{2} \left(1 - \frac{1}{M}\right)$$

che può anche scriversi:

$$D_{TD} = D_{FD} - \frac{b}{2R}(M - 1)$$

Questo dimostra che TDMA è strutturalmente superiore a FDMA dal punto di vista del ritardo.

CDMA (Code Division Multiple Access)

La tecnica di accesso multiplo a divisione di codice CDMA o SSMA (Spread Spectrum Multiple Access) è caratterizzata dal fatto che ogni trasmettitore codifica l'informazione in modo diverso dagli altri. Si usano speciali tecniche di codifica e decodifica per far sì che in ricezione si riesca a ottenere l'informazione desiderata. Le stazioni mobili trasmettono alla stessa frequenza, usando tutta la banda disponibile come nel TDMA, senza limitazioni di tempo come nel FDMA. Esistono due tipologie di SSMA: il direct sequence SSMA (DS/SSMA), detto più comunemente DS-CDMA (Direct Sequence Code Division Multiple Access) e il Frequency Hopping DS-CDMA (FH/SSMA). Le tecniche *spread spectrum*, letteralmente "spettro allargato", sono nate in ambiente militare alla fine degli anni '40 e successivamente introdotte anche in ambiente civile. L'idea di base è quella di trasformare, moltiplicandolo per una parola di codice $c(t)$, un segnale di larghezza di banda B_s , in un segnale "like-noise" avente larghezza di banda B_{ss} molto maggiore di quella originale, ma uguale potenza (quindi cala la densità spettrale media di potenza). In demodulazione, il segnale viene moltiplicato (*despreading*) di nuovo per la stessa parola di codice $c(t)$ e l'effetto è quello di ottenere al ricevitore il segnale precedente all'operazione di spreading (come si vede in Figura 8), sommato agli altri segnali che restano a larga banda.

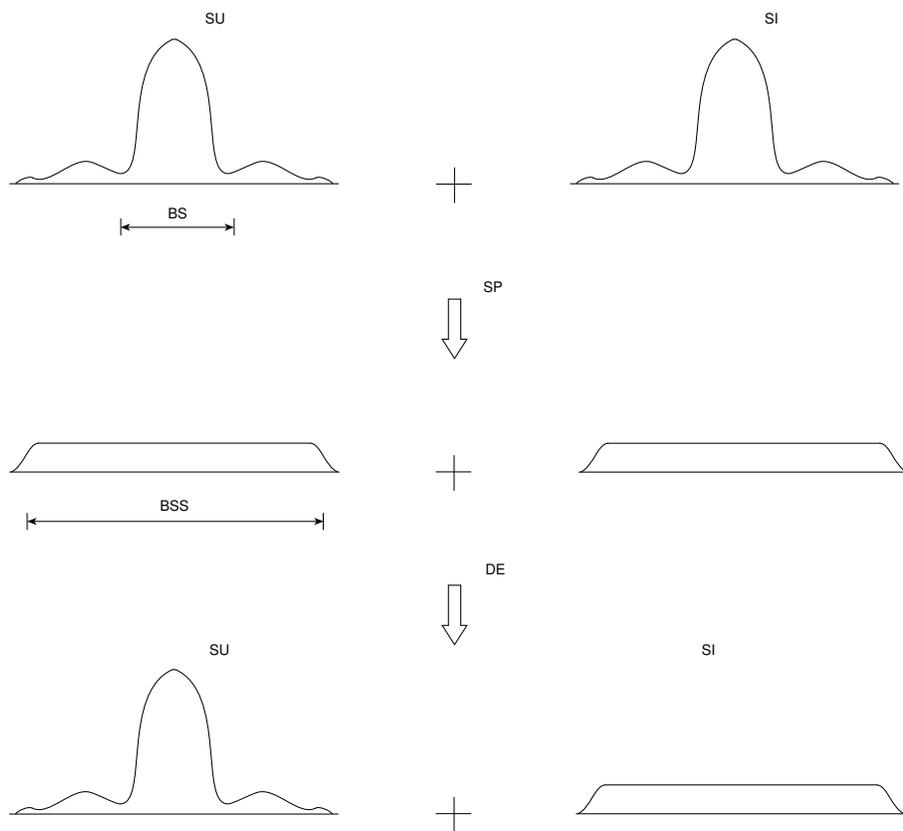


Figure 8: Principio di funzionamento della tecnica spread spectrum.

Con la tecnica CDMA tutta la banda disponibile è utilizzata in ogni parte della rete. Il numero di utenti che possono essere serviti simultaneamente dipende dal numero di parole di codice. Tale numero è limitato dai livelli di interferenza che si verificano in ricezione. L'interferenza infatti è molto critica, soprattutto a causa di trasmettitori simultanei all'interno della medesima cella (*interferenza intracella*). Per superare questo problema si impiegano controlli di potenza che tentano di far pervenire alla stazione base, segnali a livelli di potenza costanti indipendentemente dalla posizione del mobile nella cella. Sono facilitate le operazioni di hand-off, infatti, nel passaggio da una cella ad un'altra, è sufficiente mantenere la stessa parola di codice se questa non è già utilizzata nella cella di arrivo, oppure assegnarne semplicemente una nuova.

Per aver un basso contributo di interferenza in ricezione, la larghezza della banda B_{ss} deve essere molto grande, quindi, a parità di numero di canali, la banda richiesta dalla tecnica CDMA è molto superiore rispetto alle tecniche FDMA e TDMA. Una caratteristica interessante delle tecniche SSMA è che possono tranquillamente condividere la stessa banda con i sistemi di trasmissione a banda stretta senza determinare interferenze tali da limitare gli uni o gli altri sistemi.

Questa tecnica è alla base dei più evoluti sistemi radiomobili come quelli UMTS.