

UNIVERSITÀ DEGLI STUDI DI PADOVA

FACOLTÀ DI INGEGNERIA

DIPARTIMENTO DI ELETTRONICA ED INFORMATICA

TESI DI LAUREA

**PIANIFICAZIONE FREQUENZIALE PER RETI
RADIOMOBILI DI GRANDI DIMENSIONI**

RELATORE: PROF. MATTEO FISCHETTI

LAUREANDO: DAVID CIPRIAN

ANNO ACCADEMICO 1998-1999

A Lorena

SOMMARIO

Il presente lavoro di tesi è rivolto alla ricerca, analisi e sviluppo di nuove tecniche euristiche per il problema di Assegnazione Frequenziale per la Rete Telefonica Radiomobile. In questo problema, il territorio di copertura viene suddiviso in zone geografiche (denominate celle), servite da antenne di trasmissione, ciascuna equipaggiata con il numero necessario di canali. Essendo limitato il numero di frequenze disponibili, il riuso è causa di interferenza fra celle interagenti; inoltre celle adiacenti non possono utilizzare la stessa frequenza. Ci si propone, quindi, di trovare un “Piano Frequenziale” che comporti il minimo livello di interferenza complessivo, nel rispetto di vincoli assegnati. L’obiettivo di questa tesi è lo studio ed implementazione di nuovi algoritmi parametrici (basati su opportune strutture dati) che consentano un tempo di elaborazione ridotto ed una buona qualità delle soluzioni anche per reti di grandi dimensioni. I risultati computazionali mostrano l’efficacia delle tecniche implementate.

INTRODUZIONE

La comunicazione mobile è diventata una componente essenziale dell'economia, nel nostro come negli altri paesi europei. In breve tempo il numero dei terminali mobili ha superato quello dei fissi, e continua a crescere. La crescita della richiesta di servizi ha comportato un corrispondente ampliamento della rete di telecomunicazione, sia in termini di copertura geografica che come numero di comunicazioni gestite.

Il territorio viene diviso in zone di servizio denominate celle, ciascuna servita da una antenna trasmissiva tramite canali caratterizzati da una particolare frequenza: il numero di frequenze possibili, spaziate fra loro di 200kHz, è limitato a qualche decina per ogni gestore, il che comporta che sia necessario il riuso su celle distinte, che quindi interagiscono fra loro sotto forma di disturbo sul segnale utile. Ogni cella non può utilizzare frequenze che distino fra loro meno di 600kHz, o che distino meno di 400kHz da quelle assegnate a celle servite dalla stessa antenna (sito), o che corrispondano a quelle assegnate a celle adiacenti. Il problema di Assegnazione delle Frequenze consiste nel determinare il Piano di Frequenza a minima interferenza complessiva, nel rispetto di tutti i vincoli, garantendo un livello minimo di interferenza a tutte le portanti del piano.

Storicamente, dopo che è stato riconosciuto come problema NP-Completo (Metzger, 1970) e modellato come problema di ottimizzazione lineare intera, è stato affrontato con approcci di tipo esatto (Fischetti et al., 1996), euristici (Eisenblätter et al., 1998) o enumerativi (Mannino e Sassano, 1998). All'aumentare della taglia dell'istanza d'ingresso, solo le tecniche euristiche si sono dimostrate adatte allo scopo, soprattutto con l'introduzione della tecnica della Tabu Search a vincoli rilassati (Massignan, 1998); spesso, tuttavia, i tempi di esecuzione rimanevano proibitivi.

Scopo della presente tesi è lo sviluppo di tecniche euristiche che permettano l'analisi di reti di grandi dimensioni, e che nel contempo permettano tempi di esecuzione compatibili con le moderne esigenze industriali.

Nel primo capitolo si introducono i sistemi radiomobili, e si descrive a sommi capi la struttura del Sistema di Trasmissione e le caratteristiche del Sistema GSM. Il secondo capitolo si occupa della definizione del problema di Assegnazione delle Frequenze, e dei vari modelli matematici. Nel terzo capitolo verranno presentate le varie tecniche e gli approcci possibili, mentre la tecnica della Tabu Search, che occupa una parte fondamentale nella ricerca della soluzione ottima, verrà trattata in dettaglio nel quarto capitolo. Gli strumenti sviluppati verranno descritti diffusamente nel quinto capitolo, dove si confronteranno i risultati ottenuti con quelli presenti in letteratura, e verranno date alcune indicazioni sui possibili sviluppi futuri.

INDICE

SOMMARIO	I
INTRODUZIONE	III
INDICE	IV
CAPITOLO I	1
1.1 La comunicazione mobile	1
1.2 I sistemi cellulari analogici	2
1.3 Il sistema GSM	3
1.3.1 Architetture e interfacce del sistema	5
1.3.2 Mobile Station (MS)	7
1.3.3 Base Station Subsystem (BSS)	7
1.3.4 Network Switching Subsystem (NSS)	8
1.3.5 Operation and Support Subsystem (OSS)	10
1.4 Ripartizione attuale delle frequenze GSM in Italia	11
1.5 Storico delle assegnazioni	13
1.6 Evoluzione futura: il GPRS	16
1.6.1 Descrizione del sistema	17
1.6.2 Architettura di una rete GPRS	21
1.6.3 Caratteristiche del GPRS in breve	23
1.7 Inquinamento Elettromagnetico	24
CAPITOLO II: La pianificazione Frequenziale	27
2.1 Introduzione	27
2.2 Descrizione del problema	28
2.3 Vincoli del problema	30

2.3.1	Vincoli di banda utile	31
2.3.2	L'equipaggiamento delle celle	32
2.3.3	I vincoli di riuso	33
2.3.4	Vincoli interferenziali	36
2.3.5	Vincoli di frazionamento	38
2.4	Formulazione del problema	40
2.5	Complessità Computazionale	42
2.6	Rilassamento dei vincoli	43
2.7	Una Formulazione alternativa	44
 <u>CAPITOLO III: EURISTICHE NON TABU-SEARCH</u>		 46
3.1	Introduzione	46
3.2	T-coloring	47
3.3	Dual Greedy	48
3.4	DSATUR con costi	49
3.5	1-OPT iterato	51
3.6	Flusso a costo minimo (FCM)	53
3.7	Rafforzamento del vincolo di separazione	53
3.8	Un algoritmo enumerativo	54
 <u>CAPITOLO IV: LA TABU SEARCH</u>		 57
4.1	Introduzione	57
4.2	Caratteristiche generali	58
4.2.1	Criterio di aspirazione	62
4.2.2	Intensificazione e Diversificazione	62
4.3	Reactive Tabu Search	64
 <u>CAPITOLO V: DESCRIZIONE DELL'ALGORITMO</u>		 65
5.1	Generalità	65
5.2	Dati di Ingresso	67
5.3	Una nuova struttura dati di ingresso	71
5.4	Definizione del problema	73
5.5	Inizializzazione	74

5.6	Esplorazione	75
5.7	Idea di base e suo miglioramento	77
5.8	Effetto del rilassamento dei vincoli	81
5.9	L' algoritmo complessivo	83
5.10	Uso del Core	86
5.11	Vincoli di soglia	87
5.12	Analisi dei risultati	87
5.13	Dinamica dell'esplorazione	91
5.14	Possibili sviluppi futuri	93
CONCLUSIONI		94
APPENDICE A		95
BIBLIOGRAFIA		98

Capitolo I

SISTEMI RADIOMOBILI

1.1 La comunicazione mobile

Questi ultimi anni sono stati caratterizzati per la diffusione di massa della *mobile communication*, un insieme di servizi che permettono una connessione continuativa fra due o più utenti in una rete di telecomunicazione, anche in una situazione di mobilità relativa.

Storicamente si può far coincidere la nascita delle comunicazioni mobili con l'invenzione della modulazione di frequenza (FM, frequency modulation) avvenuta nel 1935 per merito di E. H. Armstrong. Notevole fu lo sviluppo di tali sistemi durante la II guerra mondiale, tanto che il numero di canali radio fu subito avvertito come risorsa scarsa, anche se inizialmente non si appoggiavano alla rete telefonica fissa. I primi sistemi di telefonia mobile furono introdotti alla fine degli anni quaranta, ed erano composti da un singolo trasmettitore FM commutato manualmente.

I sistemi non cellulari, effettuano trasmissioni di tipo *broadcast* (come radio e TV) utilizzando trasmettitori di potenza elevata per coprire una vasta area. Se il numero di utenti non è molto ridotto si ha così un enorme fabbisogno di frequenze radio, tale da impedirne l'effettiva realizzazione.

I sistemi cellulari che realizzano le reti radiomobili, invece, applicano la tecnica del riutilizzo delle frequenze: una frequenza (o canale), viene utilizzata più volte in luoghi diversi sufficientemente lontani tra loro. Il territorio viene suddiviso in aree di servizio, denominate celle, ognuna servita da una stazione radio base che trasmette su un certo insieme di canali radio, diversi da quelli utilizzati nelle celle adiacenti, per evitare interferenze.

Se durante gli spostamenti l'utente passa da una cella ad un'altra, è necessario che il terminale mobile si sintonizzi su una nuova frequenza, tipicamente quella ricevuta meglio tra le frequenze della nuova cella. Ciò è indispensabile durante una conversazione per evitare la caduta della comunicazione; la procedura con la quale si effettua il cambio di frequenza nel passare da una cella all'altra viene detta *handover*.

Nei sistemi cellulari, aumentando il numero delle celle che coprono una certa area e perciò riducendo la loro dimensione, aumenta la capacità del sistema (cioè il numero di utenti gestiti) ma diminuisce la distanza di riuso delle frequenze (cioè la distanza tra due celle che usano lo stesso canale) ed aumenta perciò l'interferenza tra canali che utilizzano la stessa frequenza (interferenza cocanale) ed aumenta il numero di handover che il sistema deve effettuare durante una conversazione. Perciò, la dimensione delle celle non può scendere al di sotto di certi valori e si ripresenta il problema del limitato numero di frequenze disponibili.

1.2 I sistemi cellulari analogici

I primi sistemi cellulari introdotti erano di tipo analogico, utilizzano la modulazione FM e avevano le seguenti limitazioni:

- Ad ogni utente che effettuasse una richiesta di connessione veniva assegnata una frequenza, che rimaneva impegnata per tutta la durata della conversazione non potendo così essere utilizzata da altri terminali.

- La capacità del sistema (numero di utenti) era limitata dal numero delle frequenze disponibili e dal limite imposto alle dimensioni delle celle a causa dell'interferenza co-canale.
- Non si potevano applicare direttamente algoritmi di crittografia se non utilizzando apparati specifici molto costosi.
- La sicurezza dell'accesso alla rete si basava solo sul riconoscimento di un numero di serie identificante il terminale mobile; erano quindi possibili clonazioni.
- Le trasmissioni dati erano praticamente inattuabili.

Una prima soluzione, sviluppata e introdotta negli USA, fu l'AMPS (*Advanced Mobile Phone Standard*) nel 1979. In Europa iniziò l'NMT (*Nordic Mobile Telephone*), in Svezia nel 1981 e subito dopo in Norvegia, Danimarca e Finlandia.

Successivamente si sviluppò, nel Regno Unito, lo standard TACS (*Total Access Communications System*), una versione modificata del sistema AMPS.

Le specifiche iniziali, che assegnavano al sistema 1000 canali centrati nella banda 890-960 MHz, sono state evolute successivamente nello standard ETACS (*Extended TACS*), che assegna 1320 canali nella banda 872-950 MHz.

1.3 Il sistema GSM

Nei primi anni '80 ogni paese europeo cercò di sviluppare il proprio sistema radiomobile, in modo però incompatibile con ogni altro sia in termini di software che di hardware: questo impose che ogni terminale fosse sì mobile, ma nel limite dei confini nazionali.

Nel 1982 la *Conférence Européenne des Postes et des Télécommunications* (CEPT) creò un gruppo di studio denominato *Groupe Spécial Mobile* (GSM) con lo scopo di studiare e sviluppare un sistema radiomobile cellulare paneuropeo, comune a tutti i paesi dell'Europa occidentale.

Il sistema proposto doveva essere incentrato su dei precisi criteri:

- elevata qualità audio della conversazione
- bassi costi
- roaming internazionale
- supporto di nuovi servizi
- compatibilità con il sistema digitale ISDN
- sicurezza e riservatezza nelle comunicazioni.

Sistemi basati su tecniche digitali presentano notevoli vantaggi rispetto ai corrispettivi analogici:

- sono meno sensibili a rumore ed interferenze e quindi consentono di ridurre le dimensioni delle celle, aumentando il numero di utenti che possono essere serviti contemporaneamente
- consentono un alto grado di riservatezza, in quanto le informazioni trasmesse sulla tratta radio possono essere cifrate direttamente dall'apparato utente
- permettono un elevato grado di sicurezza: l'identità dell'apparato che chiede l'accesso alla rete può essere controllata tramite l'applicazione di un opportuno algoritmo e di una chiave di autenticazione segreta,
- consente di effettuare trasmissioni dati (il segnale vocale stesso viene digitalizzato e poi trasmesso).

Due furono le bande di frequenza riservate: 890-915 e 935-960 MHz.

Nel 1989, la responsabilità del progetto GSM venne trasferita alla *European Telecommunication Standards Institute* (ETSI). In quella sede venne ridefinito l'acronimo GSM come *Global System for Mobile Communications*; commercializzato nel 1991 (la prima chiamata ufficiale venne effettuata l' 1/7/1991),

attivato in Italia dal 1994, il servizio è divenuto ormai uno standard non solo in Europa, ma anche in oltre 80 paesi del mondo.

Cinque sono gli obiettivi primari del GSM:

- utilizzare le frequenze radio a disposizione nel modo più efficiente possibile, tollerando le inevitabili interferenze co-canale
- qualità della fonia superiore a quella ottenibile con i sistemi analogici
- supporto della trasmissione dati sia a commutazione di circuito sia a commutazione di pacchetto
- compatibilità con la rete ISDN e con i servizi ad essa relativi
- elevato grado di sicurezza sia sui dati di abbonamento degli utenti sia delle informazioni trasmesse

In Italia il servizio ha riscosso un notevolissimo successo: a dieci anni dalla sua introduzione il numero dei terminali radiomobili attivi supera i trenta milioni, con una media di oltre un “telefonino” ogni due abitanti.

1.3.1 Architetture e interfacce del sistema

La rete si compone di quattro sottosistemi cooperanti, ciascuno comprendente un certo numero di unità funzionali comunicanti fra loro attraverso opportune interfacce. La struttura è modulare, con intelligenza distribuita, il che consente di realizzare reti adattabili alle diverse esigenze che possono presentarsi nelle varie Nazioni.

- La Stazione Mobile (*Mobile Station*) è il terminale mobile usato dall'abbonato per la richiesta di un servizio.
- La Stazione Base (*Base Station Subsystem*) gestisce la parte radio e comprende funzioni quali assegnazione delle frequenze, decisione di handover, ecc.

- Il Sottosistema di rete (*Network Switching Subsystem*), a cui fa capo il Centro di Commutazione (*Mobile services Switching Center*) realizza la connessione tra l'utente mobile e gli utenti delle altre reti, fisse o mobili.
- Sottosistema di esercizio e manutenzione (*Operation and Support Subsystem*) adibito al corretto funzionamento e controllo operativo della rete.

Ogni entità interagisce con le altre attraverso opportune interfacce, mentre il *roaming*, cioè il potersi spostare liberamente sul territorio servito dal proprio gestore, ed anche su quello servito dagli altri gestori delle nazioni che aderiscono al GSM, richiede di memorizzare in un database dinamico la posizione degli utenti. A tal scopo l'area geografica di servizio del sistema GSM è suddivisa gerarchicamente in diverse aree, dette *Network service areas*. Un operatore GSM è quindi sempre in grado di conoscere la posizione di ciascun suo abbonato.

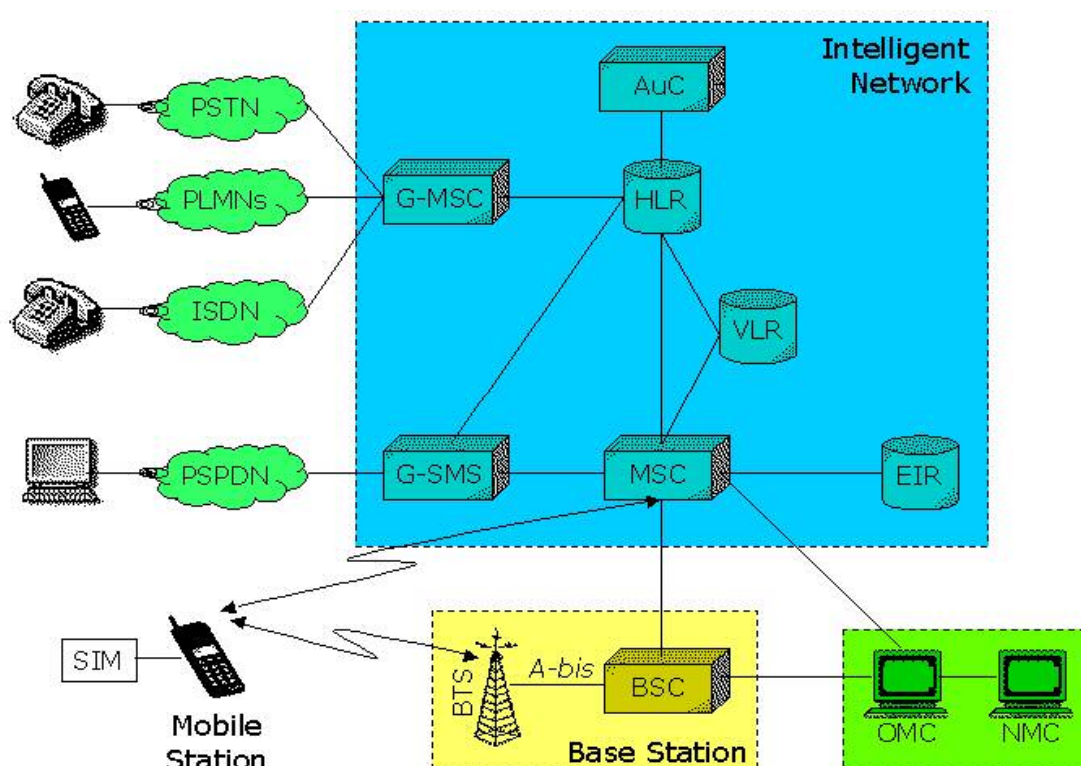


Figura 1.1 - Architettura del sistema GSM

1.3.2 Mobile Station (MS)

La *mobile station* (MS) rappresenta la stazione mobile con la quale un utente può usufruire dei servizi offerti dal GSM. È formata da due elementi:

- Un terminale mobile (*Mobile Equipment*, ME) comprendente tutto l'hardware e il software necessario, ed identificato univocamente da un numero di serie noto come IMEI (*International Mobile Equipment Identity*).
- Un modulo con chip dotato di microprocessore e memoria detto SIM (*Subscriber Identity Module*), contenente i dati di abbonamento e gli algoritmi e le chiavi di autenticazione e cifratura. Ogni SIM ha un proprio numero di serie e può essere protetto attraverso una password personale detta PIN (*Personal Identification Number*).

Tre sono i tipi definiti di Mobile Station:

- Veicolare: installata su un veicolo ha l'antenna posta all'esterno del veicolo stesso.
- Trasportabile: può essere trasportabile a mano ma l'antenna non è fisicamente inserita nell'apparato che costituisce la *Mobile Termination*.
- Portatile: è il classico "telefonino".



Figura 1.2 - SIM Card

1.3.3 Base Station Subsystem (BSS)

Il sottosistema BSS (*Base Station Subsystem*) si occupa della parte radio del sistema e di conseguenza comprende le unità funzionali che consentono di fornire la copertura radio di un'area costituita da una o più celle.

La stazione base é composta di due unità:

- La *Base Transceiver Station* (BTS), che ha il compito di realizzare i canali di comunicazione, di traffico e di segnalazione verso le MS e quindi di svolgere tutte le funzioni trasmissive e di elaborazione specificate per la Air Interface (interfaccia radio). In termini funzionali una BTS fornisce la copertura radio ad una cella.
- Una *Base Station Controller* (BSC), che controlla la parte radio della rete e svolge le funzioni correlate con il controllo del BSS e la gestione delle connessioni verso il sottosistema NSS

L'interfaccia di comunicazione tra le due entità, detta A-bis, è standardizzata. In questo modo non si è vincolati a soluzioni proprietarie e si possono utilizzare componenti prodotti da fornitori diversi.

Dal punto di vista logico appartiene al BSS anche il Trans Coder (TC), avente funzioni di transcodifica, il quale è normalmente associato ad un altro elemento noto come *submultiplexer* (SMUX)

1.3.4 Network Switching Subsystem (NSS)

Il sottosistema di rete, identificato a volte anche come *Intelligent Network* (IN), assolve ai seguenti compiti:

- Gestione del controllo delle chiamate (*Call Control*) e delle interconnessioni con le altre reti
- Gestione della mobilità degli utenti (*Mobility Management*)
- Supporto ai servizi supplementari (*Supplementary Service Support*)
- Gestione dei dati di utente relativi agli abbonamenti sottoscritti (*Subscriber Data Handling*)
- Gestione della sicurezza (*Security Management*); per l'autenticazione e la cifratura

Si compone delle seguenti unità funzionali:

- MSC (*Mobile Services Switching Center*), specializzato in
 1. Controllo delle chiamate (*Call Control*)
 2. Interworking con le altre reti (non GSM, PSTN, ISDN, ecc)
 3. Funzioni di Gateway per permettere il collegamento di un utente di rete fissa ad uno di rete mobile, tramite query all'HLR
 4. Tassazione delle chiamate (*Charging*)
 5. Controllo del sottosistema BSS
 6. Scambio della segnalazione verso il BSS e la rete telefonica.
- VLR (*Visitor Location Register*), essenzialmente un database dinamico che raccoglie tutte le informazioni necessarie per le chiamate da e verso le MS degli utenti che in un certo momento si trovano nella relativa area controllata. In particolare, ogni MS, entrando in una zona di pertinenza di un certo VLR, si registra inviando un segnale di Location Updating; dal canto suo, il VLR richiede i vari dati utente direttamente all'HLR centrale e li conserva per un periodo opportuno. Le funzioni del VLR sono:
 1. Gestione della mobilità (*Mobility Management*), comprendente funzioni di First Registration, Location Updating, Paging (chiamata), Handover tra BTS controllate da BSC diversi.
 2. Gestione della sicurezza (*Security Management*), tra cui autenticazione (Authentication), identificazione (Mobile Equipment Identification), cifratura (Ciphering), riallocazione dell'identità temporanea (*Temporary Identity Reallocation*)
 3. Trattamento dei dati di utente e dei servizi sottoscritti
- HLR (*Home Location Register*), registro che permette la localizzazione degli utenti di un gestore, detti home subscriber. Ad esso fanno riferimento il VLR per la richiesta di copia dei dati utente, ed il MSC per le informazioni necessarie per instradare una chiamata verso un MS. I principali dati in esso contenuti sono:
 1. numero di telefono della MS
 2. Identità dell'utente mobile (IMSI)
 3. parametri relativi alla gestione della sicurezza (autenticazione e cifratura)
 4. identificativo del VLR presso cui l'utente è registrato

5. servizi base e supplementari forniti

- AuC (*Authentication Center*), che gestisce la sicurezza mediante cifratura di chiavi.
- EIR (*Equipment Identity Register*), che permette di suddividere gli IMEI secondo tre liste: bianca (*White List*) per Mobile Equipment omologati e perfettamente funzionanti, grigia (*Grey List*) per i non omologati o difettosi, nera (*Black List*) per quelli rubati o che non possono accedere al sistema.

1.3.5 Operation and Support Subsystem (OSS)

Questo sottosistema si occupa di gestire, amministrare, mantenere e coordinare le varie attività. I suoi compiti sono:

- Gestione e manutenzione dei componenti della rete di telecomunicazione;
- Gestione degli abbonamenti e della “tariffazione” delle chiamate;
- Gestione e manutenzione degli apparecchi mobili e delle SIM.

Al suo interno si distingue:

- *Operation and Maintenance Center* (OMC), che controlla il corretto funzionamento di una parte di rete costituita da uno o più MSC, con i BSC e i BTS ad essi associati. L'OMC svolge le seguenti mansioni:
 1. gestione delle configurazioni e delle prestazioni di tutti gli elementi che compongono il network GSM (BSC, MSC, VLR, HLR, EIR ed AuC);
 2. gestione dei guasti, degli allarmi e dello stato del sistema con possibilità di effettuare vari tipi di test per analizzare le prestazioni e per verificare il corretto funzionamento dello stesso;
 3. gestione della sicurezza;
 4. raccolta di tutti i dati relativi al traffico degli abbonati necessari per la fatturazione
- *Network Management Center* (NMC), che fornisce la visibilità globale di tutte le attività di controllo. Coordina e gestisce tutti gli OMC presenti nel network.

1.4 Ripartizione attuale delle frequenze GSM in Italia

- Frequenze 900MHz

L'assegnazione, operativa dal 15 Settembre 1999, delle frequenze Extended GSM (880-915 MHz, 925-960 MHz) in Italia è la seguente.

- Al gestore **Tim** sono assegnati 11,4 MHz (57 portanti) con l'obbligo di utilizzare le portanti da 51 a 64 solo per il territorio all'interno dell'area urbana delle grandi città.
- Al gestore **Omnitel** sono assegnati 10.6 MHz (53 portanti), in virtù del suo minore carico d'utenza, con l'obbligo di utilizzare le portanti da 66 a 76 solo per il territorio all'interno dell'area urbana delle grandi città.
- Al gestore **Wind** sono assegnati 4.8 MHz (24 canali) solo per il territorio all'esterno dell'area urbana delle grandi città.

GSM TIM	<i>uplink</i>	891.7 - 900.1	(8.6 (11.4*) MHz canali da 8 a 50 (64*))
	<i>downlink</i>	933.7 - 945.1	(948.1*) MHz
GSM WIND	<i>uplink</i>	900.3 - 905.1 MHz	(4.8 MHz canali da 52 a 75)
	<i>downlink</i>	945.3 - 950.1 MHz	
GSM OMNITEL	<i>uplink</i>	905.3 (903.3*) - 913.7 MHz	(8.4 (10.6) MHz canali da 77 (66*) a 118)
	<i>downlink</i>	950.3 (948.3*) - 958.7 MHz	

* Le frequenze sono assegnate a Wind solo al di fuori delle 16 più grandi città.

Le tratte *Downlink* e *Uplink* si riferiscono rispettivamente ai canali usati per trasmettere dall'antenna al radiomobile e dal radiomobile all'antenna.

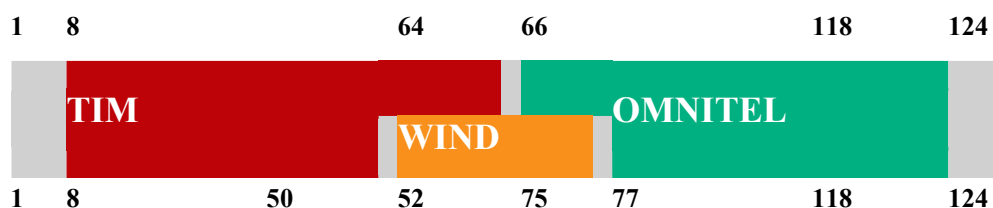


Figura 1.3 – Assegnazione delle frequenze a 900MHz in Italia

• Frequenze 1800MHz

L'assegnazione, operativa dal 1 Luglio 1999, delle frequenze DCS (1710-1785 MHz, 1805-1880 MHz) in Italia è la seguente.

- Ai gestori **Tim** e **Omnitel** sono assegnati 4.8 MHz ciascuno.
- Ai gestori **Wind** e **Blutel** sono assegnati 10 MHz ciascuno.

TIM *uplink* **1755 – 1759.8 MHz** (4.8 MHz)

downlink **1850 – 1854.8 MHz**

WIND *uplink* **1760 – 1770 MHz** (10 MHz)

downlink **1855 – 1865 MHz**

BLUTEL *uplink* **1770 – 1780 MHz** (10 MHz)

downlink **1865 – 1875 MHz**

OMNITEL *uplink* **1780.2 - 1785 MHz** (4.8 MHz)

downlink **1875.2 - 1880 MHz**

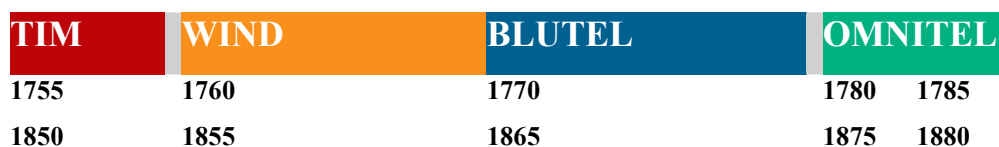


Figura 1.4 – Assegnazione delle frequenze a 1800MHz in Italia

1.5 Storico delle assegnazioni

Ecco alcuni piani di allocazione delle frequenze Extended GSM (880-915 MHz, 925-960 MHz) operativi in Italia nei mesi scorsi.

Luglio 1999

Piano frequenze operativo dal 1 Luglio 1999:

GSM TIM	<i>uplink</i>	891.7 - 900.1 (903.1*) MHz	(8.4 (11.4*) MHz canali da 9 a 50 (65*))
	<i>downlink</i>	933.7 - 945.1 (948.1*) MHz	
GSM WIND	<i>uplink</i>	900.3 - 905.1 MHz	(4.8 MHz canali da 52 a 75)
	<i>downlink</i>	945.3 - 950.1 MHz	
GSM OMNITEL	<i>uplink</i>	905.3 (903.3*) - 913.7 MHz	(8.4 (10.4) MHz canali da 77 (67*) a 118)
	<i>downlink</i>	950.3 (948.3*) - 958.7 MHz	

* Le frequenze sono assegnate a Wind solo al di fuori delle 16 più grandi città.

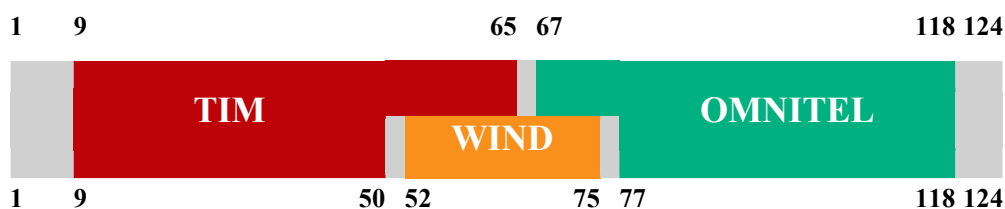


Figura 1.3

Novembre 1998

Piano frequenze operativo dal 1 Novembre 1998:

GSM TIM	<i>uplink</i>	892.3 - 900.5 (903.5*) MHz	(8.2 (11.2*) MHz canali da 12 a 50 (67*))
	<i>downlink</i>	937.1 - 945.5 (948.5*) MHz	
GSM WIND	<i>uplink</i>	900.7 - 905.3 MHz	(4.6 MHz canali da 54 a 76)
	<i>downlink</i>	945.7 - 950.3 MHz	
GSM OMNITEL	<i>uplink</i>	905.5 (903.7*) - 913.7 MHz	(8.2 (10) MHz canali da 78 (69*) a 118)
	<i>downlink</i>	950.5 (948.7*) - 958.7 MHz	

- Le frequenze sono assegnate a Wind solo al di fuori delle 16 più grandi città.

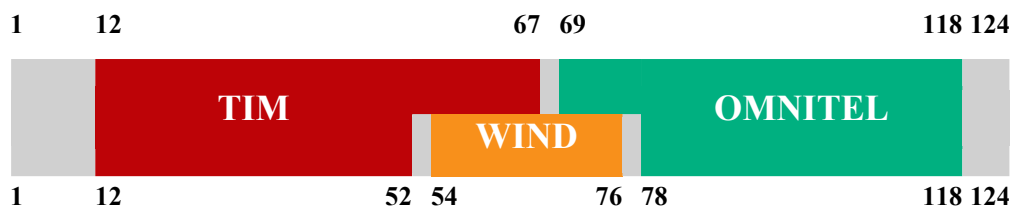


Figura 1.5

Maggio 1998

Piano frequenze operativo da **Maggio 1998**:

GSM TIM	<i>uplink</i>	894.5 - 904.5 MHz	(10 MHz canali da 23 a 72)
	<i>downlink</i>	939.5 - 949.5 MHz	
GSM OMNITEL	<i>uplink</i>	904.7 - 913.7 MHz	(9 MHz canali da 74 a 118)

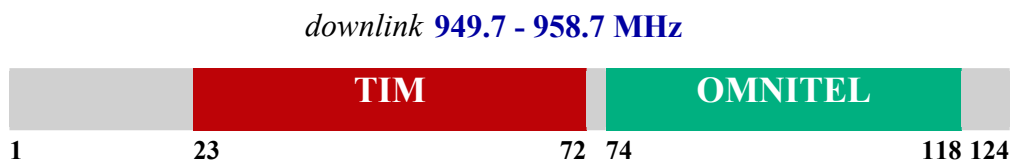


Figura 1.6

Le portanti da 64 a 72 sono assegnate a Tim solo per l'uso all'interno dell'area urbana delle grandi città.

Settembre 1997

Piano frequenze operativo dal 16 Settembre 1997:

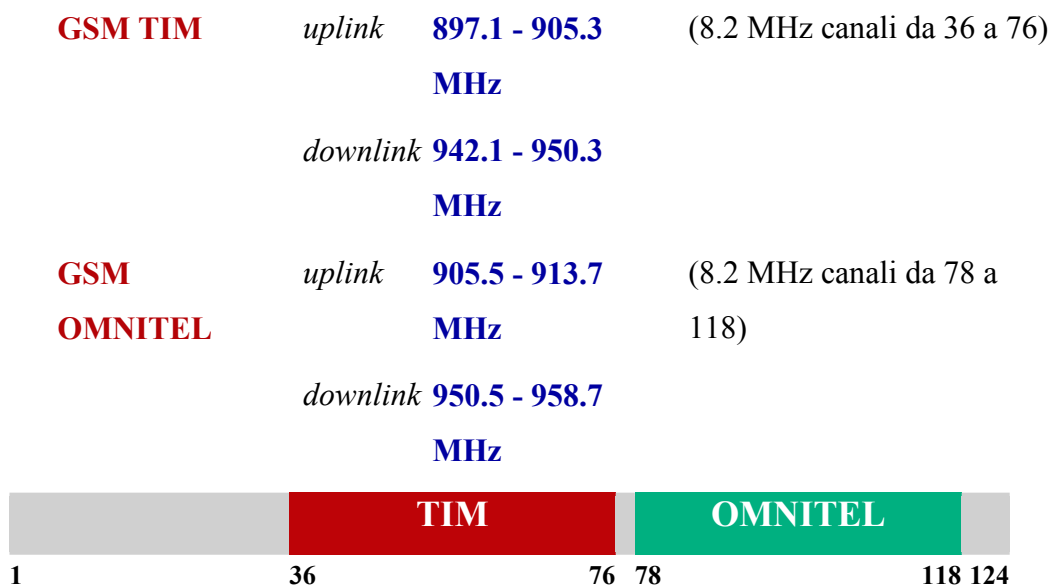
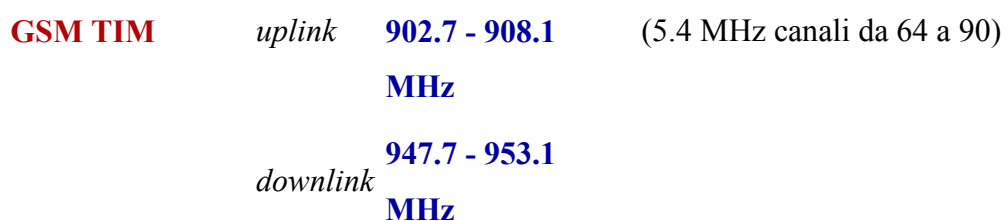


Figura 1.7

Dicembre 1994

Piano frequenze assegnato al momento della concessione ad Omnitel (Dicembre 1994):



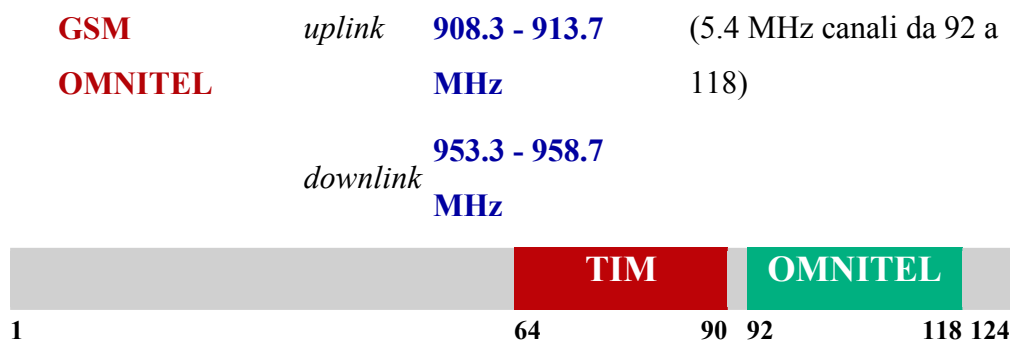


Figura 1.8

1.6 Evoluzione futura: il GPRS

Attualmente, i servizi di trasmissione dati offerti dal sistema GSM sono a commutazione di circuito e consentono velocità di trasmissioni fino a 9600 bit/s. Nell'immediato futuro sarà introdotto un nuovo servizio, denominato *High Speed Circuit-Switched Data* (**HSCSD**) che consentirà un incremento fino a circa 56 kbps. HSCSD, differentemente dal GSM normale che assegna un solo timeslot per connessione, può riservare più timeslot (fino ad un massimo di 4) ad una singola connessione. Vengono, inoltre, introdotti i concetti di *bandwidth on demand* e configurazione *uplink-downlink* asimmetrica.

Anche in questo caso, il servizio si basa su una modalità di trasferimento a *commutazione di circuito*: la rete, una volta stabilita la connessione fisica punto punto fra i due utenti, assegna le proprie risorse fino a quando il rilascio non viene esplicitamente richiesto, indipendentemente dal fatto che i due utenti si scambino dati durante la connessione: in definitiva viene mantenuto un “canale” sempre attivo.

Questa modalità di trasferimento è efficiente solo nel caso in cui i due utenti debbano scambiarsi una mole sensibile di dati (ad esempio con uno trasferimento file); diventa invece inefficiente non appena i dati da scambiare sono di modesta entità oppure, ed è il caso più frequente, il traffico dati è di tipo interattivo o *bursty*, cioè il tempo di utilizzo effettivo delle risorse di rete è solo una parte rispetto al

tempo complessivo di connessione (come ad esempio la navigazione su internet tramite World Wide Web). Trasmissioni con queste caratteristiche necessiterebbero di una connessione a commutazione di pacchetto, modalità che non è ancora offerta dal sistema GSM.

Si pone cioè per il GSM lo stesso problema sorto qualche anno fa per la PSTN (*Public Switched Telephone Network*): prevedere una modalità di trasferimento a pacchetto, in cui i dati degli utenti, racchiusi in entità di protocollo autosufficienti con indicazione del mittente e del destinatario, possono essere trasportati dalla rete senza la necessità di una stretta associazione con un circuito fisico.

Il *General Packet Radio Service (GPRS)* è stato pensato e specificato proprio per colmare questa lacuna e offrire un vero servizio di trasmissione *a commutazione di pacchetto*.

Il GPRS offre notevoli vantaggi a tutte le applicazioni che intendano sfruttare il sistema GSM come mezzo per la trasmissione dati.

1.6.1 Descrizione del sistema

Oggi il GSM offre solo un servizio a commutazione di circuito per la trasmissione dati, ciò significa che ad ogni utente viene assegnato in modo esclusivo un canale di traffico che rimane a lui assegnato per tutta la durata della sessione. Le risorse sono così gestite secondo la modalità *resource reservation*, ossia sono impegnate fino al momento in cui non ne viene fatta esplicita richiesta di rilascio.

Nello schema di figura 1.9 è schematizzato il modello di trasmissione dati a commutazione di circuito: ad ogni utente è riservato un canale (rappresentato nella figura come sequenza di timeslot) che, se la trasmissione non è continua, può

impegnare solo parzialmente. Esiste poi una entità nella rete GSM che provvede a riunire tutti i singoli canali dati convertendoli in un unico flusso a commutazione di pacchetto da instradare verso le reti esterne a pacchetto (X.25 o Internet).

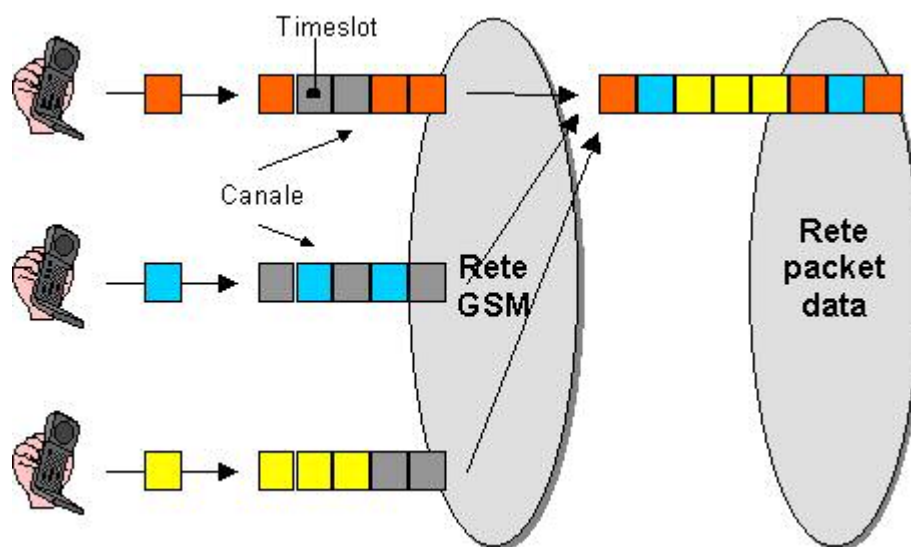


Fig 1.9 - Modello semplificato di trasmissione dati a commutazione di circuito.

Nel GPRS si adotta la tecnica del *context reservation*, le risorse radio sono effettivamente impegnate solo quando c'è la necessità di inviare o ricevere dati. Le stesse risorse radio di una cella sono così condivise da tutte le stazioni mobili (MS) presenti in quella cella, aumentando notevolmente l'efficienza del sistema.

Il servizio GPRS, pertanto, è rivolto ad applicazioni con le seguenti caratteristiche:

- Trasmissione poco frequente di piccoli o grossi volumi di dati (ad esempio applicazioni interattive).
- Trasmissione intermittente di traffico dati *bursty* (ad esempio applicazioni in cui il tempo medio tra due transazioni consecutive è di gran lunga superiore alla durata media di una singola transazione).

La commutazione di pacchetto offre diversi vantaggi:

- Permette più connessioni su un unico canale trasmissivo (*multiplexing*): ciò significa che il canale diventa un mezzo condiviso da più connessioni e non viene quindi assegnata nessuna risorsa di rete in modo esclusivo.
- Non si richiede l'instaurazione di un circuito dedicato tra sorgente e destinatario, quindi la fase di set-up della connessione risulta notevolmente velocizzata.
- È possibile eliminare le costose apparecchiature che, nel GSM base, fungevano da *gateway* per consentire l'interconnessione tra la rete GSM e le reti a commutazione di pacchetto. Queste ultime ormai costituiscono la maggior parte delle reti dati commerciali esistenti (come, per esempio, Internet).

La rete può riservare uno o più timeslot di una portante per il servizio GPRS. Ogni timeslot così configurato costituisce un canale di traffico GPRS, detto *Packet Data Traffic Channel* (PDTCH). Su questo canale possono viaggiare, in trame diverse, pacchetti di utenti diversi; l'allocazione della banda ai singoli utenti è dinamica e dipende dal throughput istantaneo richiesto dall'utente e dalla banda disponibile.

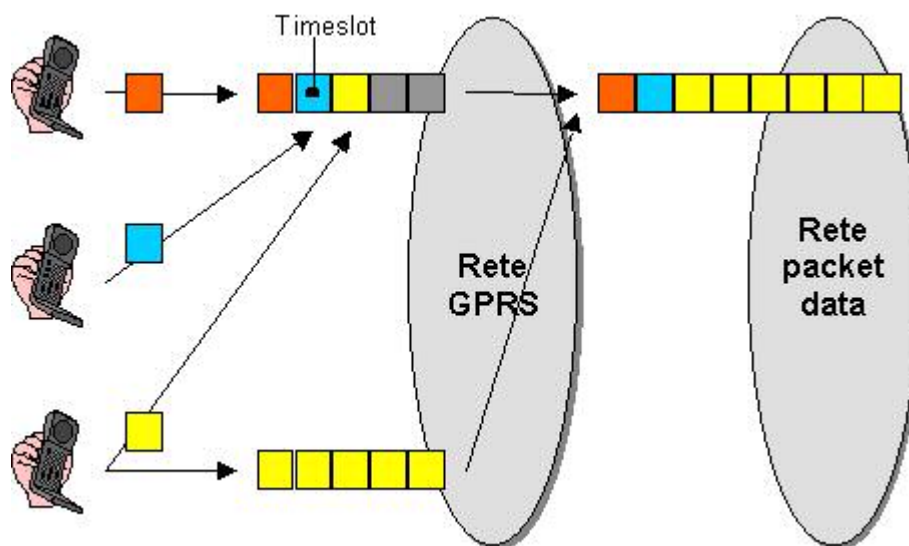


Fig 1.10 - Modello semplificato di trasmissione dati GPRS a commutazione di pacchetto.

Nello schema di figura 1.10 è schematizzato il modello di trasmissione dati GPRS. Si osservi come due utenti (rosso e azzurro), che richiedono un basso throughput, condividano lo stesso canale di traffico GPRS, mentre il terzo (giallo), che invece necessita di un'alta banda, ne utilizzi addirittura due. Ciò dimostra come

si riesca a sfruttare in modo ottimale la risorsa radio. Nel caso in esempio bastano due canali per garantire il servizio a tre utenti; nel GSM attuale a commutazione di circuito sarebbero stati necessari invece tre canali di traffico, uno per utente.

Il GPRS introduce inoltre una nuova codifica di canale che consente di raggiungere i 20 kbps su un singolo timeslot, riducendo ovviamente al minimo la ridondanza per il controllo degli errori.

In teoria, accorpendo tutti gli 8 timeslot di una portante, il GPRS può offrire un data rate di 160 kbps; in realtà l'accorpamento non supererà i 5 timeslot per un data rate comunque notevole di 100 kbps.

Assieme a tutti gli aspetti positivi appena enunciati esistono anche alcuni aspetti negativi. I pacchetti, viaggiando indipendentemente gli uni dagli altri, possono seguire percorsi diversi ed arrivare al ricevente non nello stesso ordine in cui sono stati spediti. Ogni singolo pacchetto deve quindi contenere tutte le informazioni per poter arrivare a destinazione e fornire al ricevente la possibilità di ricostruire l'esatta sequenza di trasmissione.

Il tempo di viaggio dei pacchetti ed il conseguente ritardo con cui arrivano a destinazione è però aleatorio, costituendo quindi un problema per tutte quelle applicazioni che necessitano invece di un data rate costante. Un altro problema si verifica quando tutti gli utenti che utilizzano il medesimo canale GPRS vogliono comunicare nello stesso momento, creando così una congestione con conseguente ritardo dovuto alle code di trasmissione (che si somma a quello descritto in precedenza).

Il GPRS prevede diverse soluzioni, per minimizzare i problemi dovuti all'aleatorietà del tempo di trasmissione. Tra questi l'allocazione flessibile della risorsa radio che permette al GPRS di assegnare in modo esclusivo un canale ad una connessione che richieda una determinata *Quality of Service* (QOS), creando un *circuito virtuale* (*virtual circuit*). Ad esempio per permettere una comunicazione vocale via GPRS, la rete può riservare un canale ad un solo utente, canale che quindi non è più una risorsa condivisa. Se è necessario un data rate maggiore (ad esempio

per trasmissioni video), si può arrivare ad assegnare più canali (quindi più timeslot) in modo esclusivo ad uno stesso utente.

Per garantire una certa qualità del servizio non solo nella tratta radio, ma anche durante la trasmissione attraverso un backbone GPRS di terra, sono stati previsti alcuni meccanismi di gestione intelligente delle code e di prenotazione delle risorse. Il GPRS prevede quattro tipi differenti di QOS predefiniti, ognuno con una differente priorità di trasmissione.

Tali meccanismi permettono di ottimizzare l'uso della risorsa radio e ridurre i problemi intrinseci della commutazione a pacchetto. Da notare che i medesimi problemi erano sorti anche nell'evoluzione di Internet, e che le soluzioni, attive e testate con successo ormai da decenni possono essere trasportate nel sistema GPRS.

1.6.2 Architettura di una rete GPRS

Per implementare il GPRS si possono seguire due diversi approcci: l'approccio di sistema separato oppure l'approccio di sistema integrato.

Il primo approccio prevede che tutta l'infrastruttura necessaria per il supporto del servizio sia separato rispetto a quella della rete GSM, mentre il secondo prevede l'aggiunta delle funzionalità necessarie per il supporto del GPRS alle entità che già compongono l'infrastruttura della rete GSM. In realtà anche l'approccio "integrato" richiede l'introduzione di nuove entità, garantendo comunque, dal punto di vista economico, un impatto meno vistoso sui costi necessari per l'implementazione del servizio. Ogni singolo produttore seguirà l'approccio che reputa migliore per le proprie esigenze, entrambi presentano, infatti, vantaggi e svantaggi.

La rete GSM completa può essere pensata come scomposta in due parti: una sottorete GSM a commutazione di circuito ed una sottorete GPRS a commutazione di pacchetto. Le sottoreti sono logicamente isolate, ma condividono il sottosistema BSS (*Base Station SubSystem*) e HLR (*Home Location Register*).

Il GPRS introduce alcune nuove entità di rete: il *Serving GPRS Support Node* (SGSN), il *Gateway GPRS Support Node* (GGSN) e il *Border Gateway* (BG).

I nodi SGSN possono essere visti come entità in cui sono localizzate gran parte delle funzioni necessarie per supportare il GPRS. Generalmente vi sono più nodi SGSN e l'infrastruttura che li connette, denominata *backbone network*, consente il routing dei pacchetti trasmessi dagli utenti della rete o a questi indirizzati.

Il gateway GGSN fornisce la connettività verso le altre reti dati, ad esempio X.25 o Internet, il border gateway BG invece verso le reti GPRS degli altri operatori. Le funzioni principali svolte da queste entità sono la conversione dei protocolli ed il mapping degli indirizzi di rete degli utenti coinvolti nella comunicazione dei dati.

Le funzionalità di SGSN, GGSN e BG possono essere integrate in un unico elemento della rete, che può coincidere con un MSC (*Mobile Switching Center*), oppure possono essere tutti elementi separati.

Un'altra nuova entità necessaria per il supporto del servizio è il *GPRS register* (GR), che non necessariamente va vista come una nuova entità fisica, in quanto si può pensare di estendere l'insieme di funzioni dei VLR/HLR della rete GSM. Le funzioni espletate da un GPRS register sono essenzialmente quelle di memorizzare informazioni riguardanti l'abbonato del servizio GPRS; in particolare ogni GPRS register contiene:

- Informazioni necessarie per il routing dei pacchetti indirizzati ad un mobile GPRS; ad esempio l'indirizzo di rete del mobile per un dato protocollo di rete ed il tipo di protocollo di rete cui l'indirizzo si riferisce.
- Informazioni relative al profilo di sottoscrizione dell'abbonato; ad esempio informazioni caratterizzanti la qualità del servizio richiesta dall'utente (*Quality Of Service*).

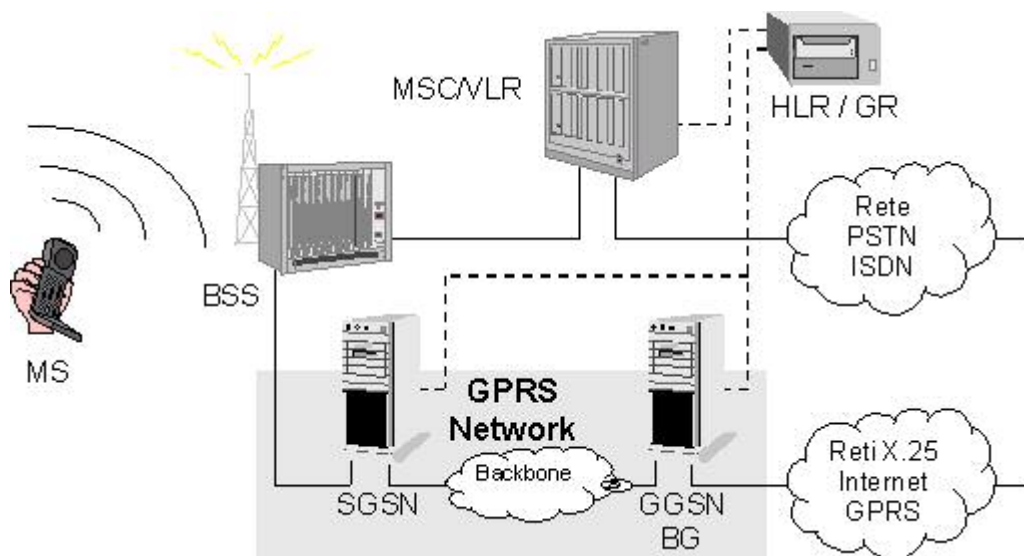


Figura 1.11 - Architettura semplificata di una rete GSM/GPRS.

1.6.3 Caratteristiche del GPRS in breve

Concludendo, le principali caratteristiche del GPRS sono:

- Gestione flessibile ed efficiente della risorsa radio.
- Possibilità per l'utente di inviare e ricevere dati nella modalità di trasferimento end-to-end a commutazione di pacchetto senza utilizzare nessuna risorsa di rete in modalità a commutazione di circuito.
- Riunisce tutti i positivi vantaggi della commutazione di pacchetto:
- più connessioni possono condividere la medesima risorsa di rete (il canale trasmissivo);
- possibilità di un servizio *connectionless*;
- maggior velocità di set-up della connessione;
- possibilità di inviare messaggi in *Multicast* o *Broadcast*;
- Le caratteristiche negative sono invece minimizzate mediante l'uso di meccanismi di priorità e QoS;
- Possibilità per l'utente di trasmettere contemporaneamente anche su più timeslot;
- Velocità di trasmissione fino a 100 kbps utilizzando 5 timeslot, con un data rate di 20 kbps per timeslot.

1.7 Inquinamento Elettromagnetico

La realizzazione di una rete radiomobile, sia essa GSM, Etacs oppure DCS, richiede l'installazione di un gran numero di stazioni radio base (BTS), alcune migliaia. E' inevitabile che la determinazione dei livelli di campo elettromagnetico prodotti nell'intorno di quest'ultime e il loro contenimento siano di fondamentale importanza nella progettazione della rete stessa.

Nel 1988 l'IRPA (*International Radiation Protection Association*) fu il primo organo internazionale che emanò delle indicazioni sugli standard di sicurezza, che furono poi recepite dalle norme protezionistiche emanate in diversi Paesi.

Nel 1992 la commissione dell'ANSI/IEEE sui pericoli delle radiazioni non-ionizzanti produsse la normativa ANSI/IEEE C95.1-1992: *Standard per i livelli di sicurezza nel rispetto dell'esposizione umana ai campi elettromagnetici da radio frequenza (Standard for safety levels with respect to human exposure to radio frequency electromagnetic fields, 3 kHz to 400 GHz)*.

Nel 1996 la FCC (*Federal Communications Commission*) di Washington stabilì le linee base per valutare gli effetti delle radiazioni elettromagnetiche (FCC 96-326).

In Italia, nel 1982 fu emanata una Circolare del Ministero della Sanità; successivamente fu redatta una proposta di Decreto della Presidenza del Consiglio che, nonostante sia stata rivista ed aggiornata numerose volte negli anni successivi, non è stata mai tradotta in norma.

In assenza di una normativa nazionale, alcune Regioni hanno emanato leggi regionali a riguardo: Abruzzo, Lazio, Molise, Piemonte e Veneto.

Gli operatori italiani TIM e Omnitel hanno sviluppato le loro reti radiomobili seguendo le normative IRPA 1988 e ANSI/IEEE 1992, e secondo le varie norme regionali. Come livello limite di esposizione per la protezione sanitaria nella banda 900 MHz è stato fissato un valore di campo elettrico pari a 20 V/m.

Le stazioni radio base sono installate in posizione sempre rialzata dal terreno, almeno 12 metri, le antenne sono posizionate in modo da irradiare il segnale a distanza, limitando il più possibile l'emissione di radiazioni elettromagnetiche nelle aree circostanti accessibili alla popolazione.

Attraverso una complicata modellizzazione matematica è possibile calcolare il “volume di rispetto” per ogni configurazione possibile di stazione radio base, sia che utilizzi antenne omnidirezionali che direttive. E' così possibile valutare con precisione *le distanze di protezione sanitaria*, cioè il volume attorno alla stazione che deve essere interdetto alla popolazione. Al di fuori di detto volume il campo elettrico è sicuramente entro i limiti stabiliti di 20 V/m.

Gli aspetti salienti si possono così riassumere:

- Le antenne convogliano tutta la loro potenza in un ristretto angolo verticale, quindi anche posizionandosi proprio sotto la stazione radio, basta mantenersi ad almeno 1.6 m dalle antenne per uscire dal volume di rispetto.
- La distanza di rispetto frontale per le antenne omnidirezionali e' di circa 5 m, oltre tale distanza i limiti sono rispettati.
- La distanza di rispetto frontale per le antenne direzionali è variabile in base al grado di apertura del cono d'illuminazione, comunque non supera mai i 10 m.

Va osservato che proprio per esigenze tecniche le stazioni radio base sono collocate su posizioni rialzate, tralicci o pali autoportanti, come tetti di edifici. Il volume di rispetto risulta quindi, il più delle volte, fisicamente inaccessibile alla popolazione.

In conclusione, la ridotta potenza emessa dalle stazioni radio base (circa 50 watt) e i particolari siti scelti per l'installazione che garantiscono il rispetto delle distanze di protezione, assicurano la loro non pericolosità per la popolazione, quantomeno alla luce degli studi finora compiuti.

Capitolo II

LA PIANIFICAZIONE FREQUENZIALE

2.1 Introduzione

La determinazione di un piano di frequenza (Pianificazione frequenziale), ossia dell'assegnazione delle frequenze disponibili ad un insieme predefinito di celle, rappresenta uno delle fasi più delicate nella progettazione di un sistema radiomobile cellulare.

L'obiettivo è quello di assegnare a ciascuna cella un numero prestabilito di frequenze (una per ogni portante radio con cui è equipaggiata la cella), tenendo conto di diversi aspetti e vincoli che influenzano la costruzione del sistema radiomobile.

In passato la pianificazione frequenziale era realizzata suddividendo il territorio da pianificare in un certo numero di *cluster* (vedi Figura 2.1), ove un cluster può essere definito come il più piccolo insieme di celle che utilizzano l'intera banda disponibile. Ad ogni cluster era associato un gruppo di frequenze costituito da un numero costante di frequenze equispaziate.

Il problema era in tal caso notevolmente semplificato in quanto si supposeva che le celle avessero dimensione e forma regolari e che interferissero tra loro nello stesso modo. La realtà è diversa in quanto le dimensioni, la forma delle celle e gli

effetti dell'interferenza reciproca tra le portanti radio dipendono direttamente dalle caratteristiche del territorio (orografia e grado di propagazione) e non possono perciò essere descritti con sufficiente approssimazione da un semplice modello geometrico. Alcuni studi indicano come anche le variazioni climatiche stagionali all'interno della stessa regione influenzino sensibilmente tali parametri.

La considerazione degli aspetti geografici per ciascuna cella e la grande espansione dei sistemi radiomobili con conseguente incremento del numero di celle dislocate sul territorio hanno reso ben presto inevitabile l'impiego di metodi automatici per la definizione di piano di frequenze.

Nel presente capitolo saranno illustrate le principali caratteristiche del problema dell'assegnazione frequenziale per i sistemi radiomobili di prima e seconda generazione.

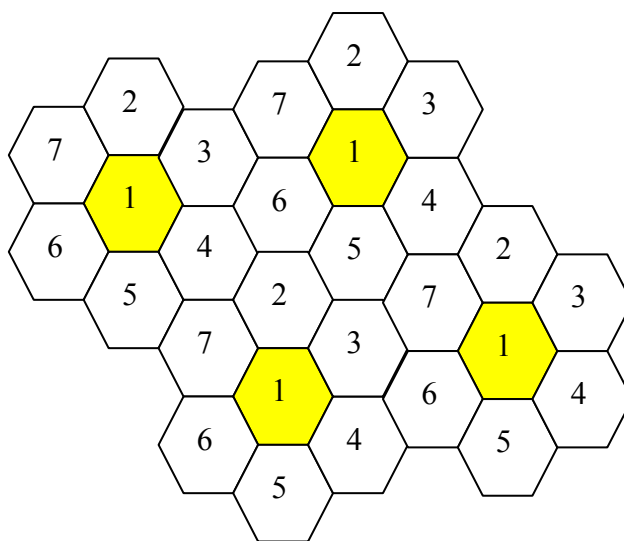


Figura 2.1 – Copertura cellulare a celle esagonali

2.2 Descrizione del problema

La progettazione di una rete radiomobile cellulare è un processo complesso e articolato. Esso richiede competenze che provengono da diversi settori: coinvolge, infatti, aspetti di trasmissione e propagazione dei segnali radio, considerazioni legate

alla tipologia delle reti di comunicazione, previsioni sull'entità del traffico offerto e sulla sua distribuzione all'interno dell'area di servizio, e allocazione ottima delle risorse (siano esse portanti radio, canali di comunicazione o apparati di Rete).

La stima del traffico offerto nelle regioni raggiunte dal servizio è di solito ottenuta utilizzando basi di dati territoriali e cartografiche, grazie alle quali è possibile individuare le vie di comunicazione principali e i punti in cui si prevede un'alta densità di utenza all'interno dei più importanti centri urbani (per esempio stazioni ferroviarie, uffici pubblici, centri commerciali). Per quanto riguarda le valutazioni legate agli aspetti trasmissivi e propagativi, i dati di maggiore interesse sono il risultato dello studio dell'orografia del territorio analizzato, del rilevamento del suo grado di urbanizzazione e di eventuali campagne di misura effettuate "sul campo". L'opportuna combinazione dei dati provenienti dalle stime di traffico, dal modello di propagazione e dalle misure, permette di selezionare un certo numero di punti del territorio in cui posizionare le stazioni radio base (BTS), cioè i tralicci su cui vengono installate le antenne direzionali. Ciascuna BTS viene successivamente "dimensionata" dal punto di vista radioelettrico; viene cioè stabilito, per essa, il numero (equipaggiamento, o dimensionamento radio) di portanti radio necessarie per smaltire il traffico offerto, opportunamente stimato, utilizzando una preassegnata qualità del servizio.

La localizzazione e il dimensionamento di ciascuna stazione radio base determinano una suddivisione (copertura) del territorio in celle, ciascuna con una propria estensione geografica ed una propria forma. Si ipotizza che per ciascun punto del territorio esista un'antenna per la quale il segnale venga captato con la minima intensità di disturbo (antenna best server), ed è quindi possibile suddividere la regione in celle assegnando a ciascuna antenna tutti i punti del territorio per i quali essa risulta essere best server. Generalmente, sul traliccio di una stazione radio base vengono installate tre antenne direzionali, in grado ciascuna di coprire un settore circolare dell'ampiezza di 120° ; si realizza così un insieme di tre celle adiacenti facenti capo alla stessa BTS, ovvero "coperte" da antenne installate sullo stesso traliccio. L'insieme delle celle facenti capo ad una stessa BTS prende il nome di *sito*. I siti trisetoriali, come quello appena descritto, sono i più comuni; esistono tuttavia anche siti a due celle (bisettoriali) o a singola cella (un'unica antenna omnidirezionale con apertura di 360°).

Una volta determinate la posizione, l'estensione della copertura ed il dimensionamento radio di ciascuna cella, è necessario progettare la rete di interconnessione fra le stazioni radio base (BTS) e gli altri elementi della rete gerarchicamente superiori (BSC, MSC, ecc.), cercando di ottimizzare le prestazioni ed il costo complessivo dovuto agli apparati e ai collegamenti veri e propri.

Terminata la fase di progettazione strutturale della rete radiomobile, è necessario procedere ad una fase di progettazione funzionale della stessa, in modo da stabilire dei criteri operativi e gestionali che ottimizzino il rendimento del sistema e minimizzino i costi complessivi. La pianificazione frequenziale è un passo cruciale di questo complesso processo di progettazione. I vincoli di cui dovrà tener conto saranno sia di ordine tecnico-costruttivo (dovuti alle caratteristiche strutturali del sistema), sia di ordine funzionale (dovuti alla qualità del servizio).

Nel successivo paragrafo si descrivono più dettagliatamente tali vincoli e le loro implicazioni nella definizione di un piano di frequenza.

2.3 Vincoli del problema

Un piano di frequenze è caratterizzato da un certo numero di vincoli da tenere in considerazione quando si voglia assegnare a ciascuna cella il numero prefissato di frequenze; fra questi, i principali sono:

- la banda utile, cioè il numero di frequenze disponibili (canali radio)
- l'equipaggiamento delle celle, cioè il numero di frequenze da assegnare a ciascuna cella;
- i vincoli di riuso, cioè l'insieme di criteri che devono essere rispettati in fase di assegnazione;
- i livelli interferenziali, cioè i disturbi a cui sono soggette le portanti di ciascuna cella;

2.3.1 Vincoli di banda utile

La banda utile è costituita dall'insieme di frequenze disponibili per la definizione di un piano di frequenza. Il sistema GSM900 (ne esiste anche una versione a 1800 MHz) opera intorno alla frequenza 900 MHz. Più precisamente, ad esso sono state riservate due bande di frequenza distinte:

- **880 MHz - 915 MHz** per le comunicazioni dal mobile verso la BT (tratta uplink);
- **925 MHz - 960 MHz** per le comunicazioni dalla BTS verso il mobile (tratta downlink).

In ciascuna sottobanda sono allocate 124 portanti, spaziate fra loro di 200 KHz (segmentazione in frequenza). L'accesso da parte dei diversi operatori radiomobili a questa risorsa comune e limitata viene disciplinato in ambito nazionale da opportune direttive ministeriali. La porzione di spettro utilizzabile da un singolo gestore dipende dal numero di licenze concesse per il sistema (ossia dal numero di gestori concorrenti) e dall'eventuale compresenza di altri servizi su parte della banda (per esempio il sistema analogico TACS, i sistemi radiomobili militari, ecc.). Ad esempio, la banda utilizzata per la pianificazione frequenziale del sistema GSM di Telecom Italia Mobile è costituita attualmente da 57 portanti. Ogni portante viene suddivisa nel dominio del tempo in trame di 8 time slot ciascuna (segmentazione nel tempo). Ciascuno di questi time slot può essere utilizzato per “ospitare” canali di traffico (destinati all'utenza) o di segnalazione (destinati ai sistemi preposti al controllo). Normalmente viene destinata a canale di segnalazione e controllo della cella quella portante che presenta il disturbo complessivo meno elevato (e quindi ha la migliore qualità trasmissiva).

2.3.2 L'equipaggiamento delle celle

L'equipaggiamento di una generica cella corrisponde al numero di frequenze che devono essere assegnate alla cella medesima; l'insieme degli equipaggiamenti associati alle celle in via di pianificazione è descritto dalla relazione

$$E = (D_1, \dots, D_N) \quad (2.1)$$

dove N è il numero di celle facenti parte del piano di frequenza da definire e D_j è l'equipaggiamento associato alla generica cella j . Esso è funzione del traffico telefonico offerto (misurato o stimato) che la generica cella deve smaltire, dipendente a sua volta dalla concentrazione degli utenti che accedono al sistema nella zona servita dalla cella in oggetto. In quanto funzione del traffico offerto, gli elementi D_j possono assumere valori differenti; attualmente, il numero di frequenze assegnate ad una cella del sistema, e quindi il numero di portanti presenti nella cella, va da un minimo di uno ad un massimo di dieci. Perciò il numero complessivo di assegnazioni da effettuare, (pari a $\sum_j D_j$), influisce sulla qualità del piano in modo combinato con l'ampiezza N della banda utile; una buona misura della complessità di un piano di frequenza è esprimibile dall'indice di riuso medio, dato dalla relazione

$$R = \frac{\sum_{j=1}^N D_j}{N} \quad (2.2)$$

corrispondente al numero di volte in cui ciascuna frequenza è mediamente utilizzata all'interno del piano in via di definizione.

A causa della limitata banda utile (che contiene poche decine di frequenze), dell'elevato numero di celle degli attuali sistemi GSM (molte migliaia) e dell'equipaggiamento medio (4 o 5 frequenze per cella), la definizione di un piano di frequenza ammissibile ha come inevitabile conseguenza il forte riuso delle frequenze disponibili (riuso frequenziale spinto). Se, ad esempio, in un generico sistema cellulare si proponga di servire 5000 celle, con un dimensionamento medio di 4

portanti per cella, utilizzando 57 frequenze, è necessario che ogni portante venga assegnata in media a circa

$$\frac{5000 \cdot 4}{57} \cong 351 \quad \text{celle.}$$

Ciò significa che una stessa frequenza viene assegnata centinaia di volte; a questo punto è evidente la possibilità di interferenze reciproche non trascurabili.

2.3.3 I vincoli di riuso

I vincoli di riuso rappresentano una serie di criteri che devono essere rispettati in fase di assegnazione delle frequenze alle celle, in quanto si delineano come vincoli di ordine tecnologico.

La procedura di *handover*, ad esempio, che prevede il passaggio della chiamata in corso da una cella ad un'altra senza l'interruzione della conversazione, può avere luogo solo se le celle interessate dalla transizione impiegano frequenze diverse. E' quindi necessario che alle portanti di celle fisicamente confinanti vengano assegnate frequenze diverse. Tale concetto può essere espresso dalla condizione che celle adiacenti fisicamente devono avere assegnazioni frequenziali che distino, in banda, almeno 200 KHz (pari a un canale).

Esiste un altro caso in cui si rende necessario equipaggiare due celle con insiemi di portanti disgiunte a due a due. Può accadere, infatti, che all'interno di una cella esista una zona in cui venga percepito il segnale utile trasmesso dall'antenna di un'altra cella, non necessariamente adiacente fisicamente; questo porta a determinare un pessimo rapporto segnale-disturbo e a compromettere la qualità delle connessioni all'interno della cella "disturbata". Questo fenomeno viene denominato *adiacenza radioelettrica* e si verifica in particolare quando l'area disturbata è piuttosto piccola rispetto all'area di servizio totale, tanto da non renderne conveniente un'eventuale separazione dalla cella. Si intuisce di conseguenza, come sia necessario equipaggiare coppie di celle adiacenti radioelettricamente con portanti diverse fra loro, per evitare

che, nella cella disturbata, una stazione possa ricevere, con intensità confrontabili, due segnali diversi alla stessa frequenza.

In generale, celle adiacenti fisicamente sono quasi sempre adiacenti anche dal punto di vista radioelettrico; non è impossibile però che celle anche molto distanti siano invece adiacenti radioelettricamente, interferendosi vicendevolmente in modo significativo.

Oltre a questa adiacenza, che possiamo chiamare “adiacenza a 200 KHz”, esistono anche coppie di celle “adiacenti a 400 KHz”, che dovranno essere assegnate frequenze distanti tra loro due canali, ossia appunto 400 KHz.

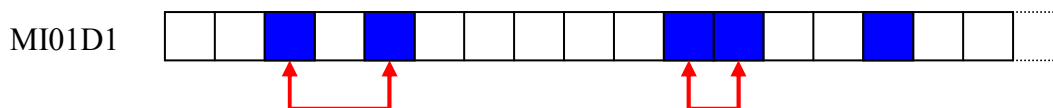


Figura 2.2 - Violazione del vincolo di spaziatura(frecce rosse) in banda di 400KHz e 200KHz nella stessa cella (MI01D1)

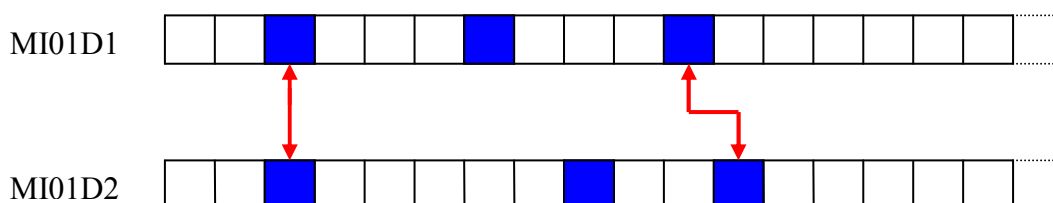


Figura 2.3 - Violazione del vincolo di adiacenza e di 200KHz per celle dello stesso sito (MI01D1 e MI01D2)

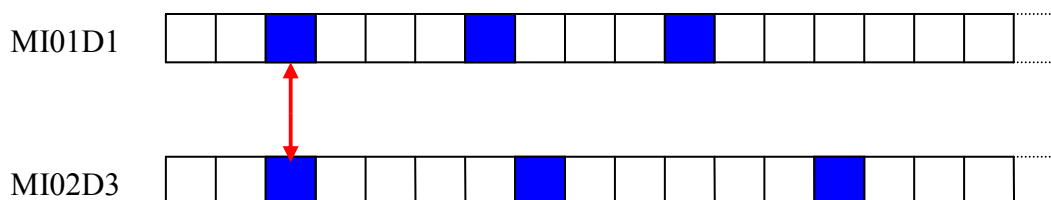


Figura 2.4 - Violazione del vincolo di adiacenza per celle “adiacenti a 200KHz” (MI01D1 e MI02D3)

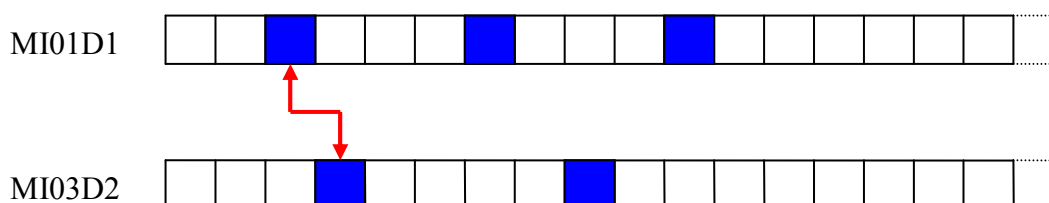


Figura 2.5 - Violazione del vincolo di adiacenza per celle “adiacenti a 400KHz” (MI01D1 e MI03D2)

L’assegnazione delle frequenze è soggetta a due ulteriori vincoli, causati da problemi di ordine tecnico. Le portanti assegnate a celle appartenenti allo stesso sito (cioè servite da antenne installate sullo stesso traliccio) devono avere frequenze che distino in banda almeno 400 kHz (2 canali), per problemi relativi alla realizzazione degli apparati di trasmissione. Le portanti, invece, appartenenti ad una stessa cella (qualora il dimensionamento radio di questa sia maggiore di uno) dovranno avere una spaziatura in banda di almeno 600 kHz (3 canali), per problemi tecnologici analoghi a quelli visti per il sito, ma più stringenti a livello di singola cella. Attualmente sono in via di commercializzazione apparati che permettono di diminuire la distanza in banda di 200 kHz, sia per la cella che per il sito; ciò significa che la spaziatura per portanti appartenenti alla stessa cella scenderebbe da 600 kHz a 400 kHz, mentre quella relativa a celle facenti capo allo stesso traliccio passerebbe dai 400 kHz ai 200 kHz. Tali apparati sono tuttavia molto costosi tanto da non renderne conveniente l’applicazione pratica.

I vincoli descritti nel presente paragrafo non sono, in linea teorica, rilassabili in alcun modo: un piano di frequenza che non li rispetti non può cioè essere realizzato fisicamente. Se per un certo problema si giungesse alla conclusione che tali vincoli non permettono alcuna soluzione, sarebbe quindi necessario procedere alla formulazione di un nuovo problema, ad esempio con diminuzione arbitraria degli equipaggiamenti su celle le cui portanti violano vincoli. L’assegnazione delle frequenze, quindi, non è separata dalla progettazione della rete ma è integrata con essa e contribuisce alla sua evoluzione.

2.3.4 Vincoli interferenziali

La qualità di un piano di frequenza è funzione del livello di interferenza subito e indotto da ciascuna portante assegnata. Due portanti interferiscono fra loro in misura proporzionale alla distanza in banda fra le frequenze a loro assegnate e in modo inversamente proporzionale alla quarta potenza della distanza fisica che separa le due stazioni emittenti. Per il calcolo del livello globale di disturbo subito o indotto da una portante generica, ai fini pratici è sufficiente considerare l'interferenza causata dalle portanti che utilizzano la stessa frequenza (interferenza co-canale) o le prime adiacenti (interferenza da canale adiacente). Si suppone, quindi, che due celle interferiscano tra loro in modo apprezzabile, (cioè percepibile durante la connessione), solo quando a due loro portanti generiche siano assegnate frequenze distanti non più di 200 KHz. I livelli interferenziali subiti e indotti da ciascuna cella del piano considerato possono essere descritti dalla cosiddetta *matrice d'interferenza* C . Quest'ultima è una matrice quadrata di ordine N (dove con N indichiamo il numero di celle facenti parte del piano) il cui generico elemento rappresenta il rapporto fra la potenza del segnale utile e la potenza del segnale interferente associato alla cella i quando la cella j utilizza la stessa frequenza (interferenza co-canale). Spesso i valori vengono espressi in decibel; una misura logaritmica permette l'uso operativo di valori diversi anche di parecchi ordini di grandezza.

$$P_{dB} = 10\log_{10}(P) \quad (2.3)$$

Attraverso la matrice di interferenza è possibile stimare, oltre all'interferenza co-canale, l'interferenza da canale adiacente. Quest'ultima è infatti approssimata all'interferenza co-canale attenuata da un fattore NFD (Net Filter Discriminator) del valore di 18dB (pari a 63.1).

Se indichiamo con $I_{i,j}$ l'inverso di $C_{i,j}$ (cioè il rapporto tra segnale interferente e segnale utile per interferenza co-canale) sarà:

$$I_{i,j}^{adj} = \frac{I_{i,j}}{NFD} \quad (2.4)$$

dove

$I_{i,j}$ è il rapporto potenza segnale interferente su potenza segnale utile per interferenza co-canale;

$I_{i,j}^{adj}$ è il rapporto potenza segnale utile su potenza segnale interferente per interferenza da canale adiacente.

L'interferenza complessiva percepita dalla generica portante p assegnata alla cella i è dovuta all'interazione con l'insieme della frequenze co-canali e adiacenti distribuite all'interno del piano di frequenza definito. Tale interferenza è valutata attraverso la relazione:

$$I_{i,p}^{tot} = \sum_{\substack{j=1 \\ j \neq i}}^N \left(I_{i,j} A_{j,p} + \frac{I_{i,j}}{NFD} A_{j,p-1} + \frac{I_{i,j}}{NFD} A_{j,p+1} \right) \quad (2.5)$$

dove

$I_{i,p}^{tot}$ è l'interferenza complessiva subita dalla portante p della cella i

$A_{j,p} = 0$ se alla cella j non è assegnata la frequenza p

$A_{j,p} = 1$ se alla cella j è assegnata la frequenza p

Per garantire un buon funzionamento del sistema, ogni portante deve essere caratterizzata da un valore complessivo del rapporto segnale-disturbo non inferiore a 9 dB. Purtroppo, a causa della limitatezza della banda e dell'elevato numero di assegnazioni da effettuare, non è sempre possibile rispettare questo vincolo per tutte le portanti del piano. L'obiettivo diventa allora quello di massimizzare il valore del rapporto segnale utile-disturbo per ogni portante, cercando di contenere contemporaneamente il numero totale di portanti che non superano la soglia dei 9dB (rilassamento del vincolo di soglia).

2.3.5 Vincoli di frazionamento

Inizialmente, le problematiche connesse alla pianificazione frequenziale riguardavano solo centri urbani particolari, come Roma, Napoli, Milano, ecc., caratterizzate dalla presenza sul territorio di alcune centinaia di celle. L'espandersi progressivo delle aree coperte dal servizio ha richiesto l'estensione su scala nazionale della procedura di pianificazione frequenziale. Ancora oggi, tuttavia, metà circa delle celle del sistema sono concentrate nelle sole città di Roma e Milano.

Fino a poco tempo fa si riteneva, quindi, che fosse impossibile trattare il problema nella sua globalità; nel tentativo di pianificare separatamente celle fisicamente distanti e considerate ad interagenzia nulla, il territorio nazionale è stato quindi suddiviso in sottoregioni, ciascuna considerata come una “zona” separata.

Applicando una pianificazione separata a ciascuna sottoregione si poteva così suddividere la complessità del problema; lo svantaggio è l'evidente sub-ottimalità della soluzione complessiva, unito alla possibilità che assegnazioni buone per una regione potessero causare problemi insormontabili alla pianificazione di una zona adiacente. Inoltre, si rende necessaria una parziale sovrapposizione fra zone adiacenti (vedi fig. 2.6) , dato che non è possibile isolare sottoinsiemi di celle che interagiscono esclusivamente fra loro. Tale suddivisione del territorio impone perciò una serie di vincoli aggiuntivi, per tener conto delle influenze tra celle contigue pur pianificandole in momenti separati.



Figura 2.6 – Suddivisione del territorio Italiano in sottoregioni

In particolare, tre sono i vincoli aggiuntivi di cui è necessario tener conto:

- cella bloccata: è una cella pianificata precedentemente e quindi non modificabile.
- frequenze proibite per adiacenza: le celle non ancora pianificate ma adiacenti a celle già pianificate (esterne alla sottoregione che si sta considerando) avranno una serie di frequenze proibite a causa dei vincoli imposti dalle celle adiacenti non modificabili.
- frequenze proibite per appartenenza al sito: le celle non ancora pianificate ma appartenenti allo stesso sito di celle già pianificate (esterne alla sottoregione che si sta considerando) avranno una serie di frequenze proibite a causa del vincolo di sito.

2.4 Formulazione del problema

Si considerino N celle ciascuna rappresentata da un numero intero $i \in \{1, \dots, N\}$ e M frequenze possibili rappresentate da un intero $p \in \{1, \dots, M\}$.

Sia $A_{i,p}$, per ogni $i \in \{1, \dots, N\}$, $p \in \{1, \dots, M\}$, una variabile intera così definita:

$$A_{i,p} = \begin{cases} 1 & \text{se la frequenza } p\text{-esima è utilizzata da una portante della cella } i \\ 0 & \text{altrimenti} \end{cases}$$

Se $A_{i,p} = 1$ si usa anche dire che il canale p è stato assegnato ad una portante della cella i ; le due espressioni sono equivalenti.

Si definisce *assegnazione di frequenza* y una funzione così definita:

$$y : A \subseteq N \times M \rightarrow \{0,1\}$$

Sia $D_i \geq 0$, per ogni $i \in \{1, \dots, N\}$, una costante intera che rappresenta il numero di portanti da equipaggiare (*dimensionamento o equipaggiamento*) in ogni cella i .

Sia $B_i \subseteq \{1, \dots, M\}$, per ogni $i \in \{1, \dots, N\}$, l'insieme delle frequenze proibite ad ogni cella i . Per ogni $i \in \{1, \dots, N\}$, per ogni $p \in \{1, \dots, M\}$ si ha che se $p \in B_i$ allora $A_{i,p} = 0$.

Sia $X_i \subseteq \{1, \dots, M\}$, per ogni $i \in \{1, \dots, N\}$, l'insieme delle frequenze bloccate (perché già assegnate) in ogni cella i . Per ogni $i \in \{1, \dots, N\}$, per ogni $p \in \{1, \dots, M\}$ si ha che se $p \in B_i$ allora $A_{i,p} = 1$.

Sia $I_{i,j} \geq 0$ una costante reale che indica l'effetto che la cella j provoca sul rapporto disturbo / segnale utile della cella i qualora ad una portante di quest'ultima fosse assegnata la medesima frequenza di una portante della prima (interferenza da cocanale). Si suppone, inoltre, che i vari effetti possano combinarsi linearmente, pertanto si definisce *interferenza complessiva da cocanale* subita da una portante il valore:

$$RDS_{i,p}^{co} = \sum_{\substack{j=1 \\ i \neq j}}^N I_{i,j} A_{j,p} \quad (2.6)$$

L'effetto di una assegnazione si fa sentire anche sul canale adiacente; questo effetto si può approssimare con il termine

$$I_{i,j}^{ad} = \frac{I_{i,j}}{NFD} \quad (2.7)$$

In definitiva si definisce *interferenza complessiva subita da una portante* il valore:

$$RDS_{i,p} = \sum_{\substack{j=1 \\ i \neq j}}^N I_{i,j} A_{j,p} + \sum_{\substack{j=1 \\ i \neq j}}^N \frac{I_{i,j}}{NFD} A_{j,p-1} + \sum_{\substack{j=1 \\ i \neq j}}^N \frac{I_{i,j}}{NFD} A_{j,p+1} \quad (2.8)$$

con la convenzione che $A_{i,p} = 0$ se $p > M$ oppure se $p < 0$, per ogni $i \in \{1, \dots, N\}$.

Si definisce *sito di i* un insieme $S_i \subseteq \{1, \dots, N\} \setminus \{i\}$, per ogni $i \in \{1, \dots, N\}$; se $h \in S_i$ allora h e i appartengono allo stesso sito; in particolare deve essere che per ogni $p \in \{1, \dots, M\}$ se $A_{i,p} = 1$ allora per ogni $k \in \{-2, -1, 0, 1, 2\}$, $A_{h,p+k} = 0$.

Si definisce *insieme degli adiacenti di i a 400kHz* un insieme $A_i^{400} \subseteq \{1, \dots, N\} \setminus \{i\}$, per ogni $i \in \{1, \dots, N\}$; se $h \in A_i^{400}$ allora i è adiacente a 400kHz ad h ; in particolare deve essere che per ogni $p \in \{1, \dots, M\}$ se $A_{i,p} = 1$ allora per ogni $k \in \{-2, -1, 0, 1, 2\}$, $A_{h,p+k} = 0$.

Si definisce *insieme degli adiacenti di i a 200kHz* un insieme $A_i^{200} \subseteq \{1, \dots, N\} \setminus \{i\}$, per ogni $i \in \{1, \dots, N\}$; se $h \in A_i^{200}$ allora i è adiacente a 200kHz ad h ; deve essere che per ogni $p \in \{1, \dots, M\}$ se $A_{i,p} = 1$ allora per ogni $k \in \{-1, 0, 1\}$, $A_{h,p+k} = 0$.

A questi si aggiungono i vincoli interni di cella: per ogni $p \in \{1, \dots, M\}$ per ogni $i \in \{1, \dots, N\}$ se $A_{i,p} = 1$ allora per ogni $k \in \{-2, -1, 1, 2\}$, $A_{i,p+k} = 0$; inoltre la stessa frequenza non può essere assegnata a due portanti della stessa cella: deve essere per ogni $i \in \{1, \dots, N\}$

$$D_i = \sum_{p=1}^M A_{i,p} \quad (2.9)$$

Un ulteriore vincolo riguarda il valore massimo dell'interferenza complessiva di ciascuna portante, che deve essere inferiore di una soglia ξ : per ogni $i \in \{1, \dots, N\}$ per ogni $p \in \{1, \dots, M\}$ se $A_{i,p} = 1$ deve essere $RDS_{i,p} < \xi$.

Si definisce *rete di telecomunicazione radiomobile* R la settupla $(D, I, S, A^{400}, A^{200}, B, X)$.

Si definisce *piano di frequenze* P l'insieme delle variabili $A_{i,p}$ per ogni $i \in \{1, \dots, N\}$ e per ogni $p \in \{1, \dots, M\}$. Sia Π l'insieme di tutti i piani possibili.

Un piano di frequenze $P \in \Pi$ si definisce *ammissibile* se non è violato alcun vincolo in esso definito.

Si definisce *costo* di un piano P il seguente valore reale:

$$C(P) = \sum_{i=1}^N \sum_{p=1}^M A_{i,p} RDS_{i,p} \quad (2.10)$$

Si definisce infine il *problema dell'assegnazione delle frequenze (FAP)*:

Data una rete $R=(D, I, S, A^{400}, A^{200}, B, X)$, trovare un piano $P^* \in \Pi$ ammissibile e di costo $C(P^*)$ minimo.

2.5 Complessità Computazionale

Il problema dell'Assegnazione Frequenziale (FAP) è noto come NP-completo, con una riduzione dal problema di colorazione minima dei grafi. Data quindi la difficoltà intrinseca del problema, e l'enorme dimensione dello spazio delle

soluzioni, gli algoritmi che possono efficacemente essere applicati sono tipicamente di tipo euristico.

2.6 Rilassamento dei vincoli

A livello teorico, qualunque insieme di vincoli originari potrebbe essere rimosso (rilassamento), purchè se ne tenga conto nella funzione obbiettivo. Questa operazione potrebbe rivelarsi inutile, perché potrebbe portare a piani di frequenza inconsistenti perché irrealizzabili tecnicamente o legalmente. Ad esempio, il vincolo di banda, non potrebbe essere rilassato in quanto qualsiasi piano che utilizzi più banda del consentito sarebbe inutilizzabile.

Un approccio migliore potrebbe essere il rilassamento degli equipaggiamenti: cercare di assegnare quante più frequenze possibile, magari assegnando un diverso livello di priorità sia alle varie celle, sia alle portanti della cella. Questo metodo richiede quindi informazioni ulteriori, senza le quali un algoritmo potrebbe privilegiare una cella di un'area quasi disabitata a una cella di un centro città.

La scelta di quali vincoli rilassare è quindi cruciale, anche perché comporta una valutazione dell'importanza relativa di un vincolo. Si è scelto quindi di rilassare tutti i vincoli interferenziali e di riuso, assegnando un peso differente ad ognuno; questo richiede una modifica alla definizione della funzione costo.

Sia $W(i,j,d)$ il costo derivato dal fatto che le celle i e j usano due portanti a distanza in banda d , definito per ogni $i, j \in \{1..N\}$, $d \in \{0,1,2\}$. Nel nostro caso, $W(i,j,d)$ è diverso da zero solo nei seguenti casi: per ogni $i, j \in \{1..N\}$ con $i = j$, oppure per ogni $j \in S_i$ o $j \in A_i^{400}$ con $d \in \{0,1\}$, oppure per ogni $j \in A_i^{200}$ con $d = 0$.

Sia NPSS il numero delle portanti sotto soglia della soluzione corrente, e WPSS il peso relativo per ognuna di esse.

La funzione costo diventa la seguente:

$$\begin{aligned}
 C(P) = & \sum_{i=1}^N \sum_{p=1}^M A_{i,p} RDS_{i,p} + \sum_{k=0}^2 \sum_{i=1}^N \sum_{p=1}^M A_{i,p} \sum_{\substack{j=0 \\ j \neq i}}^N W(i, j, k) A_{j,p \pm k} + \\
 & + \sum_{k=1}^2 \sum_{i=1}^N \sum_{p=1}^M A_{i,p} W(i, i, k) A_{i,p \pm k} + \sum_{i=1}^N (D_i - \sum_{p=1}^N A_{i,p}) W(i, i, 0) + NPSS \cdot WPSS \quad (2.11)
 \end{aligned}$$

Si noti la presenza nella funzione costo dell'interferenza complessiva, dei vincoli di riuso per sito e adiacenza, dei vincoli di riuso all'interno della stessa cella, del vincolo di unicit  delle assegnazioni, e del numero di portanti sotto soglia. Sono naturalmente possibili anche altre definizioni della funzione costo.

2.7 Una Formulazione alternativa

Sia (V, E) un grafo non orientato, i cui nodi rappresentino le portanti da assegnare. Si definisce un insieme C (*spettro*) di numeri consecutivi non negativi che rappresenti i possibili canali. Per ogni $v \in V$, sia $B_v \subseteq C$ l'insieme dei *canali bloccati*.

Si definiscono tre funzioni:

$$d : E \rightarrow Z_+ \quad (2.12)$$

$$c^{co} : E \rightarrow R_0 \quad (2.13)$$

$$c^{ad} : E \rightarrow R_0 \quad (2.14)$$

dove per ogni $\{v, w\} \in A$ $d(\{v, w\})$ indica la *distanza* da rispettare nell'assegnazione di v e di w , c^{co} e c^{ad} indicano l'*interferenza di co-canale* e l'*interferenza di canale adiacente*, rispettivamente, con $c^{ad} \leq c^{co}$; R_0 indica l'insieme dei numeri reali maggiori od uguali a zero e Z_+ l'insieme dei numeri interi maggiori di zero.

Una *rete* N (*carrier network*) pu  quindi essere rappresentata con una settupla

$$N = (V, A, C, \{B_v\}_{v \in V}, d, c^{co}, c^{ad}) \quad (2.15)$$

Si definisce *assegnazione di frequenza* per N una funzione

$$y : V \rightarrow C \quad (2.16)$$

Una assegnazione si definisce *ammissibile* se ad ogni portante $v \in V$ viene assegnato un canale $y(v) \in C \setminus B_v$, se per ogni $\{v, w\} \in A$ $|y(v) - y(w)| \geq d(\{v, w\})$.

Definizione

Data una rete N , si definisce *problema di assegnazione delle frequenze* (FAP) il seguente problema di ottimizzazione:

$$\min_y \sum_{\substack{\{v, w\} \in E: \\ y(v)=y(w)}} c^{\text{co}}(\{v, w\}) + \sum_{\substack{\{v, w\} \in E: \\ |y(v)-y(w)|=1}} c^{\text{co}}(\{v, w\}) \quad , \text{ con } y \text{ ammissibile.} \quad (2.17)$$

Si può dimostrare che anche tale problema è NP-completo usando ancora una volta una riduzione dal problema di colorazione minima dei grafi.

Capitolo III

EURISTICHE NON TABU-SEARCH

3.1 Introduzione

Molte euristiche sono composte da due fasi, inizializzazione (*starting heuristic*) e miglioramento (*improvement heuristic*). La prima si occupa di ricercare una soluzione di partenza, possibilmente ammissibile, mentre la seconda tenta di migliorarla. Nel seguito vengono descritte brevemente alcune tecniche usate in pratica, che estendono il caso banale di una assegnazione iniziale casuale.

Si fa riferimento al problema dell'assegnazione delle frequenze così come formulato nella Sezione 2.7

3.2 T-coloring

L'idea di base prende spunto da una procedura per la colorazione dei grafi sviluppata da D. Costa e D.Br  laz, e si concentra sull'ammissibilit   piuttosto che sul costo dell'assegnazione.

```
// inizializzazione
per ogni  $v \in V$ 
     $\text{satdeg}[v] = |B_v|$ 
     $\text{spadeg}[v] = \sum_{\{v,w\} \in E} d(\{v,w\})$ 

// assegnazione
 $U = V$ 
finch    $U \neq \emptyset$ 
    scegli  $v \in U$  tale che  $\text{satdeg}[v]$  sia massimo fra quelli che hanno massimo
         $\text{spadeg}[v]$ 
     $U = U \setminus \{v\}$ 
    sia  $y(v)$  il canale libero di minimo indice per  $v$ 
    se tale canale non esiste
        "nessuna assegnazione ammissibile trovata"
    per ogni  $w \in U$  con  $(v,w) \in A$ 
        aggiorna  $\text{satdeg}[w]$  ,  $\text{spadeg}[w]$ 
```

Figura 3.1 – T-Coloring

Per ogni portante non assegnata v , la variabile $\text{satdeg}[v]$ (*saturation degree*) tiene traccia di quanti canali non sono pi   liberi (   il numero di canali proibiti). La variabile $\text{spadeg}[v]$ (*spacing degree*) quantifica l'effetto che l'assegnazione ai vicini di v non ancora fissati provoca sulla assegnazione stessa di v . Se tale valore    grande,

si preferisce dare priorità a lei piuttosto che ai suoi vicini. Lo stesso vale per le portanti che hanno un alto grado di saturazione.

Con una implementazione che utilizzi un albero binario per stabilire quale sia la prossima portante da assegnare, il tempo di esecuzione risulta $O(|C||E| + |E|\log|V|)$, con una occupazione di memoria pari a $O(|C||V| + |E|)$.

3.3 Dual Greedy

L'idea di base di questo algoritmo (sviluppato da M. Hebermehl) è di cercare di identificare le combinazioni particolarmente svantaggiose di canali e portanti, e prendere delle precauzioni per poterne evitare.

Partendo da tutte le possibili combinazioni di portanti e canali, l'algoritmo elimina quelle che “sembrano” peggiori; per ciascuna portante l'ultima coppia rimanente viene usata effettivamente.

M e W sono dei parametri di controllo. Usando una struttura ad heap, si ottiene un tempo $O(|C|^2|V|\log(|C||V|) + |C|^2|E|)$ ed uno spazio $O(|C||V| + |E|)$. Questa tecnica non si è rivelata buona in pratica, soprattutto per il tempo di calcolo proibitivo.

//Inizializzazione

U = insieme di tutte le combinazioni permesse (v, f)

Per ogni (v, f) in U

Penalty $[v, f] = 0$

finché $U \neq \emptyset$

//Assegnazione

per ogni (v, f) in U che siano le uniche combinazioni in U in cui compaia v

per ogni combinazione (w, g) in U dove

$\{v, w\} \in E$ AND $((f = g$ AND $c^{co}(v, w) > 0)$ OR

$(|g - f| = 1$ AND $c(v, w) > 0)$ OR

$(|g - f| = d(v, w))$))

penalty $[w, g] = \text{penalty}[w, g] + W$

$y(v) = f$

$U = U \setminus \{(v, f)\}$

Se $U \neq \emptyset$ elimina da U la combinazione (v, f) di peso peggiore

$$\sum_{\substack{(w, f) \in V: \\ w \in \delta(v) \wedge d(v, w) = 0}} c^{co}(v, w) + \sum_{\substack{(w, f \pm 1) \in V: \\ w \in \delta(v) \wedge d(v, w) \leq 1}} c^{ad}(v, w) + \sum_{\substack{(w, g) \in V: \\ w \in \delta(v) \wedge |f - g| < d(v, w)}} M + \text{penalty}[v, f]$$

Figura 3.2 – Dual Greedy

3.4 DSATUR con costi

L'obiettivo perseguito da questo algoritmo (anch'esso formulato da D. Brélaz e D. Costa) è ottenere una assegnazione ammissibile usando un prefissato intervallo di canali a basso costo. A tal proposito si crea una matrice bidimensionale

di costo in cui le righe vengono indicizzate dalle portanti in V e le colonne dai canali in C , gli elementi rappresentano le differenti combinazioni possibili, eventualmente invalidati con un apposito marcatore (BLOCKED).

```
// Inizializzazione
Per ogni  $v \in V$ 
    Se  $f \in C \setminus B_v$   $\text{Cost}[v][f] = 0$  altrimenti  $\text{Cost}[v][f] = \text{BLOCKED}$ 
    Inserisci  $v$  nello heap con chiave  $|B_v|$ 
// Assegnazione
finché lo heap non è vuoto
    estrai la portante con chiave massima
    assegna a  $y(v)$  un canale non bloccato  $f$  di minimo costo nella riga  $\text{cost}[v]$ 
    aggiorna la matrice costo aggiungendo  $\Delta(v, f)$ 
    aggiorna le chiavi di tutte le portanti ancora nello heap
```

Figura 3.3 – DSATUR con costi

Si definisce “cattivo” (*bad*) per una portante un canale se il suo elemento corrispondente nella matrice è grande almeno quanto BAD , una altra costante opportunamente fissata. Ogni elemento dello heap corrisponde ad una portante non ancora assegnata.

Si ha inoltre che

$$\text{key}(v) = |B_v| \cdot BAD + \sum_{f \in C \setminus B_v} \min\{\text{cost}[v, f], BAD\} \quad (3.1)$$

Finché lo heap non è vuoto, si estrae la portante v con la chiave massima, e le si assegna il canale libero f a costo minore. Esiste la possibilità che così facendo si violi dei vincoli di separazione, ma in tal caso (e se BAD viene scelto sufficientemente elevato) questo succederebbe anche per gli altri canali. Al passo

successivo, vengono aggiornate tutte le righe indicizzate dalle portanti adiacenti a v , e tutte le chiavi nello heap (relative a portanti non ancora assegnate). Formalmente, alla matrice dei costi viene sommata una matrice $\Delta(v, f)$, così formata:

$$\Delta_{w,g}(v, f) = \begin{cases} \text{BAD} & \text{se } \{v, w\} \in E, g \in C \setminus B_v, |f - g| < d(v, w) \\ c^{\text{co}}(\{v, w\}) & \text{se } \{v, w\} \in E, d(v, w) = 0, f = g \in C \setminus B_v \\ c^{\text{ad}}(\{v, w\}) & \text{se } \{v, w\} \in E, d(v, w) \leq 1, f \pm 1 = g \in C \setminus B_v \\ 0 & \text{altrimenti} \end{cases} \quad (3.2)$$

Come prestazioni, si ottiene $O(|C||A| + |V|\log(|V|))$ per il tempo, e $|V| = O(|A|)$, $O(|C||V| + |A|)$ per lo spazio. Le varie implementazioni mostrano che la scelta della prima portante da assegnare assume una importanza cruciale.

3.5 1-OPT iterato

Questa euristica di miglioramento usa una struttura dati basata sul concetto di “distanza 1” fra assegnazioni: due assegnazioni si definiscono *distanti a 1* se una può essere ottenuta dall'altra variando il canale di una singola portante. Ad ogni singolo passo, nota la struttura di adiacenza, l'algoritmo determina una assegnazione y' adiacente di costo minore (o che viola meno vincoli) a partire dall'assegnazione attuale y .

Si definisce *assegnazione parziale* una funzione $y : A \rightarrow C$ con $A \subseteq V$

Si definisce *costo* di una coppia portante-canale (v, f) , $v \in V, f \in C$, rispetto ad una assegnazione parziale y su A , denotata con y_A , come

$$c(y_A; (v, f)) = \frac{1}{2} \sum_{\substack{w \in \delta(v) \cap A: \\ f = y_A(w)}} c^{\text{co}}(\{v, w\}) + \sum_{\substack{w \in \delta(v) \cap A: \\ |f - y_A(w)| = 1}} c^{\text{ad}}(\{v, w\}) \quad (3.3)$$

dove $\delta(v)$ indica l'insieme dei nodi incidenti a v in (V, E) .

Si definisce inammissibilità (*infeasibility*) di una coppia portante-canale (v, f) , $v \in V, f \in C$, rispetto a y_A , come

$$\text{infeas}(y_A; (v, f)) = \sum_{\substack{w \in \delta(v) \cap A: \\ |f - y_A(w)| \leq d(\{v, w\})}} 1 + \begin{cases} 1 & \text{se } f \in B_v \\ 0 & \text{altrimenti.} \end{cases} \quad (3.4)$$

```
//inizializzazione
y'=y
A'=A
ordina tutte le portanti v in V in ordine decrescente secondo
    infeas(y_A;(v,y(v))) e c(y_A;(v,y(v)))
metti davanti tutte le portanti non ancora assegnate

//passo i-esimo
per ogni portante v in V ordinata come prima
    se v ∉ A' aggiungi v ad A'
    attribuisce y'(v) al canale f tale che (v,f) sia minima fra tutte le coppie
        disponibili rispetto a infeas(y_A;(v,y(v))) e c(y_A;(v,y(v)))
```

Figura 3.4 – I-OPT iterato

Si definisce *passata* una sequenza di passi in cui ogni portante viene selezionata una volta. Le portanti vengono ordinate in senso decrescente secondo l'inammissibilità e il costo; da notare che l'algoritmo si può definire di discesa semplice e può restare "intrappolato" in minimi locali.

Se si usa uno heap di Fibonacci per determinare la prossima portante da considerare, le prestazioni sono $O(|C||E| + |V|\log(|V|))$ per il tempo, e $O(|C||V| + |E|)$ per lo spazio. Si potrebbe pensare ad una tecnica che permetta di considerare assegnazioni a distanza k , ma le implementazioni pratiche mostrano tempi di esecuzioni inaccettabili a livello pratico.

3.6 Flusso a costo minimo (FCM)

Questa euristica di miglioramento (di A. Eisenblätter) prova a migliorare una assegnazione y producendone un'altra y' sotto la seguente condizione:

per un prefissato ordinamento sull'insieme V :

$$\begin{array}{l} \forall (v, w) \in E \text{ con } v < w \text{ vale} \\ \quad (y(v) \leq y(w)) \Rightarrow (y'(v) \leq y'(w)) \\ \quad \text{AND} \\ \quad (y(v) > y(w)) \Rightarrow (y'(v) \geq y'(w)) \end{array} \quad (3.5)$$

Se una assegnazione di frequenza non usa tutti i canali disponibili, l'algoritmo tenta di produrre un assegnamento a costo inferiore utilizzando più canali. In pratica, ci si riconduce ad un problema di flusso a costo minimo, dove le funzioni d , c^{co} e c^{ad} vengono usate per valutare i coefficienti dei costi e le capacità degli archi. Questo algoritmo richiede uno spazio $O(|E|)$, ma è esponenziale nella taglia dell'ingresso se valutato nel caso peggiore.

Buoni sono i risultati di un uso combinato 1-OPT con MCF, grazie al vantaggio che offre quest'ultima nel “fuggire” dai minimi locali.

3.7 Rafforzamento del vincolo di separazione

Un approccio utile per il controllo dell'interferenza è escludere gli assegnamenti che causino grosse interferenze fra coppie di portanti. Per far questo, si introduce una *soglia* t , e la si applica nella costruzione dei vincoli di separazione. In particolare

$$d^t(v, w) = \begin{cases} \max\{1, d(v, w)\} & \text{se } (c^{co}(v, w) > t) \wedge (c^{ad}(v, w) < t), \\ \max\{2, d(v, w)\} & \text{se } c^{ad}(v, w) > t \\ d(v, w) & \text{altrimenti} \end{cases} \quad (3.6)$$

Nella costruzione di un nuovo problema, è possibile che non vi siano più soluzioni ammissibili; tuttavia, in pratica, si ottengono risultati migliori.

3.8 Un algoritmo enumerativo

L'ipotesi che sta alla base di questa euristica (sviluppata da C. Mannino e A. Sassano) è che esista un sottoinsieme di cardinalità ridotta rispetto al problema complessivo di celle le cui portanti siano fortemente “congestionate” (*core*), il quale riassume in sé la complessità totale del sistema. Se si riuscisse a trovare una soluzione buona per questo sottoinsieme, essa potrebbe essere “allargata” senza difficoltà alle restanti portanti.

La ricerca di una soluzione ammissibile in ciascun *cluster* di volta in volta selezionato, viene implementata attraverso un *Algoritmo Enumerativo*, capace di esplorare rapidamente un largo numero di alternative. Non trovare alcuna soluzione ammissibile per un cluster è condizione sufficiente per la non ammissibilità dell'intero problema.

Sia V l'insieme delle portanti
 W l'insieme delle portanti assegnate
 A_v l'insieme dei canali permessi a v

Algoritmo BBfreq()

1. se $V = W$
 - i. AMMISSIBILE = vero
 - ii. Stop
2. AMMISSIBILE = falso
3. Se $A_j = \emptyset$ per qualche $j \in V - W$ allora stop
4. Scegli una portante i che abbia il numero più basso di frequenze disponibili
5. Rimuovi da A_i tutte quelle frequenze che causerebbero su i o sulle portanti già assegnate un livello interferenziale inammissibile
6. finché $(A_i \neq \emptyset)$ AND (AMMISSIBILE = falso) AND (TIMEOFF = false)
 - i. se possibile, scegli la frequenza f di indice minimo con un livello interferenziale minore di $\alpha \cdot \text{SOGLIA}$, con $0 \leq \alpha \leq 1$; altrimenti quella con il livello interferenziale minimo.
 - ii. $A_i = A_i \setminus \{f\}$
 - iii. $W = W \cup \{i\}$
 - iv. $y(i) = f$;
 - v. Aggiorna il piano delle assegnazioni eliminando da tutte le portanti non assegnate le frequenze che non soddisfano i vincoli di riuso frequenziale, e i livelli di interferenza.
 - vi. Chiama ricorsivamente BBfreq()

Figura 3.5 – Schema enumerativo BBFreq

La procedura ricorsiva *BBFreq* viene invocata dopo ogni assegnazione, e lavora di volta in volta con istanze più piccole. Essa può essere interpretata come visita di un albero (Branchig Tree), i cui nodi rappresentano i sottoproblemi. Se la dimensione dell'istanza è troppo elevata, il numero dei nodi diviene troppo elevato così come il tempo per esplorarli. Si usa allora il *Restricted Backtracking*, che consiste nel visitare solo un sottoinsieme opportuno dell'albero originario, e il *Frequency Fixing* in cui si considera solo un sottoinsieme ridotto delle frequenze disponibili, sempre con la condizione che se il sottoproblema è irrisolvibile anche il problema originario lo è.

Cruciale inoltre diviene la ricerca di un *core* con un procedimento euristico efficiente (la ricerca del core stesso è un problema NP-Hard); in pratica si ricerca una zona che sia abbastanza piccola da poter essere analizzata con la procedura BBFreq ma che sia anche rappresentativa della complessità globale del problema originario.

Alcune implementazioni hanno dimostrato che tale algoritmo è particolarmente efficace solo con istanze di dimensione medie (fino ad un migliaio di celle).

Capitolo IV

LA TABU SEARCH

4.1 Introduzione

Nel campo scientifico ed ingegneristico esistono molti problemi di ottimizzazione intrinsecamente complessi, nei quali la ricerca della soluzione ottima mediante metodi esatti risulta improponibile, in quanto non sono noti algoritmi capaci di ottenerla senza impiegare un tempo di calcolo eccessivo. Diventa quindi fondamentale l'applicazione di algoritmi euristici che consentono, nella maggioranza dei casi, di ottenere soluzioni accettabili in tempi ragionevoli.

La Tabu Search è una metodologia euristica molto flessibile, che può essere applicata a un'estesa classe di problemi. Ciò che la contraddistingue è la capacità di superare i confini generalmente trattati come barriere - imposti dall'ammissibilità di una soluzione o dall'ottimalità locale, e di permettere (mediante la rimozione temporanea di alcuni vincoli o l'accettazione di soluzioni intermedie peggiorative) una valida esplorazione dello spazio delle soluzioni.

La sistemazione teorica della Tabu Search e la sua divulgazione nella forma con cui è conosciuta oggi è attribuibile principalmente al lavoro di Fred Glover, dell'Università di Colorado, a partire dal 1986. Da allora essa è stata applicata con successo a numerosi problemi di ottimizzazione, come ad esempio la minimizzazione quadratica, lo scheduling, l'instradamento ottimo, la colorazione dei grafi.

Al momento, non esiste alcuna prova formale (se non per alcune forme probabilistiche) della convergenza dell'algoritmo di Tabu Search.

4.2 Caratteristiche generali

Da un punto di vista formale, un problema di ottimizzazione può essere posto in questi termini: dato un insieme S di soluzioni ammissibili ed una funzione obiettivo $f : S \rightarrow \mathcal{R}$, trovare una soluzione i^* tale che $f(i^*)$ sia ottima rispetto a qualche criterio; generalmente, si richiede che sia $f(i^*) \leq f(i), \forall i \in S$.

Fra le tecniche di ottimizzazione, le tecniche *iterative* rivestono un ruolo di primaria importanza; infatti, per molti problemi non è noto alcun modo “analitico” per ottenere direttamente una soluzione “ottima”. Il passo generale di una procedura iterativa consiste nel costruire una nuova soluzione j a partire dalla soluzione corrente i e valutare se fermarsi o attuare un nuovo passo. I metodi di *ricerca in un intorno* (Neighborhood Search) sono un sottoinsieme delle tecniche iterative in cui si definisce, per ogni soluzione ammissibile i , un insieme $N(i)$ a cui deve appartenere la nuova soluzione j . A questa classe di tecniche fa parte il famoso *algoritmo di discesa* per la determinazione di una approssimazione di un minimo di una funzione reale f definita su un insieme S :

- a) Scegli una soluzione $i \in S$.
- b) Trova la miglior soluzione $j \in N(i)$ tale che $f(j) \leq f(k), \forall k \in N(i)$
- c) Se $f(j) \geq f(i)$ allora STOP, altrimenti $i = j$ e torna al passo b)

Questo metodo può arrestarsi ad un minimo locale anche molto peggiore del minimo globale.

La Tabu Search (TS) è una tecnica di discesa che prevede un uso efficiente della “memoria” per uscire da minimi locali alla ricerca di soluzioni migliori. A questo proposito viene tenuta traccia non solo delle informazioni locali (come il valore corrente della funzione obiettivo) ma anche di parametri che hanno caratterizzato il

processo di esplorazione, allo scopo di guidare, di volta in volta, la scelta della soluzione j . Per questo motivo, l'euristica è stata inserita fra le *tecniche a ricerca dinamica in un intorno* (*Dynamic Neighborhood Search*).

Il ruolo della “memoria” è quello di aiutare a definire un insieme

$$\overline{N}(i) \subseteq N(i) \quad (4.1)$$

detto *intorno ridotto* della soluzione i , entro il quale scegliere la soluzione successiva j , evitando così di dover estendere la ricerca a tutto l'intorno $N(i)$, che in molti casi risulta di taglia considerevole.

La semplice tecnica di discesa va incontro al rischio di rimanere “intrappolata” in un punto di minimo locale, che potrebbe essere anche molto distante dal punto di minimo globale. Un metodo per risolvere questo problema è quello di accettare, in alcune circostanze, delle soluzioni peggiorative. Questo accorgimento ha però lo svantaggio di rendere possibile che l'algoritmo cicli su un numero finito di soluzioni. Si supponga infatti che, al passo k -esimo, la soluzione corrente sia i , e non esista alcun $j \in \overline{N}(i)$ tale che $f(j) \leq f(i)$. Accettando una mossa peggiorativa, l'algoritmo sceglie come soluzione corrente (in base ad un qualche criterio) un opportuno p tale che $f(p) > f(i)$. Al passo $k+1$ -esimo, se la “vecchia” soluzione i appartiene all'intorno ridotto di p , e se non esiste alcuna soluzione $q \in \overline{N}(p)$ tale che $f(q) \leq f(p) < f(i)$, allora l'algoritmo individua i come soluzione migliorativa e compie un ciclo, riportandosi alla situazione del passo k -esimo.

Per ridurre la probabilità di compiere dei cicli, è necessario l'impiego da parte dell'algoritmo delle informazioni relative al “passato”, ovvero relative al percorso seguito per giungere alla soluzione corrente. Queste informazioni possono essere usate per proibire a priori la valutazione da parte dell'algoritmo delle soluzioni in $\overline{N}(i)$ visitate nei passi precedenti. La composizione dell'intorno ridotto di una soluzione i viene quindi a dipendere dall'iterazione k , oltre che dalla soluzione i .

D'ora in avanti, perciò, l'intorno ridotto di una soluzione verrà indicato con la notazione

$$\bar{N}(i, k) . \quad (4.2)$$

Formalizzando i concetti introdotti, si può riscrivere la generica iterazione dell'algoritmo di discesa modificandola in questo modo:

- a) scegli una soluzione iniziale $i \in S$. Poni $i^* = i$ e $k = 0$
- b) $k = k + 1$; genera il sottoinsieme $\bar{N}(i, k)$
- c) scegli il miglior $j \in \bar{N}(i, k)$ rispetto a f oppure a f_m
- d) se $f(j) \leq f(i^*)$ allora $i^* = j$
- e) fermati se si verifica una condizione di termine, altrimenti torna al passo b)

Per condizione di termine si intende una clausola che, se verificata, determina la fine del processo di ricerca, come ad esempio:

- $\bar{N}(m, k + 1) = \emptyset$
- k è maggiore del massimo numero di iterazioni consentite;
- il numero di iterazioni compiute dall'ultimo miglioramento di i^* è maggiore di un numero prestabilito.

Rimane da chiarire in che modo la conoscenza dell'itinerario percorso per arrivare alla soluzione corrente possa essere utilizzata convenientemente per modificare la struttura di $\bar{N}(i, k)$. Le soluzioni recentemente visitate vengono escluse dall'intorno ridotto della soluzione corrente; tale accorgimento ha l'effetto di impedire l'instaurazione di cicli. Quando una soluzione viene eliminata dall'insieme delle soluzioni candidate, si dice che essa è una soluzione "tabù" (*tabu solution*).

Si definisce lista tabù (tabu list) l'insieme T delle ultime soluzioni esplorate. Si ha allora:

$$I(i,k) = I(i) \setminus T(k). \quad (4.3)$$

È chiaro che il mantenimento di una *lista tabù* contenente le ultime $|T|$ soluzioni esplorate previene la formazione di cicli di lunghezza non superiore a $|T|$. Di qui il vantaggio di utilizzare liste tabù di notevoli dimensioni.

La gestione di una lista di soluzioni può essere molto dispendiosa dal punto di vista delle risorse di calcolo (memoria utilizzata, tempo impiegato), soprattutto nei casi in cui le soluzioni stesse siano oggetti molto complessi.

Si preferisce allora descrivere il processo di esplorazione delle soluzioni in termini delle “mosse” effettuate per passare da una soluzione ad un'altra. Con il termine generico di *mossa* si intende l'insieme delle operazioni elementari di modifica da compiere per ottenere una soluzione da un'altra.

Formalmente, si definisce $M(i)$ l'insieme delle mosse m che possono essere applicate a i per ottenere una nuova soluzione j , e si usa la notazione $j = i \oplus m$. Da ciò discende che $N(i) = \{j \mid \exists m \in M(i) \text{ con } j = i \oplus m\}$. In generale le mosse sono reversibili: per ogni m esiste una mossa m^{-1} tale che $(i \oplus m) \oplus m^{-1} = i$.

Così, invece di tenere traccia delle ultime $|T|$ soluzioni esplorate, è sufficiente memorizzare le ultime $|T|$ mosse effettuate, o le ultime $|T|$ mosse inverse associate a quelle effettivamente compiute. È chiaro che questa restrizione comporta una perdita di informazione, e non garantisce quindi a priori che non si possano instaurare cicli di lunghezza minore o uguale a $|T|$.

Per efficienza, può essere conveniente usare diverse liste T_r allo stesso tempo; un parametro di i o m può avere un valore tale da impedire una particolare mossa. Si possono quindi formulare le *condizioni tabù* come segue:

$$t_r(i, m) \in T_r, \quad r \in \{1, \dots, t\} \quad (4.4)$$

Una mossa è definita *tabù* se tutte le condizioni tabù sono soddisfatte.

4.2.1 Criterio di aspirazione

La perdita di informazione causata dalla semplificazione adottata (il fatto cioè di considerare le mosse invece che le soluzioni) determina anche un altro svantaggio: può succedere in alcuni casi che una soluzione m , non ancora esplorata, per la quale si avrebbe

$$f(m) < f(m^*) \tag{4.5}$$

divenga irraggiungibile a causa delle restrizioni introdotte dall'interdizione di alcune mosse. È necessario quindi introdurre un criterio che, in questi casi, consenta il rilassamento dei vincoli imposti dalla lista tabù.

Per ogni *mossa tabù* m che, applicata alla soluzione corrente i , darebbe luogo ad una soluzione j migliorativa, viene allora definito il *livello di aspirazione* e si impone che la mossa tabù possa essere effettuata quando il miglioramento del costo è maggiore di una soglia $A(i,m)$ assegnata

4.2.2 Intensificazione e Diversificazione

Nei paragrafi precedenti è stato illustrato come le informazioni sulle soluzioni esplorate nelle ultime $|T|$ iterazioni possano essere utilizzate per “proibire” l'effettuazione di mosse che potrebbero riportare a soluzioni già visitate, instaurando un ciclo. Si parla in questo caso di uso della memoria per *l'orientamento a corto termine*.

La conoscenza delle informazioni relative alle mosse precedenti consente inoltre di guidare il processo di ricerca anche in altri modi. Può infatti essere vantaggioso, in alcune circostanze, intensificare la ricerca in alcune regioni nelle quali è stata rilevata un'elevata concentrazione di punti di minimo locale. Questo obiettivo può essere raggiunto assegnando un'alta priorità alle soluzioni che presentano delle caratteristiche in comune con la soluzione corrente. È necessario

allora introdurre nella funzione obiettivo un termine aggiuntivo f_i che penalizzi le soluzioni “lontane” da quella corrente, e che abbia valore diverso da zero solo per quelle iterazioni per cui si ritiene opportuno effettuare una fase di *intensificazione*. In altre circostanze, tipicamente al termine di una fase di intensificazione, può risultare vantaggioso guidare la ricerca verso zone inesplorate dello spazio delle soluzioni. La *diversificazione* può essere forzata introducendo nella funzione obiettivo un termine aggiuntivo f_d che penalizzi le soluzioni che siano troppo “vicine”, cioè simili a quella corrente.

La funzione obiettivo modificata viene indicata con f_m e si ha per essa:

$$f_m = f + f_i + f_d \quad (4.6)$$

dove ai termini aggiuntivi f_i ed f_d sono associati opportuni pesi variabili che permettono all'algoritmo di alternare fasi di intensificazione a fasi di diversificazione.

Relativamente a queste due modalità di ricerca, si parla talora di uso della memoria per *l'orientamento a lungo termine*.

- a) Scegli una soluzione iniziale $i \in S$. Poni $i^* = i$ e $k = 0$.
- b) Poni $k = k + 1$ e genera un sottoinsieme $N(i, k)$ tale che almeno una delle condizioni $t_r(i, m) \in T_r$, $r \in \{1, \dots, t\}$ sia violata oppure valga almeno una fra le condizioni di aspirazione $a_r(i, m) \in A_r(i, m)$, $r \in \{1, \dots, a\}$.
- c) Scegli il miglior $j = (i \oplus m) \in \overline{N}(i, k)$ rispetto a f oppure a f_m e poni $i = j$
- d) Se $f(i) < f(i^*)$ allora poni $i^* = i$.
- e) Aggiorna le condizioni tabù e di aspirazione.
- f) fermati se si verifica una condizione di termine, altrimenti torna al passo b)

4.3 Reactive Tabu Search

La ricerca reattiva (*Reactive Search*), proposta da R. Battiti, propone l'introduzione di schemi di retroazione (feedback) nelle euristiche per l'ottimizzazione discreta. Si tratta di applicare una semplice forma di *apprendimento con rinforzo* al sistema, in modo che i parametri di ricerca possano dinamicamente modificarsi a seguito dell'influenza del comportamento del sistema negli eventi passati; si introduce, quindi, il concetto di esperienza nella ricerca. Spesso, infatti, i vari parametri (costi, lunghezza delle liste, ecc) possono variare secondo uno schema indipendente dall'istanza trattata; il che comporta che soluzioni accettabili possano essere ottenute dopo una accurata taratura dei valori tramite procedure "trial and error" (per tentativi). Risulta evidente, quindi, l'opportunità di utilizzare tecniche a parametri adattativi (*self tuning*) , per garantire una sufficiente flessibilità del metodo.

Si prenda, ad esempio, la lunghezza della lista tabù. Più grande è T , più lunga sarà la "traiettoria" che l euristica dovrà compiere prima di permettere che una soluzione possa essere ripresa in considerazione. Aumentare T quindi è un modo valido per attuare la fase di diversificazione. Tuttavia, non si può far crescere troppo T , in quanto potrebbe verificarsi il caso in cui nessun'altra soluzione sia possibile. La variazione adattativo del parametro T può quindi permettere notevoli miglioramenti della tecnica complessiva.

Capitolo V

DESCRIZIONE DELL'ALGORITMO

5.1 Generalità

Una qualsiasi applicazione che abbia come scopo la soluzione euristica del problema dell'assegnazione delle frequenze, così come è stato formulato nel capitolo 2, deve tenere in considerazione i seguenti aspetti operativi:

- Tempo di CPU
- Taglia delle istanze
- Qualità della soluzione
- Modularità

Tempo di CPU

Per tempo di CPU si intende l'intervallo di tempo necessario affinché l'algoritmo elabori una soluzione a partire da un'istanza d'ingresso. A parità delle caratteristiche hardware dell'elaboratore elettronico impiegato, e dell'istanza di ingresso, è ovviamente preferibile un tempo di CPU inferiore. Tipicamente sono tollerati tempi dell'ordine di poche ore. Spesso, inoltre, tale parametro dipende fortemente dalla taglia dell'istanza di ingresso; sarebbe preferibile un algoritmo che minimizzi questa dipendenza.

Taglia delle istanze

I dati di ingresso contengono tutte le informazioni necessarie per la determinazione della soluzione; per ogni cella vengono elencati il numero di portanti da assegnare (*equipaggiamento*), la lista delle celle adiacenti a 200kHz o 400KHz, la composizione del sito, la lista delle frequenze bloccate (prefissate o proibite), il livello di interferenza co-canale. La mole di questi dati aumenta circa quadraticamente rispetto al numero delle celle, soprattutto a causa della matrice delle interferenze. L'enorme sviluppo del mercato radiomobile comporta l'installazione di un gran numero di portanti e quindi di antenne. Le istanze di media dimensione, quindi, sono composte già oggi da migliaia di celle ciascuna equipaggiata mediamente con sei-sette portanti.

Risultano quindi necessari algoritmi e strutture dati che siano in grado di gestire efficientemente moli consistenti di dati, tenendo conto anche della possibile limitatezza delle risorse di calcolo (ad esempio della memoria).

Qualità della soluzione

Una soluzione al problema, per essere utilizzabile nella pratica, deve innanzi tutto rispettare i vincoli di natura tecnologica imposti (deve cioè essere ammissibile), e deve inoltre garantire al maggior numero di portanti un livello di interferenza inferiore ad una certa soglia.

Modularità

A causa delle diverse tipologie di istanze, invece di creare un unico algoritmo, è preferibile realizzare un insieme di funzioni di libreria (magari con la struttura modulare tipica della programmazione orientata agli oggetti), in modo che sia possibile adattare le procedure di elaborazione alle varie classi di istanze, ottimizzandone le prestazioni. Questo permette inoltre di testare ed integrare tecniche diverse fra loro.

Questi requisiti sono alla base delle scelte progettuali del pacchetto software sviluppato, e ne condizionano i vari aspetti. Vengono ora prese in esame le caratteristiche principali.

5.2 Dati di ingresso

- NUMERO DI CELLE DA PIANIFICARE N
- MATRICE DI INTERFERENZA M : è la struttura più vasta, in quanto raccoglie le informazioni dell'interferenza di co-canale fra due celle, intesa come rapporto (in decibel) tra la potenza del segnale utile e la potenza del segnale interferente. Formalmente, se la cella i e la cella j , con $i \neq j$, utilizzano la stessa portante p , allora l'elemento $M_{i,j}$ della matrice rappresenta l'interferenza che la cella j provoca sulla portante p della cella i . Si noti che la matrice *non* è simmetrica, cioè $M_{i,j} \neq M_{j,i}$, per i motivi visti in (2.3); può anzi succedere che i due termini differiscano e di molto.

Per motivi storici, sulla diagonale principale della matrice è presente il termine fittizio -99.00 che non è significativo.

La stessa matrice può essere utilizzata per ottenere l'interferenza da canale adiacente, in quanto si suppone che sia pari all'interferenza di co-canale divisa per il termine NFD (*Net Filter Discriminator*) del valore di 18 dB (pari a circa 63.1).

-99.00	13.00	14.50	18.00	40.50	41.00	...
14.50	-99.00	12.00	25.00	47.00	49.00	...
19.00	7.00	-99.00	16.50	40.00	37.00	...
11.00	21.00	19.50	-99.00	22.00	22.00	...
31.50	15.00	23.00	26.00	-99.00	16.50	...
41.00	20.50	14.50	14.50	10.50	-99.00	...
...	...	...	...	...	...	...

Figura 5.1 – Particolare della matrice delle interferenze

- LISTA DI EQUIPAGGIAMENTO DELLE CELLE: Vengono indicati rispettivamente
 1. numero identificativo della cella (all'interno del piano);
 2. sigla identificativa della cella (universale); celle dello stesso sito hanno lo stesso identificativo tranne, l'ultima cifra.
 3. un numero di servizio, influente ai fini dell'applicazione;
 4. la sigla UND oppure OVL che indica che le assegnazioni sono da considerarsi di *underlay* oppure *overlay*¹;
 5. il numero di portanti da equipaggiare;
 6. elenco delle frequenze bloccate perché già assegnate (-1 indica che tali frequenze non sono ancora state assegnate);
 7. la lista delle frequenze proibite, preceduta dal loro numero.

¹ Underlay (UND) si differenzia da Overlay (OVL) in quanto in quest'ultimo non sono attivi i vincoli di adiacenza a 200KHz e 400KHz.

1	2	3	4	5	6										7		
1	AA01D1	1	UND	7	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	0		
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
1364	PP24D2	1	UND	5	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	15	52	53
1365	PP24D3	1	UND	5	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	15	52	53
1366	PP29D1	1	UND	4	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	15	52	53
1367	PP29D2	1	UND	5	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	15	52	53
1368	PP29D3	1	UND	6	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	15	52	53
1369	PP30D1	1	UND	6	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	0		
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
1789	NN40D1	1	UND	6	14	21	33	-1	-1	-1	-1	-1	-1	-1	0		
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
2009	KK61D3	1	UND	3	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	0		

Figura 5.2 – Particolare del file degli equipaggiamenti

- LISTA DELLE ADIACENZE A 200kHz: per ogni cella si riportano le seguenti informazioni:

1. numero identificativo della cella;
2. sigla identificativa della cella;
3. lista delle celle adiacenti a 200 kHz alla stessa cella, geograficamente o radioelettricamente, preceduta dalla lunghezza della lista stessa.

1	AA01D1	4	2	3	4	5						
...	...	...	...	...	...	...	...	...	...	...	...	...
1364	PP24D2	6	861	929	1107	1261	1363	1365				
1365	PP24D3	6	861	964	1107	1261	1363	1364				
1366	PP29D1	7	891	893	1007	1008	1367	1368	1378			
1367	PP29D2	6	893	945	947	1366	1368	1379				
1368	PP29D3	9	843	947	977	1008	1366	1367	1378	1379	1380	
1369	PP30D1	6	728	730	1189	1190	1370	1371				
...	...	...	...	...	...	...	...	...	...	...	...	...
2009	KK61D3	4	378	469	2007	2008						

Figura 5.3 – Particolare del file delle adiacenze a 200 kHz

- LISTA DELLE ADIACENZE A 400kHz : composta di
 1. numero identificativo della cella;
 2. sigla identificativa della cella;
 3. lista delle celle adiacenti a 400 KHz alla stessa cella, geograficamente o radioelettricamente, preceduta dalla lunghezza della lista stessa

1	AA01D1	1	3				
...	...	...	...	...	...	...	...
1491	QQ30D1	3	1433	1493	1507		
1492	QQ30D2	1	1497				
1493	QQ31D1	3	1433	1491	1497		
1494	QQ31D2	3	1446	1497	1498		
1495	QQ32D1	1	1442				
1496	QQ32D2	0					
...	...	...	...	...	...	...	...
2009	KK61D3	0					

Figura 5.4 – Particolare del file delle adiacenze a 400 kHz

Si noti che la matrice di interferenza se relativa a problemi di dimensione medio-grande ($N > 1000$) può richiedere uno spazio di memoria eccessivo. Inoltre si può presupporre che tale matrice sia piuttosto sparsa, e cioè che una cella non interferisca significativamente con molte altre. È quindi preferibile, associare ad ogni cella una lista delle celle interferenti, con il loro rispettivo valore, soprattutto per le reti di grandi dimensioni.

5.3 Una nuova struttura dati di ingresso

Un algoritmo non può essere efficiente senza strutture dati efficienti. In seguito viene descritta una nuova struttura, proposta in alternativa alle strutture già esistenti, che si compone di due parti: un *header* che contiene informazioni generali riguardo l'istanza, (come il numero delle celle, l'intervallo dei canali possibili, ecc) e un corpo (*body*) contenente le informazioni relative alle singole celle. Queste ultime sono a loro volta raccolte in una sottostruttura, definita per ogni cella, divisa in:

- *informazioni specifiche della cella*, tra cui
 1. numero identificativo della cella
 2. sigla identificativa della cella
 3. numero di portanti da equipaggiare
 4. numero di celle dello stesso sito
 5. numero di celle adiacenti a 400kHz
 6. numero di celle adiacenti a 200kHz
 7. numero complessivo di celle interferenti
 8. numero di frequenze bloccate
 9. numero di frequenze proibite
- *lista degli interferenti*, che si distingue in
 1. lista delle celle dello stesso sito
 2. lista delle celle adiacenti a 400kHz

3. lista delle celle adiacenti a 200kHz
4. lista delle altre celle interferenti (ad un livello $> \lambda$).

Ogni elemento di una lista si compone del numero identificativo della cella interferente e del valore dell'interferenza. Nella lista 4 sono presenti solo quelle celle che presentano una interferenza maggiore di una soglia prestabilita λ . Al diminuire di λ aumenta l'occupazione di memoria; al limite per $\lambda=0$ sono presenti tutte le celle come interferenti. All'aumentare di λ , invece, diminuisce la precisione con cui si stima le interferenze totali di canale, in quanto i valori risultano troppo ottimistici e distanti dai valori reali. La scelta di λ rappresenta quindi un compromesso; soluzioni accettabili si ottengono con $\lambda = 10^{-4}$.

- *Lista (eventuale) di frequenze bloccate e/o proibite.*

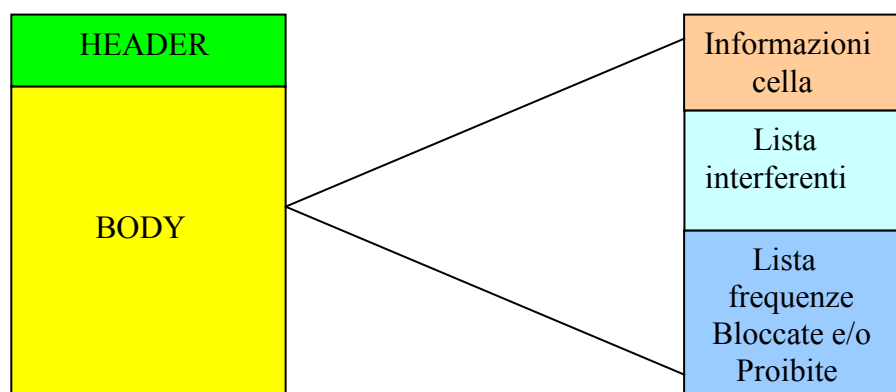


Figura 5.5 – Struttura di ingresso

5.4 Definizione del problema

Si fa riferimento alla descrizione del capitolo 2.6, e si considera il rilassamento dei vincoli tecnologico-realizzativi mediante penalizzazione nella funzione costo. Ciò permette di ottenere i seguenti vantaggi:

- Si ottiene sempre un piano di partenza iniziale
- Implementazioni più semplici ed efficienti
- Prestazioni elevate
- Nel caso non si riesca a trovare (o non esistano) soluzioni ammissibili, si può indicare la soluzione che viola meno vincoli.

Definizioni (richiamo)

$I_{i,j}$	interferenza subita dalla cella i ad opera della cella j se ad entrambe viene assegnato il medesimo canale ²
$A_{i,p} = 1$	se alla cella i viene assegnato il canale p
$A_{i,p} = 0$	se alla cella i non viene assegnato il canale p
$RDS_{i,p}$	interferenza complessiva a cui è soggetta la portante p della cella i
$EFF_{i,p}$	somma degli effetti causati dall'assegnazione stessa
$LD_{i,p} = RDS_{i,p} + EFF_{i,p}$	
I^{tot}	interferenza complessiva del piano
N	numero di celle del piano

² Spesso si usa il termine frequenza come sinonimo di canale, anche se il canale, essendo un segmento di banda, è in realtà un insieme di frequenze.

M numero di canali

Si ha quindi

$$RDS_{i,p} = \sum_{\substack{j=1 \\ j \neq i}}^N (I_{i,j} A_{j,p} + \frac{I_{i,j}}{NFD} A_{j,p-1} + \frac{I_{i,j}}{NFD} A_{j,p+1}) \quad (5.1)$$

$$EFF_{i,p} = \sum_{\substack{j=1 \\ j \neq i}}^N (I_{j,i} A_{j,p} + \frac{I_{j,i}}{NFD} A_{j,p-1} + \frac{I_{j,i}}{NFD} A_{j,p+1}) \quad (5.2)$$

$$I^{tot} = \sum_{i=1}^N \sum_{p=1}^M A_{i,p} \sum_{\substack{j=1 \\ j \neq i}}^N I_{i,j} \left(A_{j,p} + \frac{A_{j,p-1} + A_{j,p+1}}{NFD} \right) \quad (5.3)$$

5.5 Inizializzazione

La prima fase da affrontare è l'inizializzazione, cioè la ricerca di una soluzione (possibilmente ammissibile) di partenza. In letteratura sono state proposte numerose tecniche. La semplice assegnazione casuale, spesso permette, se iterata, una discreta esplorazione dello spazio delle soluzioni. Al crescere del rapporto portanti proibite su portanti permesse conviene invece utilizzare una tecnica che dia priorità di assegnazione a portanti che abbiano “minore possibilità di scelta”.

Sia Q_i , definito per ogni $i \in \{1, \dots, N\}$ l'insieme dei canali liberi della cella i . Per canali liberi si intendono quelle frequenze che, se assegnate alla cella i , non provocano nessuna violazione di vincolo. Per ogni $i \in \{1, \dots, N\}$ si ordinino le celle in ordine decrescente secondo il rapporto D_i su $|Q_i|$; utilizzando tale ordine si equipaggino le celle assegnando a ciascuna portante la frequenza di interferenza minima. Ad ogni assegnazione si aggiorna l'ordine delle celle (si potrebbe anche usare una struttura ad heap).

In questo modo, soprattutto nel caso di piani fortemente congestionati, è possibile che alcuni vincoli non siano rispettati. Il compito della prima fase è, tuttavia, solo di ottenere un piano di partenza, del quale non è richiesta una particolare efficienza (non è detto, infatti, che buone soluzioni di partenza conducano a risultati migliori).

5.6 Esplorazione

Una volta ottenuta una soluzione (piano) di partenza, inizia la fase migliorativa. L'idea di base è considerare tutti i piani a distanza 1 da quello attuale (che differiscano cioè per una sola assegnazione di portante). Se esiste un piano a costo inferiore, allora quello diviene il nuovo piano ed il procedimento viene iterato. In questo modo è però possibile rimanere “intrappolati” in un punto di minimo locale; si fa ricorso allora alla tecnica *Tabu Search*.

```

Pk = piano corrente
Cmin = C(Pk)
Pmin = Pk
TROVATO = falso
Per ogni i ∈ {1..N}
    Per ogni p ∈ {1..M}
        if Ai,p = 1 then
            begin
                Ai,p = 0
                Per ogni q ∈ {1..M}
                    if Ai,q = 0 then
                        begin
                            Ai,q = 1
                            Pk+1 = piano corrente
                            if C(Pk+1) < Cmin then
                                begin
                                    TROVATO = vero
                                    Cmin = C(Pk+1)
                                    Pmin = Pk+1
                                end
                            end
                            Ai,q = 0
                        end
                    end
                Ai,p = 1
            end
        end
    if TROVATO = falso then
        Pk+1 = Pk
    else
        Pk+1 = Pmin

```

Figura 5.6 - Algoritmo di discesa semplice, passo k-esimo

5.7 Idea di base e suo miglioramento

L'idea di base è usare la definizione di costo di un piano e ricalcolarlo ad ogni passo per ogni assegnazione (vedi fig 5.6); ad ogni mossa si valutano tutti i piani a distanza uno da quello corrente, e si sceglie quello di costo minimo.

L'uso della Tabu Search (vedi fig 5.7) previene il blocco da minimo locale, in quanto si accettano anche mosse peggiorative. Per prevenire l'insorgere di cicli, si usa una lista di mosse effettuate; in definitiva una mossa è permessa se non appartiene alla *Lista Tabù* oppure se verifica il criterio di aspirazione (cfr 3.). La lista Tabù è una struttura (in genere una lista circolare) che permette di ricostruire la storia delle mosse effettuate.

Definizioni

Sia Π l'insieme di tutti i piani possibili.

Siano $P_A, P_B \in \Pi$, con $P_A \neq P_B$, si definisce *mossa* m la coppia (P_A, P_B) , e si usa la seguente notazione: $P_A \xrightarrow{m} P_B$, che si legge: P_B si ottiene da P_A tramite la mossa m .

Si definisce *intorno ridotto* di P_A l'insieme $U_A = \{m : P_A \xrightarrow{m} P_B : P_A \text{ disti } l \text{ da } P_B\}$

Si definisce *Lista Tabù al passo k* l'insieme T_k delle coppie (i, p) , $i \in \{1..N\}$, $p \in \{1..M\}$, tale che P_k sia il piano corrente e per ogni coppia esista una successione di mosse $\{m_1, m_2, \dots, m_r\}$ con $r \leq |T_k|$ tale per cui :

$$P_k \xrightarrow{m_1} P_{k-1} \xrightarrow{m_2} \dots \xrightarrow{m_r} P_{k-r}, m_h \in U_{k-h+1} \text{ per ogni } h \in \{1..r\},$$

$$A_{i,p} = 1 \text{ in } P_{k-r}, A_{i,p} = 0 \text{ in } P_{k-r+1}.$$

Sia $K_{i,j} = I_{i,j} + I_{j,i}$ con $j, i \in \{1..N\}$. Ovviamente la matrice K è simmetrica.

Si definisce *interferente* a i una coppia $(j, K_{i,j})$, con $I_{i,j} > \lambda$ e/o $I_{j,i} > \lambda$.

Sia definisce Γ_i l'insieme degli interferenti relativi alla cella i , per ogni $i \in \{1..N\}$

Si definisce $Z_i = \{p : p \in \{1..M\}, A_{i,p} = 1\}$.

```

 $C_{curr} = +\infty$ 
Per ogni  $i \in \{1..N\}$ 
    Per ogni  $p \in \{1..M\}$ 
        if  $A_{i,p} = 1$  then
            begin
                 $A_{i,p} = 0$ 
                Per ogni  $q \in \{1..M\}$ 
                    if  $(A_{i,q} = 0)$  then
                        begin
                             $A_{i,q} = 1$ 
                             $P = \text{piano corrente}$ 
                            if  $(C(P) < C_{curr})$  and  $((i,q) \notin T_k \text{ or } C(P) < C_{best})$  then
                                begin
                                     $C_{curr} = C(P)$ 
                                     $P_{curr} = P$ 
                                     $index = i$ 
                                     $oldf = p$ 
                                end
                             $A_{i,q} = 0$ 
                        end
                    end
                 $A_{i,p} = 1$ 
            end
         $Piano\ corrente = P_{curr}$ 
        Se  $C_{best} < C_{curr}$ 
             $C_{best} = C_{curr}$ 
             $P_{best} = P_{curr}$ 
             $T_{k+1} = (T_k \setminus T_{k-|T|+1}) \cup \{(index, oldf)\}$ 

```

Figura 5.7 - Algoritmo con Tabu Search, passo k-esimo

Per semplicità si è supposto $|T|$ costante al variare di k ; implementazioni più complesse suggeriscono di variare anche tale parametro a seconda dello stato attuale e della “storia” recente, applicando in tale modo le fasi di diversificazione e intensificazione (cfr). A tale tecnica ci si riferisce come *Reactive Tabu Search*.

Il passo seguente consiste nel creare e utilizzare strutture dati efficienti allo scopo di ottimizzare la fase di miglioramento. Un primo risultato si ottiene con il seguente teorema.

Teorema 5.1

Siano P_1 e P_2 due piani di frequenza che differiscano solo nell’ assegnazione di una frequenza in una cella h (siano cioè a distanza 1). In particolare sia per P_1 : $A_{h,f} = 1$, $A_{h,g} = 0$ e per P_2 $A_{h,f} = 0$, $A_{h,g} = 1$. Si supponga trascurabile, per semplicità, l’interferenza da canale adiacente (il teorema generale viene riportato in appendice, in quanto i passaggi sono alquanto laboriosi).

Allora vale

$$I_{P_2}^{\text{tot}} = I_{P_1}^{\text{tot}} - LD_{h,f} + LD_{h,g} \quad (5.4)$$

con $LD_{i,p} = RDS_{i,p} + EFF_{i,p}$.

Dimostrazione

Trascurando l’interferenza da canale adiacente si ha che:

$$RDS_{i,p} = \sum_{\substack{j=1 \\ j \neq i}}^N I_{i,j} A_{j,p} \quad (5.5)$$

$$EFF_{i,p} = \sum_{\substack{j=1 \\ j \neq i}}^N I_{j,i} A_{j,p} \quad (5.6)$$

$$I^{\text{tot}} = \sum_{i=1}^N \sum_{p=1}^M A_{i,p} \sum_{\substack{j=1 \\ j \neq i}}^N I_{i,j} (A_{j,p}) \quad (5.7)$$

Esplicitando l'espressione (5.6) rispetto ad una cella h con i suoi canali f e g si ottiene:

$$\begin{aligned}
 I^{\text{tot}} &= \sum_{\substack{i=1 \\ i \neq h}}^N \sum_{\substack{p=1 \\ p \neq h}}^M A_{i,p} \sum_{\substack{j=1 \\ j \neq i}}^N I_{i,j} A_{j,p} + \sum_{p=1}^M A_{h,p} \sum_{\substack{j=1 \\ j \neq h}}^N I_{h,j} A_{j,p} \\
 I^{\text{tot}} &= \sum_{\substack{i=1 \\ i \neq h}}^N \sum_{\substack{p=1 \\ p \neq h}}^M A_{i,p} \sum_{\substack{j=1 \\ j \neq i, h}}^N I_{i,j} A_{j,p} + \sum_{\substack{i=1 \\ i \neq h}}^N \sum_{\substack{p=1 \\ p \neq h}}^M A_{i,p} I_{i,h} A_{h,p} + \sum_{p=1}^M A_{h,p} \sum_{\substack{j=1 \\ j \neq h}}^N I_{h,j} A_{j,p} \\
 I^{\text{tot}} &= \sum_{\substack{i=1 \\ i \neq h}}^N \sum_{\substack{p=1 \\ p \neq h}}^M A_{i,p} \sum_{\substack{j=1 \\ j \neq i, h}}^N I_{i,j} A_{j,p} + \sum_{\substack{i=1 \\ i \neq h}}^N \sum_{\substack{p=1 \\ p \neq f, g}}^M A_{i,p} I_{i,h} A_{h,p} + \sum_{\substack{p=1 \\ p \neq f, g}}^M A_{h,p} \sum_{\substack{j=1 \\ j \neq h}}^N I_{h,j} A_{j,p} + \\
 &\quad + \sum_{\substack{i=1 \\ i \neq h}}^N A_{i,f} I_{i,h} A_{h,f} + \sum_{\substack{i=1 \\ i \neq h}}^N A_{i,g} I_{i,h} A_{h,g} + A_{h,f} \sum_{\substack{j=1 \\ j \neq h}}^N I_{h,j} A_{j,f} + A_{h,g} \sum_{\substack{j=1 \\ j \neq h}}^N I_{h,j} A_{j,g}
 \end{aligned}$$

e quindi :

$$I^{\text{tot}} = K + A_{h,f} \text{EFF}_{h,f} + A_{h,g} \text{EFF}_{h,g} + A_{h,f} \text{RDS}_{h,f} + A_{h,g} \text{RDS}_{h,g} \quad (5.8)$$

$$\text{con } K = \sum_{\substack{i=1 \\ i \neq h}}^N \sum_{\substack{p=1 \\ p \neq h}}^M A_{i,p} \sum_{\substack{j=1 \\ j \neq i, h}}^N I_{i,j} A_{j,p} + \sum_{\substack{i=1 \\ i \neq h}}^N \sum_{\substack{p=1 \\ p \neq f, g}}^M A_{i,p} I_{i,h} A_{h,p} + \sum_{\substack{p=1 \\ p \neq f, g}}^M A_{h,p} \sum_{\substack{j=1 \\ j \neq h}}^N I_{h,j} A_{j,p}$$

Nel nostro caso, quindi:

$$I_{P_1}^{\text{tot}} = K + LD_{h,f} \quad I_{P_2}^{\text{tot}} = K + LD_{h,g}$$

e in definitiva

$$I_{P_2}^{\text{tot}} = I_{P_1}^{\text{tot}} - LD_{h,f} + LD_{h,g} \quad \square$$

Tale teorema può essere sfruttato introducendo la matrice LD , che ha dimensione $N \times M$. In particolare, se due piani sono distanti 1 perché nella cella h essi usano il canale f o g , la differenza dei loro costi vale $LD_{h,f} - LD_{h,g}$. La mossa ottimale ad ogni passo può essere quindi determinata come:

$$(h^*, f^*, g^*) = \arg \max_{h, f, g} (LD_{h, f} - LD_{h, g}), \quad (5.9)$$

con $A_{h, f} = 1$ e $A_{h, g} = 0$, con $h \in \{1..N\}$; $f, g \in \{1..M\}$

Molto utile risulta l'uso di uno heap per contenere il massimo miglioramento relativo ad ogni cella.

Si definisce, per ogni $h \in \{1..N\}$, una terna $\Delta_h = (w, p, q)$ con $w \in \mathbb{R}_0$; $p, q \in \{1..M\}$; con

$$(p, q) = \arg \max_{f, g} (LD_{h, f} - LD_{h, g}), \quad (5.10)$$

con $A_{h, f} = 1$ e $A_{h, g} = 0$, $f, g \in \{1..M\}$ e $w = \max_{f, g} (LD_{h, f} - LD_{h, g})$.

Lo heap userà come chiave il termine w .

5.8 Effetto del rilassamento dei vincoli

Il teorema vale ancora nell'ipotesi di rilassamento dei vincoli a patto di modificare la definizione della matrice LD , per tenere conto della possibile violazione.

Sia $W(i, j, d)$ il costo derivato dal fatto che le celle i e j usano due portanti a distanza in banda d , definito per ogni $i, j \in \{1..N\}$, $d \in \{0, 1, 2\}$. Nel nostro caso, $W(i, j, d)$ è diverso da zero solo nei seguenti casi: per ogni $i, j \in \{1..N\}$ con $i = j$, oppure per ogni $j \in S_i$ o $j \in A_i^{400}$ con $d \in \{0, 1\}$, oppure per ogni $j \in A_i^{200}$ con $d = 0$.

Sia NPSS il numero delle portanti sotto soglia della soluzione corrente, e WPSS il peso relativo per ognuna di esse.

La funzione costo diventa la seguente:

$$\begin{aligned}
 C(P) = & \sum_{i=1}^N \sum_{p=1}^M A_{i,p} RDS_{i,p} + \sum_{k=0}^2 \sum_{i=1}^N \sum_{p=1}^M A_{i,p} \sum_{\substack{j=0 \\ j \neq i}}^N W(i, j, k) A_{j,p \pm k} + \\
 & + \sum_{k=1}^2 \sum_{i=1}^N \sum_{p=1}^M A_{i,p} W(i, i, k) A_{i,p \pm k} + \sum_{i=1}^N (D_i - \sum_{p=1}^M A_{i,p}) W(i, i, 0) + NPSS \cdot WPSS \quad (5.11)
 \end{aligned}$$

Tutto ciò porta a considerare il peso di una violazione come parte dell'interferenza stessa, un aspetto fondamentale di tutta l'euristica. Ecco che si può definire l'interferenza modificata $I'_{ij} = I_{ij} + W_{ij}$ e considerare questa come interferenza nei calcoli; la matrice dei pesi combinati LD, la matrice delle interferenze RDS e degli effetti EFF vengono quindi ridefinite di conseguenza, mantenendo la stessa struttura; mentre RDS perde il significato intrinseco di interferenza complessiva (spettro).

Infine, per tenere conto dei vincoli di riuso all'interno della stessa cella, si deve procedere, ad ogni assegnazione o rimozione della stessa, all'incremento o decremento relativo del valore di LD nelle frequenze adiacenti. In formule:

sia p la frequenza assegnata ad una portante della cella h .

Allora per ogni $k \in \{-2, -1, 0, 1, 2\}$ $LD_{h,p \pm k} = LD_{h,p} \pm W_{h,h \pm k}$.

Questo garantisce che se la cella avesse ancora portanti da equipaggiare, ad esse può essere assegnata una frequenza f con un incremento nel costo del piano di $LD_{h,f}$. Lo stesso discorso vale nella ricerca dell'assegnazione migliore in una cella: prima di stabilire quale sia è necessario rimuovere i costi dei vincoli di cella alla portante da sostituire e alle sue quattro adiacenti.

5.9 L'algoritmo complessivo

L'“accensione” (assegnazione) e lo “spegnimento” (revoca) di una frequenza ad una portante di una cella è rappresentata dalle due procedure TurnON e TurnOFF rispettivamente; si noti la rapidità con cui permettono di aggiornare la matrice LD.

L'algoritmo inizia calcolando la matrice dei pesi combinati LD accendendo tutte le frequenze coinvolte nel piano iniziale; per la linearità degli effetti, ogni elemento $LD_{h,p}$ rappresenta la variazione del costo del piano (tolti i pesi dovuti ai vincoli di cella). Come completamento della fase di inizializzazione, si costruisce uno heap per mantenere informazioni relative a quali frequenze e in quali celle conviene scambiare.

Procedura TurnON

Input : $i \in \{1..N\}$, $p \in \{1..M\}$

$$A_{i,p} = 1$$

Per ogni $(j, K_{i,j}) \in \Gamma_i$

$$LD_{j,p} = LD_{j,p} + K_{i,j}$$

Figura 5.8 – Procedura TurnON

Procedura TurnOFF

Input : $i \in \{1..N\}$, $p \in \{1..M\}$

$$A_{i,p} = 0$$

Per ogni $(j, K_{i,j}) \in \Gamma_i$

$$LD_{j,p} = LD_{j,p} - K_{i,j}$$

Figura 5.9 – Procedura TurnOFF

La fase di miglioramento ed esplorazione opera sul complesso delle celle e di volta in volta sceglie lo scambio migliore. Ogni passo richiede tempo $O(1)$ per determinare lo scambio migliore, $O(\bar{\Gamma})$ per l'aggiornamento di LD, dove $\bar{\Gamma}$ è il numero medio di interferenti (tipicamente $\bar{\Gamma}$ può ritenere costante); e tempo $O(\bar{\Gamma} \log N)$ per l'aggiornamento dello heap (questa è una stima molto pessimistica, visto che raramente una variazione dell'interferenza dovuta ad un'unica portante modifica la decisione di uno scambio migliore, se non nelle celle adiacenti, le quali sono a tutti gli effetti mediamente di numero costante). In definitiva una iterazione dipende da N in maniera molto blanda: nei casi pratici, per reti con oltre un migliaio di celle, la singola iterazione richiede un tempo praticamente costante ed indipendente dal numero di celle.

$\bar{\Gamma}$ è un parametro di ingresso e dipende fortemente dal parametro λ descritto nel paragrafo (5.3). Questo significa che aumentando λ non solo si ha una occupazione minore di memoria, ma anche una riduzione diretta del tempo di calcolo e della dipendenza da N .

L'iterazione termina con l'aggiornamento della lista tabù, secondo le modalità già descritte.

Inizializzazione

scegli un piano di frequenza iniziale

$P_{\text{curr}} = P_{\text{best}} = \text{piano iniziale}$

$C_{\text{curr}} = C_{\text{best}} = \text{costo del piano iniziale}$

per ogni $i \in \{1..N\}$, **per ogni** $p \in \{1..M\}$: $LD_{i,p} = 0$

per ogni $i \in \{1..N\}$, **per ogni** $p \in Z_i$: $\text{TurnON}(i,p)$

Costruisci lo heap determinando Δ_i per ogni $i \in \{1..N\}$

Passo k-esimo dell'esplorazione

Leggi l'elemento $\Delta_h = (w_h, f_h, g_h)$ massimo dallo heap

$\text{TurnOFF}(h, f_h)$

$\text{TurnON}(h, g_h)$

$C_{\text{curr}} = C_{\text{curr}} - \Delta_h$

if $C_{\text{curr}} < C_{\text{best}}$ **then**

begin

$C_{\text{best}} = C_{\text{curr}}$

$P_{\text{best}} = P_{\text{curr}}$

end

Aggiorna l'elemento Δ_h dello heap

per ogni j tale che $(j, K_{h,j}) \in \Gamma_h$

$\Delta_j' = \Delta_j$

calcola $\Delta_j = (w_j, f_j, g_j)$ con la condizione

$((f_j, g_j), \notin T_k$ a meno che $C_{\text{curr}} - w_j < C_{\text{best}}$)

if $\Delta_j' \neq \Delta_j$ **then** aggiorna l'elemento Δ_j dello heap

$T_{k+1} = (T_k \setminus T_{k-|T|+1}) \cup \{(h, f_h)\}$

Fig 5.10 – Algoritmo completo, inizializzazione e iterazione k-esima

5.10 Uso del Core

Si supponga che, nel suo procedere, l'euristica arrivi ad un punto di minimo locale. Si supponga inoltre che nel piano esista una zona “centrale” molto congestionata, configurata in modo che esistano dei vincoli violati ma che non possano essere eliminati se non con una serie di mosse inizialmente peggiorative, inizialmente introducenti a loro volta ulteriori violazioni di vincoli; si consideri, inoltre, l'esistenza di una vasta zona “periferica” a bassa interferenza in cui alle portanti possano essere assegnate moltissime frequenze senza che il costo complessivo vari sostanzialmente. In questa situazione, l'euristica potrebbe essere indotta a considerare solo il secondo tipo di celle, in quanto esse garantiscono le minori variazioni negative del costo; sarebbe conveniente, invece, forzare l'algoritmo a considerare l'insieme di celle più congestionato, in modo da ottenere un effettivo miglioramento.

A questa necessità, risponde la *tecnica del Core*, simile a quella introdotta da Mannino e Sassano [8], che consiste nel selezionare un insieme di celle e concentrare su questo la Tabu Search. A questo proposito si è scelto di selezionare la cella con la portante a costo peggiore, e di formare un insieme costituito da questa cella e le sue dirette adiacenti, e le adiacenti di queste ultime e così via ricorsivamente. Il numero totale di celle o il livello di ricorsione viene controllato tramite un parametro; buoni risultati si ottengono con *core* di circa cinquanta celle. Questa tecnica si è rivelata particolarmente efficace nel risolvere violazioni di vincoli.

Una semplice tecnica di miglioramento, quindi, consiste nell'alternare fasi di ricerca estesa a tutte le celle a fasi di ricerca relativi a determinati *core*. È bene inoltre proibire temporaneamente la ricerca su un core appena visitato: questo permette di non ripetere inutilmente l'analisi e di concentrarsi anche su altre zone del piano.

5.11 Vincolo di soglia

Il vincolo di soglia, cioè il fatto che il maggior numero possibile di portanti debba avere il valore di interferenza complessiva minore di una certa soglia (tipicamente a 9dB), è un vincolo “soft” ritenuto meno importante (molti autori addirittura non lo considerano, e indicano l’interferenza media come unico parametro di qualità). Avendo quindi in mente come scopo principale il miglioramento medio del piano, e tenendo in considerazione che piani “mediamente” buoni lo sono molto spesso anche dal punto di vista del vincolo di soglia, si è scelto di non gestire esplicitamente durante la Tabu Search tale vincolo, ma di rafforzare la naturale tendenza dell’euristica a distribuire le interferenze, nel limite del possibile, in misura bilanciata.

Si vuole infatti evitare che un piano in cui coesistano molte portanti con interferenza pressoché nulla e alcune con una interferenza intollerabile venga preferito ad un piano in cui a tutte le portanti venga assegnata una interferenza accettabile, solo perché il primo ha un costo minore del secondo (un caso patologico di questo tipo è in realtà poco probabile, in quanto le interferenze sono distribuite in uno spazio geografico reale)

A tal fine, si utilizza un contatore di soglie violate e si considera un piano degno di essere memorizzato se il suo costo è minore della migliore soluzione corrente (a parità del numero di soglie violate) oppure se viola un numero inferiore di soglie (ovviamente a parità di violazioni degli altri vincoli).

5.12 Analisi dei risultati

Di seguito vengono riportati alcuni risultati ottenuti (contrassegnati da una C) con delle istanze fornite da CSELT³ per un confronto qualitativo e quantitativo con

³ CSELT (Centro Studi e Laboratori di Telecomunicazioni) con sede a Torino, è il centro ricerca del gruppo STET.

l'algoritmo enumerativo descritto nel capitolo III, come implementato dagli autori, Mannino e Sassano [8]. In Figura 5.11 sono riportati i calcolatori.

Calcolatore BBfreq	Calcolatore Tesi
• SUN ULTRA 250 • RAM 256 Mb	• PENTIUM II/400 • RAM 128 Mb

Figura 5.11 – Caratteristiche tecniche principali degli elaboratori usati nei risultati rispettivi

Per ogni istanza analizzata, in Tabella 5.1 vengono riportate il nome identificativo, il numero delle celle e il numero complessivo delle portanti (equipaggiamenti); le colonne contrassegnate da BB_B e BB_A si riferiscono a risultati ottenuti da due versioni dell'algoritmo enumerativo (BBFreqB e BBFreqA), così come descritto in [12]. Si riportano il tempo complessivo di esecuzione e il rapporto segnale disturbo peggiore trovato: una soluzione è tanto migliore quanto più alto è il peggior rapporto segnale disturbo (RSD).

Inoltre, in Tabella 5.2 si fornisce il rapporto segnale disturbo medio e il numero di portanti sotto soglia per l'algoritmo enumerativo (versione B) e per l'algoritmo usato nella tesi.

Confrontando le Tabelle 5.1 e 5.2 si può notare il sensibile miglioramento relativo alla qualità delle soluzioni e ai tempi di esecuzione.

In Tabella 5.3 si riportano risultati relativi a reti di grandi dimensioni: le prime due si riferiscono a casi reali, mentre l'ultima è un caso fittizio creato giustapponendo quindici volte la penultima istanza, che rappresenta l'istanza reale più grande e complessa fornita dai laboratori dello CSELT. Ovviamente, al crescere delle dimensioni del problema, il numero delle iterazioni necessarie deve crescere conseguentemente per garantire il medesimo livello qualitativo; tuttavia, si può notare la costanza del tempo necessario per eseguire ogni mossa.

Sicuramente, la velocità nell'esplorazione ha consentito l'analisi di un numero maggiore di piani di frequenza; inoltre il fatto che la mossa migliore sia scelta su un intorno ridotto con cardinalità poco minore dell'intorno completo

Nome	N° Celle	N° port.	Tempo (sec)			RSD massimo (dB)		
			BB_A	BB_B	C	BB_A	BB_B	C
Aa11	139	328	23.4	54.7	2.2	9.40	9.08	10.05
Aa12	160	375	65.4	68.7	2.3	8.88	8.64	9.04
Aa13	181	425	74.3	75.6	2.3	8.64	8.94	9.02
Aa14	200	470	65.4	120.3	2.3	8.36	8.55	9.01
Bb10	121	316	12.6	23.0	3.8	9.82	9.82	12.08
Bb11	162	430	14.9	23.5	4.6	9.22	9.22	9.93
Bb12	179	473	76.2	24.0	4.9	9.19	9.22	9.17
Bb13	203	536	72.0	31.7	5.4	8.36	8.54	9.14
Cc13	140	352	10.5	29.0	4.2	9.84	9.84	9.15
Cc14	161	396	27.0	24.6	4.0	8.55	8.55	9.02
Cc15	181	444	112.0	37.9	4.2	8.36	8.36	9.14
Cc16	200	486	95.4	32.7	4.4	9.07	9.07	9.26
Dd11	141	370	10.6	23.7	4.4	9.84	9.84	14.16
Dd12	159	419	15.7	29.8	5.0	9.04	9.04	14.53
Dd13	181	476	29.5	67.9	5.5	9.84	9.84	13.67
Dd14	201	525	76.0	30.4	5.9	8.91	8.91	12.92
Ee12	139	349	15.4	22.6	2.5	9.07	9.84	9.18
Ee13	161	408	32.0	22.0	1.9	8.91	8.65	9.26
Ee14	179	457	65.0	25.4	2.0	8.28	8.28	8.45
Ee15	201	514	160.4	100.0	2.1	8.65	8.28	8.43
Ff10	141	314	74.6	102.5	4.3	9.89	9.89	9.09
Ff11	161	357	63.7	63.7	4.7	8.87	8.87	8.95
Ff12	182	404	65.4	66.9	5.0	8.87	8.87	8.93
Ff13	201	446	140.9	250.9	5.1	8.65	8.65	8.74

Tabella 5.1- Risultati relativi a 27 frequenze disponibili e 15000 mosse

suggerisce che la probabilità di effettuare la scelta “migliore” sia alta.

La scelta di utilizzare la tecnica Tabu Search a vincoli rilassati si è quindi rivelata particolarmente efficace, in quanto è sempre possibile avere in tempi rapidi un buon piano; la tecnica del core è inoltre efficace nel caso di violazione di vincoli

tecnologici, in quanto si è sempre dimostrata in grado di ripristinare l'ammissibilità del piano.

Da notare che in qualche caso non è stato possibile garantire il livello minimo di 9 dB a tutte le portanti (soprattutto nelle istanze di grande dimensione) a causa dell'elevato numero di celle adiacenti e dell'elevata richiesta di equipaggiamento.

Nome	Ncelle	Nport.	BBFreqB				Nuovo			
			Tempo (sec.)	Interf. Max.	Interf. media	Npss	Tempo (sec.)	Interf. Max.	Interf. Media	Npss
Bb13	203	536	31.7	8.54	14.42	7	5.7	9.14	16.80	0
Cc14	161	396	24.6	8.55	13.80	14	4.0	9.02	15.98	0
Cc15	181	444	37.9	8.36	12.87	21	4.2	9.14	15.63	0
Ee13	161	408	22	8.65	12.73	24	1.9	9.26	15.84	0
Ee14	179	457	25.4	8.28	13.07	17	2.0	8.45	14.70	2
Ee15	201	514	100	8.28	12.23	50	2.1	8.43	14.26	3
Ff11	161	357	63.7	8.87	14.45	12	4.7	8.95	16.58	1
Ff12	182	404	66.9	8.87	13.03	30	5.0	8.93	16.72	1
Ff13	201	446	250.9	8.65	13.71	5	5.1	8.74	16.36	2
ZonaA	985	4472	>7200	-	-	-	20.9	7.43	15.66	42

Tabella 5.2

Nome	N° celle	N° port.	Tempo (sec.)	N° iter (mosse)	Interf. Media	N° pss
ZonaA	985	4472	20.9	50000	15.66	42
Vuoto900	2009	8795	44.9	100000	15.27	186
Z2000	30135	131925	762.2	1500000	15.14	2863

Tabella 5.3 – Risultati per reti di grandi dimensione.

5.13 Dinamica dell'esplorazione

Una semplice tecnica di ricerca che utilizza la libreria sviluppata è la seguente:

1. Genera un piano utilizzando un assegnazione casuale
2. Procedi con la Tabu Search in forma base per J iterazioni
3. Se hai raggiunto il numero massimo di iterazioni KMAX oppure se nell'ultimo ciclo la soluzione ottima corrente non è stata aggiornata allora prosegui altrimenti torna al passo 2
4. Per H iterazioni genera un *core* a partire dalla cella con la portante a costo peggiore e applica la Tabu Search per K iterazioni utilizzando solo le celle del core e per Z utilizzando tutte le celle.
5. Se il risultato viene ritenuto soddisfacente (secondo certi parametri) allora fermati, altrimenti genera un piano di partenza alternativo con una tecnica qualunque e torna al passo 2

Il risultato viene ritenuto soddisfacente se la soluzione è conforme alle specifiche date dall'utente. Ad esempio si può fissare un tempo massimo di esecuzione, o un numero di iterazioni complessivo. Se le specifiche lo consentono, l'algoritmo potrebbe cercare di migliorare ulteriormente la soluzione. Altrimenti, se all'utente interessa una soluzione ammissibile nel tempo minore possibile, si potrebbe imporre l'arresto non appena si giunga ad un minimo locale.

La tecnica a scelta (dell'utente o dell'algoritmo stesso) potrebbe essere fissata a priori, oppure scelta (anche casualmente) fra un insieme di procedure, o anche determinata a posteriori mediante semplice procedura "trial and error".

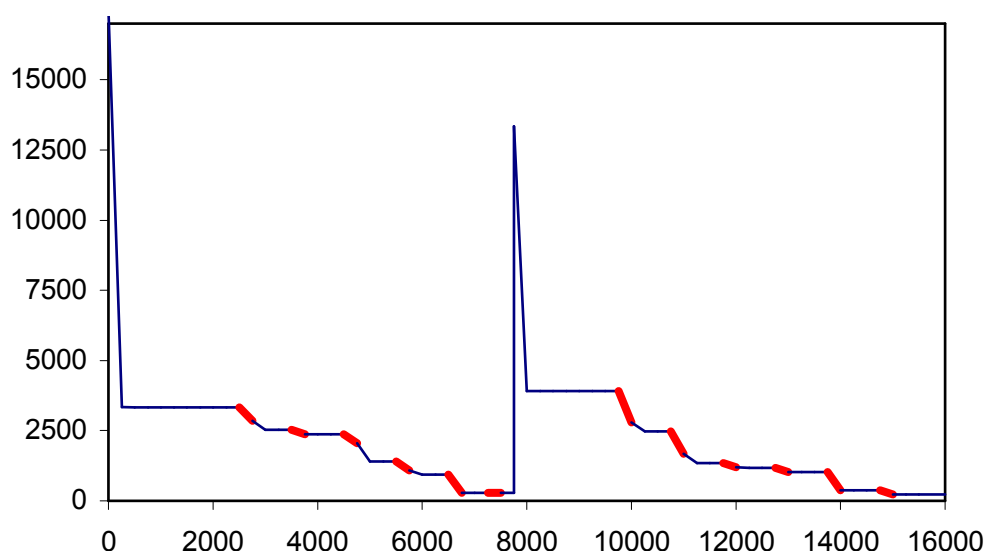


Figura 5.12 – Esempio di andamento del costo di un piano in funzione delle iterazioni svolte

In figura 5.12 viene presentato un tipico andamento del costo corrente di un piano (l'esempio è tratto dall'istanza "ZonaA") al variare delle iterazioni: tale esempio ha solo valore illustrativo: in una situazione reale verrebbero effettuati più cicli con più sottoricerche, e quest'ultime verrebbero sensibilmente anticipate. In questo caso sono presenti due cicli di ricerca con sei sottoricerche nei core (evidenziati in rosso). Si vede che la Tabu Search velocemente migliora il piano ma si arresta ben presto attorno un valore. A questo punto conviene che la fase di *diversificazione* inizi, intensificando la ricerca in una zona particolarmente critica. Inizialmente si avrà un aumento del costo, con la violazione anche di qualche vincolo ma, come si può notare, i benefici risultano alla fine significativi. Quando non esistono più sottozone critiche o si raggiunge il numero prestabilito di iterazioni, si può provvedere ad una nuova fase di *diversificazione* molto radicale (secondo picco). In questo caso, si utilizzano le informazioni raccolte dalla precedente ricerca nel senso che si privilegiano le celle a costo maggiore, come precedentemente spiegato. Questo porta ad una soluzione di partenza sensibilmente migliore della precedente, e aumenta la probabilità di individuare una nuova soluzione ottima. Questa tecnica, ovviamente, può essere iterata.

5.14 Possibili sviluppi futuri

Gli strumenti software sviluppati si sono dimostrati pienamente rispondenti scopi prefissi, in quanto permettono l'analisi di reti di grandi dimensioni aumentando il numero di iterazioni necessarie in base alla taglia dell'istanza. La durata di una singola iterazione si può ritenere quasi costante rispetto alle dimensioni dell'istanza, ed in ogni caso sufficientemente breve nell'ottica di un uso applicativo. La tecnica del *core*, inoltre, si è dimostrata molto efficiente nell'eliminare le violazioni di vincoli contenendo nel contempo il numero delle portanti sotto soglia.

Un possibile sviluppo è l'integrazione degli strumenti sviluppati in un algoritmo capace di adattarsi automaticamente ai vari tipi di istanze, magari con autoregolazione dei parametri di ricerca. È auspicabile una integrazione efficiente con le tecniche di ricerca basate sul *core*, come ad esempio il metodo BBfreq di Mannino e Sassano [8], al fine di analizzare a fondo le sotto-zone altamente congestionate, ed infine l'uso di tecniche efficaci ed efficienti che consentano una ulteriore riduzione del numero di portanti sotto-soglia, sempre entro tempi di esecuzione accettabili.

Conclusioni

In questa tesi è stato affrontato il complesso problema di assegnazione frequenziale, elencando e descrivendo i possibili approcci e le possibili tecniche risolutive, sia euristiche che esatte. Sono stati sviluppati e presentati nuovi strumenti operativi, che consentono un miglioramento sostanziale per quanto riguarda rapidità di esecuzione e qualità di soluzione; permettono, inoltre, l'analisi di reti di grandi dimensioni, in linea teorica limitate solo dalla memoria dell'elaboratore e comunque di decine di migliaia di celle.

Con una semplice implementazione si sono ottenuti risultati che dimostrano la validità degli strumenti e delle strutture dati, sia per istanze di piccola taglia che per esempi reali di notevole dimensione.

Questi strumenti, considerati in un'ottica generale, si dimostrano adatti non solo per il problema per cui sono stati creati, ma anche per una vasta classe di problemi di ottimizzazione.

Appendice A

Teorema

Siano P_1 e P_2 due piani di frequenza che differiscano solo nell' assegnazione di una frequenza in una cella h . In particolare per ipotesi sia per P_1 $A_{h,f} = 1$, $A_{h,g} = 0$ e per P_2 $A_{h,f} = 0$, $A_{h,g} = 1$.

Allora vale:

$$I_{P_2}^{\text{tot}} = I_{P_1}^{\text{tot}} - LD_{h,f} + LD_{h,g}$$

con $LD_{i,p} = RDS_{i,p} + EFF_{i,p}$.

Dimostrazione

Esplicitando l'espressione (5.3) rispetto ad una cella h con i suoi canali f e g si ottiene:

$$\begin{aligned} I^{\text{tot}} &= \sum_{\substack{i=1 \\ i \neq h}}^N \sum_{p=1}^M A_{i,p} \sum_{\substack{j=1 \\ j \neq i}}^N I_{i,j} \left(A_{j,p} + \frac{A_{j,p-1} + A_{j,p+1}}{NFD} \right) + \sum_{p=1}^M A_{h,p} \sum_{\substack{j=1 \\ j \neq h}}^N I_{h,j} \left(A_{j,p} + \frac{A_{j,p-1} + A_{j,p+1}}{NFD} \right) \\ I^{\text{tot}} &= \sum_{\substack{i=1 \\ i \neq h}}^N \sum_{p=1}^M A_{i,p} \sum_{\substack{j=1 \\ j \neq i,h}}^N I_{i,j} \left(A_{j,p} + \frac{A_{j,p-1} + A_{j,p+1}}{NFD} \right) + \sum_{\substack{i=1 \\ i \neq h}}^N \sum_{p=1}^M A_{i,p} I_{i,h} \left(A_{h,p} + \frac{A_{h,p-1} + A_{h,p+1}}{NFD} \right) + \\ &+ \sum_{p=1}^M A_{h,p} \sum_{\substack{j=1 \\ j \neq h}}^N I_{h,j} \left(A_{j,p} + \frac{A_{j,p-1} + A_{j,p+1}}{NFD} \right) \end{aligned}$$

$$\begin{aligned}
 I^{\text{tot}} = & \sum_{\substack{i=1 \\ i \neq h}}^N \sum_{\substack{p=1 \\ p \neq h}}^M A_{i,p} \sum_{\substack{j=1 \\ j \neq i,h}}^N I_{i,j} \left(A_{j,p} + \frac{A_{j,p-1} + A_{j,p+1}}{\text{NFD}} \right) + \sum_{\substack{i=1 \\ i \neq h}}^N \sum_{\substack{p=1 \\ p \neq f-1,f,f+1 \\ p \neq g-1,g,g+1}}^M A_{i,p} I_{i,h} \left(A_{h,p} + \frac{A_{h,p-1} + A_{h,p+1}}{\text{NFD}} \right) + \\
 & + \sum_{\substack{p=1 \\ p \neq f,g}}^M A_{h,p} \sum_{\substack{j=1 \\ j \neq h}}^N I_{h,j} \left(A_{j,p} + \frac{A_{j,p-1} + A_{j,p+1}}{\text{NFD}} \right) + \sum_{\substack{i=1 \\ i \neq h}}^N A_{i,f-1} I_{i,h} \left(A_{h,f-1} + \frac{A_{h,f-2} + A_{h,f}}{\text{NFD}} \right) + \\
 & + \sum_{\substack{i=1 \\ i \neq h}}^N A_{i,f} I_{i,h} \left(A_{h,f} + \frac{A_{h,f-1} + A_{h,f+1}}{\text{NFD}} \right) + \sum_{\substack{i=1 \\ i \neq h}}^N A_{i,f+1} I_{i,h} \left(A_{h,f+1} + \frac{A_{h,f} + A_{h,f+2}}{\text{NFD}} \right) + \\
 & + \sum_{\substack{i=1 \\ i \neq h}}^N A_{i,g-1} I_{i,h} \left(A_{h,g-1} + \frac{A_{h,g-2} + A_{h,g}}{\text{NFD}} \right) + \sum_{\substack{i=1 \\ i \neq h}}^N A_{i,g} I_{i,h} \left(A_{h,g} + \frac{A_{h,g-1} + A_{h,g+1}}{\text{NFD}} \right) + \\
 & + \sum_{\substack{i=1 \\ i \neq h}}^N A_{i,g+1} I_{i,h} \left(A_{h,g+1} + \frac{A_{h,g} + A_{h,g+2}}{\text{NFD}} \right) + A_{h,f} \sum_{\substack{j=1 \\ j \neq h}}^N I_{h,j} \left(A_{j,f} + \frac{A_{j,f-1} + A_{j,f+1}}{\text{NFD}} \right) + \\
 & + A_{h,g} \sum_{\substack{j=1 \\ j \neq h}}^N I_{h,j} \left(A_{j,g} + \frac{A_{j,g-1} + A_{j,g+1}}{\text{NFD}} \right)
 \end{aligned}$$

$$\begin{aligned}
 I^{\text{tot}} = & \sum_{\substack{i=1 \\ i \neq h}}^N \sum_{\substack{p=1 \\ p \neq h}}^M A_{i,p} \sum_{\substack{j=1 \\ j \neq i,h}}^N I_{i,j} \left(A_{j,p} + \frac{A_{j,p-1} + A_{j,p+1}}{\text{NFD}} \right) + \sum_{\substack{i=1 \\ i \neq h}}^N \sum_{\substack{p=1 \\ p \neq f-1,f,f+1 \\ p \neq g-1,g,g+1}}^M A_{i,p} I_{i,h} \left(A_{h,p} + \frac{A_{h,p-1} + A_{h,p+1}}{\text{NFD}} \right) + \\
 & + \sum_{\substack{p=1 \\ p \neq f,g}}^M A_{h,p} \sum_{\substack{j=1 \\ j \neq h}}^N I_{h,j} \left(A_{j,p} + \frac{A_{j,p-1} + A_{j,p+1}}{\text{NFD}} \right) + A_{h,f-2} \sum_{\substack{i=1 \\ i \neq h}}^N I_{i,h} \frac{A_{i,f-2}}{\text{NFD}} + A_{h,f-1} \sum_{\substack{i=1 \\ i \neq h}}^N I_{i,h} \left(A_{i,f-1} + \frac{A_{i,f}}{\text{NFD}} \right) + \\
 & + A_{h,f+1} \sum_{\substack{i=1 \\ i \neq h}}^N I_{i,h} \left(A_{i,f+1} + \frac{A_{i,f}}{\text{NFD}} \right) + A_{h,f+2} \sum_{\substack{i=1 \\ i \neq h}}^N I_{i,h} \frac{A_{i,f+2}}{\text{NFD}} + A_{h,g-2} \sum_{\substack{i=1 \\ i \neq h}}^N I_{i,h} \frac{A_{i,g-2}}{\text{NFD}} + \\
 & + A_{h,g-1} \sum_{\substack{i=1 \\ i \neq h}}^N I_{i,h} \left(A_{i,g-1} + \frac{A_{i,g}}{\text{NFD}} \right) + A_{h,g+1} \sum_{\substack{i=1 \\ i \neq h}}^N I_{i,h} \left(A_{i,g+1} + \frac{A_{i,g}}{\text{NFD}} \right) + A_{h,g+2} \sum_{\substack{i=1 \\ i \neq h}}^N I_{i,h} \frac{A_{i,g+2}}{\text{NFD}} +
 \end{aligned}$$

$$\begin{aligned}
 & + A_{h,f} \sum_{\substack{i=1 \\ i \neq h}}^N I_{i,h} \left(A_{i,f} + \frac{A_{i,f-1} + A_{i,f+1}}{NFD} \right) + A_{h,g} \sum_{\substack{i=1 \\ i \neq h}}^N I_{i,h} \left(A_{i,g} + \frac{A_{i,g-1} + A_{i,g+1}}{NFD} \right) + \\
 & + A_{h,f} \sum_{\substack{j=1 \\ j \neq h}}^N I_{h,j} \left(A_{j,f} + \frac{A_{j,f-1} + A_{j,f+1}}{NFD} \right) + A_{h,g} \sum_{\substack{j=1 \\ j \neq h}}^N I_{h,j} \left(A_{j,g} + \frac{A_{j,g-1} + A_{j,g+1}}{NFD} \right)
 \end{aligned}$$

$$I^{\text{tot}} = K + A_{h,f} \text{EFF}_{h,f} + A_{h,g} \text{EFF}_{h,g} + A_{h,f} \text{RDS}_{h,f} + A_{h,g} \text{RDS}_{h,g}$$

con K costante comprendente tutte le sommatorie tranne le ultime quattro.

Nel nostro caso, sfruttando l'ipotesi si ha che:

$$I_{p_1}^{\text{tot}} = K + LD_{h,f} \quad I_{p_2}^{\text{tot}} = K + LD_{h,g}$$

e in definitiva

$$I_{p_2}^{\text{tot}} = I_{p_1}^{\text{tot}} - LD_{h,f} + LD_{h,g} \quad \square$$

Bibliografia

- [1] Aart E., J.K. Lenstra, *Local search in combinatorial optimization*, J.K. Lenstra Editions, 1992.
- [2] Colombo G., Lepschy C., Menolascino R., Settimo F., *Un metodo autoadattativo per l'assegnazione frequenziale a sistemi cellulari per servizi mobili*, Rapporto Tecnico, CSELT, 1994.
- [3] Dong-wan Tcha, Yong-joo Chung, Taek-jin Choi, *A new Lower Bound for the Frequency Assignment Problem*, IEEE/ACM Transaction on Networking, 1997.
- [4] Fischetti M., Lepschy C., Minerva G., Romanin Jacur G., Toto E., *Frequency Assignment in mobile radio system using Branch-and-Cut Techniques*, Università di Padova, 1996.
- [5] Glover F., Laguna M., *Modern Heuristic Techniques for Combinatorial Problems*, 1992.
- [6] Hertz A., Taillard E., De Werra D., *Local Search in Combinatorial Optimisation*, Wiley & Sons, 1997.
- [7] Hurley S., Smith D.H., Thiel S.U. *FASoft : A System for Discrete Channel Frequency Assignments*, Radio Science, 1997.
- [8] Mannino C., Sassano A., *An Enumerative Algorithm for the Frequency Assignment Problem*, Università La Sapienza di Roma, 1998.
- [9] Massignan M., *Pianificazione Frequenziale nelle Reti Radiomobili Cellulari*, Tesi di Laurea, Istituto di Ingegneria Gestionale dell'università di Padova, Vicenza 1998. (G. Romanin Jacur relatore)

- [10] Metzger B. H., *Spectrum Management Techniques*, 38th National O.R.S.A. Meeting, Detroit, 1970
- [11] Minerva G., *Metodi di Ricerca Locale per Problemi di Ottimizzazione Combinatoria*, Rapporto Tecnico, CSELT, 1991.
- [12] Berdin E. S., *Pianificazione Frequenziale per il Servizio Radiomobile Regionale: Modello e Algoritmo Risolutivo*, Tesi di Laurea, Istituto di Ingegneria Gestionale dell'università di Padova, Vicenza 1999. (G. Romanin Jacur relatore)
- [13] Battiatì R., *Reactive Search: Toward Self-Tuning Heuristics*, Dipartimento dell'Università di Trento, 1996.
- [14] Mouly M., Pautet M.B., *The GSM System for Global Communications*, Palaiseau, France, 1992.
- [15] Del Favero A., *Tecniche Numeriche Avanzate ed Algoritmi Efficienti per l'Intelligenza Artificiale*, Mc Graw-Hill, 1999.
- [16] Borndörfer R., Eisenblätter A., Grötschel M., Martin A., *Frequency Assignment in Cellular Phone Networks*, Konrad-Zuse-Zentrum für Informationstechnik, Berlino, 1997
- [17] Borndörfer R., Eisenblätter A., Grötschel M., Martin A., *The Orientation Model for Frequency Assignment Problems*, Konrad-Zuse-Zentrum für Informationstechnik, Berlino, 1998.
- [18] Sedgewick R., *Algorithms in C++*, Princeton University, Addison Wesley Publishing Company, 1992.
- [19] Cormen T., Leiserson C. E., Rivest R. L., *Introduction to Algorithms*, MIT Press, 1990.

- [20] Vocale A., *Progettazione e Sviluppo di un Algoritmo per la Pianificazione Frequenziale nelle Reti Radiomobili Cellulari*, Tesi di Laurea, Dipartimento di elettronica ed informatica dell'università di Padova, 1996. (M. Fischetti e G. Romanin Jacur relatori)
- [21] Wei Wang, Craig K. Rushforth, *An Adaptative Local-Search Algorithm for the Channel-Assignment Problem (CAP)*, IEEE Transaction on Vehicular Technology, 1996.
- [22] “ETSI – Standarding Telecommunications Products and Services”, ETSI, URL: <http://www.etsi.org/>.
- [23] “Global System for Mobile Communications, tecnica, teoria e pratica del sistema GSM”, URL: <http://www.gsmworld.netcity.it/>.