

## MSA - 23.03.2021

Continuiamo e con oggi chiudiamo il capitolo relativo agli aspetti strettamente connessi alla realizzazione di un canale di comunicazione wireless. Oggi completiamo questo capitolo.

Nelle lezioni passate abbiamo presentato lo standard per comunicazioni wireless in ambito locale, 802.11. Nel capitolo di oggi spostiamo lo sguardo su un'altra tipologia di reti che utilizzano questo tipo di tecnologia, quelle cellulari. Non scenderemo molto nei dettagli. Vi darò delle informazioni sull'architettura generale delle reti, necessarie per introdurre il capitolo successivo, la gestione della mobilità all'interno di queste reti. Queste reti sono reti che si poggiano sulle infrastrutture di reti cellulari. Sono infrastrutture la cui diffusione nel mondo risale ad alcuni decenni fa. Inizialmente questo tipo di infrastruttura è stata pensata per supportare unicamente traffico di tipo vocale in mobilità. Di conseguenza la prima generazione di reti cellulari installata sul territorio risentiva fortemente di questa caratterizzazione. Tenendo presente le caratteristiche di un traffico di tipo locale, adottava una metodologia di definizione dei percorsi di tipo circuit switch e le bande erano dell'ordine delle pochissime decine di Kb per secondo. Su questa rete pensata per questa specifica di traffico, poi si è sviluppata un'infrastruttura più generale dedicata anche al trasporto di dati di diversa natura. Con esigenze operative e scenari di riferimento diversi da quelli delle reti locali. Dal punto di vista delle prestazioni offerte, giusto per rimarcare che tutt'ora, a parità di evoluzione della tecnologia, le reti locali offrono qualcosa in più rispetto alle reti ad alta copertura che si poggiano su infrastrutture cellulari.

Partiamo dalla prima generazione di queste reti, fino ad arrivare alle reti attuali. Partiamo dalla seconda generazione, che è la prima in cui si adottano tecniche digitali e non analogiche come nella prima, e si inizia a porre per la prima volta l'idea di trasportare anche traffico dati e non solo vocale. La figura schematizza l'architettura generale. Su questo tipo di sistemi ci ritorneremo più avanti. Nella prima trattazione vi darò notizie più strettamente connesse con il capitolo che stiamo coprendo, con qualche accenno alla parte più infrastrutturale che vedremo nel capitolo sulla gestione della mobilità.

Guardando nella figura porto anche il vocabolario utilizzato dallo standard. Nel mondo delle reti cellulari un dispositivo mobile viene indicato con MS, mentre l'AP viene chiamato BSS. La figura evidenzia la strutturazione gerarchica dell'architettura delle reti. Al di là dei nodi mobili veri e propri, abbiamo un livello costituito dai BTS, che corrispondono alle antenne disperse sul territorio, agli AP di una rete wireless locale. A questo primo livello infrastrutturale corrisponde un secondo livello, saliamo nella gerarchia, BSC a cui fanno capo un certo numero di BTS. Collettivamente BSC e BTS che fanno capo ad esso viene indicato con BSS, la stessa usata per il protocollo 802.11 anche se il significato cambia. E poi si sale di un ulteriore livello nella gerarchia, MSC, controllori di livello a cui fanno capo questi BSC. Rimarco l'aspetto della strutturazione gerarchica in quanto è un tipo di organizzazione che persisterà per varie generazioni successive di queste reti, fino ad essere abbandonata recentemente in favore di un diverso tipo di organizzazione. Tolta la parte wireless e quella che riguarda le comunicazioni tra i nodi mobili e le stazioni base, tutte le

altre comunicazioni tipicamente avvengono via cavo tramite interfaccia A definita nello standard.

Riguardo le modalità di organizzazione del canale wireless è un'organizzazione TDMA, in modalità circuit switch. Ogni entità che si connette ad un'antenna, un nodo BTS, e vuole comunicare una porzione del canale wireless gestito da quel BTS, viene assegnato uno slot temporale che occupa una posizione fissa nel tempo. Il tempo è diviso in 8 slot successivi, un frame complessivamente ha durata complessiva di poco più di 4 millisecondi e mezzo. Ad ogni partecipante viene assegnato uno specifico slot. All'interno dello slot riescono ad essere codificati un centinaio di bit di carico utile. Considerando che questo dimensionamento è ritagliato per traffico di tipo locale, il segnale che viene emesso in 4 millisecondi e mezzo richiede un centinaio di bit per essere codificato e trasmesso. C'è anche la possibilità, per aumentare la capacità del canale, si possono assegnare un frame sì ed uno no, raddoppiando la capacità del canale ma con un leggero peggioramento della qualità.

Stiamo parlando di un'organizzazione di tipo cellulare. Oltre alla condivisione del canale rispetto al tempo si abbina ad una condivisione anche rispetto allo spazio per consentire il riuso delle stesse frequenze e tipologia di canale wireless in aree sufficientemente distanti da quella in cui viene utilizzata una certa parte del canale wireless. Avendo questo tipo di ripartizione spaziale, e non volendo inibire la mobilità degli utenti che utilizzano questi dispositivi, l'infrastruttura si deve carico della necessità di gestire la transizione tra una cella e quella adiacente dell'utente. Guardando all'infrastruttura gerarchica che sta sopra alle celle, queste transizioni da una cella a quella adiacente possono essere classificate secondo quattro tipologie che dipendono dalle entità che controllano le 4 celle adiacenti tra cui avviene la transizione per effetto della mobilità.

Tipo 1. Un nodo mobile passa da una cella ad un'altra che ha esattamente le stesse caratteristiche. Sono due celle adiacenti che fanno capo allo stesso BTS. Una singola antenna può coprire più celle diverse adiacenti tra di loro utilizzando la tecnologia di antenna direzionale.

Tipo 2. Il passaggio avviene da una cella coperta da un'antenna, ad una cella adiacente coperta da un BTS diverso, dove però i due BTS fanno capo allo stesso controllore BSC.

Tipo 3. Il passaggio avviene da una cella ad un'adiacente che fanno capo a BTS diversi, a BSC diversi, che però fanno capo allo stesso MSC.

Tipo 4. Il passaggio avviene tra due celle adiacenti che hanno BTS diverso, BSC diverso ed anche un MSC diverso. Questo lo chiariremo quando andremo avanti, gli handoff di tipo 3 andrebbero distinti in due sottocasi. Ma a questo livello non introduciamo questa ulteriore caratterizzazione.

Lo scopo dell'handoff è garantire la non interruzione del canale wireless. La transizione senza interruzioni sfrutta il fatto che il segnale wireless si attenua in modo *dolce*, per cui nelle zone di confine tra due celle c'è una zona di margine in cui i segnali di entrambi le antenne sono percepibili. All'interno di questa zona c'è il tempo per rilasciare il vecchio canale wireless ed agganciarsi al segnale wireless della nuova antenna.

Lo schema definisce in una rete di seconda generazione, ma non solo, quello che succede quando c'è un handoff di tipo 3, che comporta il passaggio ad un BTS che fa capo ad un nuovo BSC, che fanno capo allo stesso MSC.

Problemi di questa architettura. I problemi sono problemi se io voglio usare questa infrastruttura per trasmettere dati. Se non mi pongo il problema, la utilizzo unicamente per gli scopi per cui era stata pensata, non ha bisogno di nessuna evoluzione particolare. Se voglio trasmettere anche traffico dati, i tassi di trasmissione garantiti sono relativamente bassi. Da un punto di vista di modello economico, quello prevalente era di addebitare costi relativi alla durata della comunicazione, poco plausibile per il traffico dati.

Prima evoluzione, da questa seconda generazione è la 2.5G, perché è proprio un'evoluzione incrementale minimale rispetto alla generazione 2. Proprio considerando che il traffico vocale è un traffico che tipicamente è caratterizzato dalla presenza di tante pause mentre si svolge. L'idea di questa evoluzione è quella di non cambiare nulla nell'infrastruttura, semplicemente nei buchi che si creano all'interno di un canale utilizzato per emettere voce, in questi buchi infilo del traffico dati. Usando questo semplice espediente si riescono a raggiungere velocità nominali di traffico dati salendo di un ordine di grandezza, arrivando alle centinaia di Kb per secondo. Questa seconda generazione e mezzo è la base per l'evoluzione successiva, la terza generazione, UMTS.

Giusto per dare un'idea del tipo di prestazioni che la generazione 2.5 garantiva, giusto anche per misurare la distanza percorsa da allora, le prestazioni minime garantite erano abbastanza grossolane. Al meglio che si riusciva a garantire dal punto di vista delle prestazioni si offrivano latenze dell'ordine del mezzo secondo, mediamente.

Dal punto di vista infrastrutturale le novità più rilevanti introdotti dalla prima evoluzione riguardano la parte colorata in verde, ma che non vedremo nel dettaglio. Guardando la parte più ai bordi della rete, sostanzialmente se guardate le sigle organizzate tra di loro, non cambia molto. Rimangono i BSS ripartito in un BSC a cui fanno capo vari BTS.

Dal punto di vista della pila protocollare, quello si evolve a livello di canale wireless, quello che cambia è la gestione del canale wireless, dove nella generazione 2G era un TDMA secco usato in modalità circuit switch, qui l'idea è di usare slot inutilizzati per codificare e trasmettere pacchetti dati. Oltre ad utilizzare gli slot non utilizzati, per trasmettere dati si può anche scegliere di riservare permanentemente alcuni slot per trasmettere traffico dati, per garantire una banda minima unicamente al traffico dati. La condivisione di questi slot per il traffico dati avviene con un protocollo a contesa. Dal punto di vista della rete dati generale, una rete cellulare organizzata in questo modo consideriamo che questa rete dal punto di vista dell'infrastruttura generale che avvolge tutto il pianeta, una rete cellulare estesa quanto vogliamo viene vista dal punto di vista architetturale come una rete di livello 2, su cui poi si poggia una rete tipicamente IP. X.25 nella slide era un concorrente di IP.

Arriviamo ai primi anni 2000, terza generazione, UMTS. Il risultato di una promozione fatta da questo organismo internazionale, ITU, che pochi anni prima del passaggio di secolo e di millennio aveva lanciato l'idea di definire una nuova evoluzione con caratteristiche più evolute. L'idea è di supportare delle bande di un'ordine di grandezza superiore, Mbit per secondo, per soggetti con velocità minore ai 10 km/h. Per velocità più elevate le prestazioni

garantite si abbassano. L'idea è di fare affidamento forte sull'infrastruttura pre-esistente. E' ancora un'evoluzione incrementale rispetto alla generazione 2G, dove così come abbiamo visto avvenire nello standard 802.11, l'evoluzione ed il conseguente miglioramento delle prestazioni ha riguardato la parte strettamente wireless, della modalità della gestione del canale.

Facendo riferimento alla figura che descrive l'architettura di una rete di terza generazione, evidenzia che all'interno della generazione si continua a mantenere l'idea che reti di questo tipo devono supportare le reti precedenti, ed anche la parte innovativa, dove i BTS diventano Node B ed i BSC diventano RNC, ma la configurazione architetturale rimane di un'organizzazione fortemente gerarchica a tre livelli. Parte core rimasta invariata rispetto alla generazione precedente.

Passiamo da una gestione del canale TDMA ad una gestione a cui alla condivisione in tempo si può affiancare una condivisione in frequenza del canale wireless. Dal punto di vista protocollare cambia poco, una delle poche differenze di rilievo, il protocollo GTP utilizzato per far parlare le varie entità in gioco, prima era limitato a gestire il livello UDP/IP nel core della rete, e qui si estende anche per maggiore uniformità anche alle frange della rete, i controllori di ultimo livello. Qui inizia a manifestarsi la spinta verso l'uniformizzazione della rete che poi si realizzerà pienamente a partire dalla generazione successiva.

Quarta generazione. Altro salto di qualità dal punto di vista delle prestazioni, banda e latenza. Dal punto di vista architetturale si introduce una novità importante. Novità dal punto di vista wireless, propriamente detto. eNB sta per Node B, stesso nome che si utilizzava nella generazione precedente, con l'aggiunta del prefisso e. Se ci facciamo caso, questo nodo B, come si rapporta al core? Direttamente. E' sparito un livello nella gerarchia. Non c'è più un livello intermedio. L'antenna si connette direttamente alla parte core, quella che si occupa della gestione della mobilità. La rete si è appiattita.

Dal punto di vista, se vogliamo riassumere brevemente questa carrellata storica, quello che abbiamo visto è che siamo partiti da una generazione 2G, dove si trasmetteva solo traffico vocale più un traffico dati che era gestito con la tecnologia di quelli che conosciamo come SMS, e dove il modo di gestire le connessioni è di tipo circuit switch. Le generazioni 2.5 e 3 si mantengono la parte legacy, pienamente supportata, ma a questa si affianca la parte più nuova migrata alle esigenze specifiche di traffico dati che non sia puramente vocale. Si adotta anche una tecnologia di trasporto diverso. Per passare poi alle generazioni di quarta, quinta, sesta, in cui non viene più garantito il supporto delle generazioni precedenti e tutto, incluso il traffico vocale, viene gestito in modalità packet switching.

Dal punto di vista dell'infrastruttura passiamo dall'infrastruttura gerarchica a tre livelli, si passa nelle generazioni dalla quarta in poi ad una struttura estremamente più piatta, in cui ogni nodo si interfaccia direttamente con la parte core, che consente la comunicazione su grande scala.

---

## Mobility management

Cambiamo capitolo veramente. Finora abbiamo coperto problematiche relative alla realizzazione in senso stretto di un canale di comunicazione wireless. Non ci siamo posti più di tanto domande a problematiche sul se e se i nodi che vogliono usare il canale sono entità mobili o fissa. A questo punto cambiamo punto di vista e ci dedichiamo all'esplorazione di una serie di questioni che riguardano l'aspetto della mobilità.

Dobbiamo chiarirci su che tipo di mobilità vogliamo considerare. Un primo livello potrebbe essere la mobilità dei terminali, di un dispositivo di calcolo e comunicazione, che per le sue caratteristiche intrinseche fisiche può essere agevolmente cambiato di posizione. La mobilità di questo oggetto crea dei problemi che andranno risolti. Un altro tipo di mobilità è la mobilità personale. L'accento non è su un oggetto fisico, ma è un tipo di mobilità con la prospettiva umano centrica. Quello che si immagina è che noi essere umani siamo esseri dotati di mobilità. In generale la mobilità personale non implica la mobilità di dispositivi che portiamo fissi con noi. Sono scenari in cui dato che una persona può trovarsi in vari posti nella giornata, come riuscire a fare arrivare una comunicazione in ogni caso a questa persona, a prescindere dal fatto che porti continuamente con sé un dispositivo che faccia da terminale di questa comunicazione. Mobilità di sessione si intendono problematiche relative al mantenimento di una sessione in corso, streaming video ad esempio, che ad un istante di tempo ha come endpoint un certo dispositivo, e come conseguenza di una delle due mobilità precedenti, si vuole garantire la persistenza di questa sessione anche nel caso del cambiamento di posizione dell'entità su cui questa sessione si poggia. Ultima mobilità dei servizi, umano centrica, ma vista in una visione complementare rispetto a quella personale. In quella prima si guarda al mantenimento della capacità di raggiungimento di un essere umano per comunicazioni che sono iniziate da una parte terza che vuole raggiungere un essere umano a prescindere dalla sua posizione. Qua invece si intende l'offrire ad un essere umano, nonostante i suoi cambi di posizione causa mobilità, la possibilità di trovare ovunque vada un ambiente sempre uniforme e familiare sui diritti di accesso, profili d'uso, costi da pagare ecc. Di tutti questi tipi di mobilità ci focalizzeremo sul primo. Come consentire la comunicazione tra entità fisiche, dispositivi dotati di capacità di calcolo e posizione, la cui posizione cambia con frequenze ed ampiezze relativamente grandi.

Quali sono i problemi fondamentali da affrontare per dare una soluzione completa al problema della mobilità dei terminali? Continuità della sessione. Se un terminale è coinvolto in una iterazione in corso, come garantire il mantenimento della continuità di questa sessione nonostante la mobilità del terminale. Qua parliamo di una sessione di lavoro che è in corso quando il terminale cambia posizione. L'altro aspetto del problema è come fare a raggiungere un terminale. Come avviare una sessione indirizzata ad un terminale la cui posizione non è fissa nel tempo. Sono due problemi distinti. Si può risolvere uno senza risolvere necessariamente l'altro. Una soluzione globale dovrebbe riguardare questi due aspetti. Li abbiamo chiamati problemi, ma che siano problemi per i quali vale la pena impegnarsi a risolverli, dipende dallo scenario di riferimento che abbiamo. Non è detto che sempre e comunque possa valere la pena offrire una soluzione da eventuali problemi causati dalla mobilità dei terminali nei confronti di questi due aspetti.

Per fare un esempio, se siamo impegnati in una sessione di navigazione sul web, saltando da un sito all'altro, ed intanto ci stiamo spostando, ci darebbe fastidio se ogni tot la nostra connessione cade e dobbiamo ricominciare tutto daccapo. Sarei interessato ad avere una gestione della mobilità che mi garantisca la continuità della mia sessione di lavoro. Se invece sono occupato in attività di risoluzione di un nome di dominio, DNS, a causa dello spostamento perdo la sessione interattiva con il server DNS, rifaccio l'interrogazione dal punto in cui mi trovo e comporterà solamente un piccolo ritardo nella risoluzione del nome. Stesso per la raggiungibilità. Diverso è il caso di un dispositivo che svolge il ruolo di server, ci sono tanti potenziali interessati a contattare un servizio installato su questo server, o se il mio nodo è il terminale di una VPN, sono interessato a mantenere il suo contatto con la rete nonostante i cambi di posizione. Se invece il mio dispositivo che si sposta è il client per una sessione di navigazione sul web che deve ancora iniziare, il cambio di posizione non mi interessa che venga registrato perché essendo io a contattare il server, sono io nel momento in cui lo contatto rendere nota al server la mia posizione attuale.

La mobilità può essere gestita a vari livelli. Un primo livello come abbiamo visto quando abbiamo parlato del protocollo 802.11 in particolare, può essere gestita entro certi limiti a livello 2. Questa è un'animazione, la percorriamo rapidamente giusto per dare un'idea di cosa significa da un punto di vista visivo, a cosa potrebbe voler far fronte una gestione fatta a livello 2 della mobilità. Immaginiamo di avere un dispositivo mobile che inizialmente è connesso ad una rete 3G, rientra a casa dotata di un hotspot 802.11, si rileva la presenza di una rete wifi quindi un eventuale sessione di lavoro utilizzando la rete 3G viene commutata sull'interfaccia 802.11, quindi continua a svolgersi sulla rete 802.11. Poi in successione esce da casa, si dirige verso l'aeroporto, la connessione WiFi inizia a deteriorarsi, si esce all'esterno e c'è una copertura di tipo cellulare ma anche la tecnologia WiMax, quindi di nuovo si commuta la comunicazione su questa nuova interfaccia, si continua ad andare avanti, però a questo punto il sistema si rende conto che il livello di carica sta diminuendo, e sapendo che questo tipo di tecnologia WiMax ha consumi energetici superiori, spegne l'interfaccia WiMax e si riconnette alla rete 3G. Poi si rientra in aeroporto, ci si può connettere alla presa di corrente e ci si sgancia dalla rete 3G e si utilizza la rete offerta dall'aeroporto.

Tutti questi handoff possono essere tutti gestiti a livello 2, in maniera più o meno semplice. In caso di transizioni tra reti omogenee la decisione su quando e come effettuare l'handoff è basata solo su considerazioni relative alla qualità del segnale offerto, non ci sono altre ragioni che possono suggerire la necessità di sganciarsi dal BTS ed agganciarsi ad un'antenna diversa se non legate alla diversa qualità che io percepisco, come conseguenza tipicamente dei miei spostamenti. Se invece mi trovo in una situazione eterogenea, in questo caso come l'esempio voleva illustrare, in alcuni casi la decisione se effettuare o meno l'handoff non è detto che debba essere guidata solo da considerazioni sulla qualità, ma e/o considerazioni su aspetti sul risparmio energetico, costo monetario, sicurezza. Tutte le informazioni che si vorrebbero tenere conto per decidere se fare o meno handoff, possono essere informazioni che corrispondono al rilevamento di elementi particolari (rilevamento o meno della presenza e caratteristiche di un AP) o dal valore che assumono certi parametri (potenza segnale, caratteristica antenne disponibili, condizioni economiche o di altra natura di accesso al servizio offerto ecc).

Tutto questo se ci limitiamo a livello 2. Ma esistono anche degli standard che definiscono il modo di reagire a queste situazioni. Però il livello 2 da solo non è sufficiente per affrontare tutti i problemi relativi alla mobilità. E' tipicamente necessario aver anche un supporto dai livelli superiori.

Vediamo di allargare la prospettiva e passare all'altro aspetto della gestione della mobilità, quello su come garantire la raggiungibilità dei nodi. Come si fa a raggiungere all'interno di un'infrastruttura più o meno ampia, prima dell'apparizione dei nodi mobili? Soluzione tipicamente utilizzata, nei tempi in cui rete voce e dati erano distinte, in entrambi le reti il principio di base per il raggiungimento di nodi fissi era lo stesso: ad ogni entità che era l'endpoint di una comunicazione veniva associato un identificativo unico valido su scala globale, che svolgeva un doppio ruolo. Non solo quello di identificare il nodo ma anche un'informazione sul cammino dentro la rete per arrivare a quel nodo. Questa è l'idea dei numeri di telefono fisso, che avevano una struttura gerarchica per aree geografiche sempre più piccole, e stessa cosa per gli indirizzi internet. L'idea è che l'indirizzo globale lo posso spezzare in pezzi che servono a restringere la zona via via verso cui mi devo indirizzare per individuare il terminale di una certa comunicazione. Tutto questo va benissimo se i nodi non si spostano, ma chiaramente nel momento in cui questo non è più vero ed i terminali cominciano a spostarsi non posso più pensare di associare ad un unico codice questo doppio ruolo di essere identificativo unico indipendente dalla posizione e portare incastrato con se il percorso da seguire per arrivare a quella posizione che però cambia nel tempo.

In una rete internet, nella rete statica internet, l'indirizzamento avviene così: quando un pacchetto con un certo indirizzo di destinazione arriva al router, ha la sua tabella di routing di indirizzamento che associa ogni interfaccia ad uno o più prefissi di rete. Se ad esempio arriva un pacchetto con questo destinatario, il router individua qual'è l'interfaccia che corrisponde ad un prefisso presente nell'indirizzo, fa uscire il pacchetto da questa interfaccia ed il pacchetto viaggia sulla sottorete corrispondente e verrà poi catturato ed arriverà alla destinazione prevista. Se il nodo cambia il suo punto di aggancio, se continuo a riferirlo usando l'indirizzo IP, nella struttura di indirizzamento statica il pacchetto continuerà a seguire lo stesso percorso di prima, ma il routing statico di internet non mi offre mezzi per reindirizzare il pacchetto per quella che è attualmente la locazione del nodo individuata dall'indirizzo IP. Come faccio, per esempio in una rete internet o in una rete cellulare, come faccio ad arrivare un pacchetto alla destinazione giusta, assumendo che questo destinatario logico cambi la sua posizione.

Esploreremo varie soluzioni. Per costruire una cornice comune di riferimento che ci consenta di inquadrare varie soluzioni che discuteremo a questo problema generale, adotteremo uno schema classificatorio di questo tipo: immagineremo che tutte le soluzioni che esploreremo le possiamo pensare come punti all'interno di uno spazio multidimensionale. Le dimensioni che caratterizzano questo spazio sono le quattro che vedete: in che modo si sceglie di tenere traccia della posizione occupata dal terminale mobile in un certo istante di tempo, qual'è l'ampiezza della regione entro cui può avvenire la mobilità supportata da una certa soluzione, a quale livello nella pila architetturale si colloca la soluzione che esploreremo, quali entità vengono coinvolte nella realizzazione di una certa soluzione. Per ogni soluzione diremo, rispetto ad ognuna delle quattro, qual'è il

valore assunto dalla rispettiva coordinata. Interpretiamo ogni soluzione come un punto all'interno di questo spazio.

Modalità con cui si può tenere traccia della posizione di un terminale. Riguarda l'algoritmo che si usa per mantenere aggiornata quest'informazione. Possiamo distinguere due tipologie di approccio: pro attivi e reattivi. Tracciamento pro attivo, l'idea di base è quello che l'infrastruttura dedicata a questo problema della raggiungibilità di un nodo è sempre e comunque impegnata a mantenere aggiornata l'informazione sulla posizione attuale del nodo. A prescindere dall'esistenza o meno di una richiesta di localizzare quel nodo. I vantaggi sono la riduzione al minimo la latenza tra il nascere dell'esigenza comunicativa e la capacità di soddisfare quest'esigenza. Lato negativo, c'è il rischio di un overhead eccessivo delle risorse impegnate dal sistema, soprattutto considerando che alcuni degli sforzi fatti dal sistema per tenere aggiornate queste informazioni sulla posizione dei nodi potrebbero essere inutili, se alcuni nodi sono raramente contattati. Questo spreco di risorse può essere legato a vari fattori: tempo di calcolo, traffico sulla rete perché tipicamente devo far viaggiare informazioni in continuo per stare dietro alla posizione dei nodi mobili, soprattutto se questi sono caratterizzati da una mobilità elevata, ed anche occupazione di memoria notevole per tenere aggiornate tutte le tabelle che tengono traccia della posizione di ogni singolo nodo. E' una soluzione che in alcuni casi si ritiene vantaggioso adottare.

All'estremo opposto abbiamo il tracciamento reattivo. In prima battuta il sistema non fa nulla per conoscere e mantenere l'informazione della posizione dei nodi che fanno capo ad esso a meno che non ci sia una richiesta di rintracciare il nodo. Il sistema reagisce alle richieste di localizzazione. L'idea è che quando il sistema fa qualcosa lo fa perché qualcuno lo chiede, si vuole ovviare al problema di sprecare tempo, spazio, banda, per azioni che non vanno a beneficio di nessuno. Questo comporta dei costi, un aumento della latenza, del ritardo nel rintracciare un nodo, e non è detto che comporti un'effettiva riduzione, a meno che non agisca in modo intelligente calibrando bene lo scenario in cui applicare questa metodologia di tracciamento, non è detto che l'overhead rispetto al caso precedente si riduca in modo effettivo. In casi sfortunati potrebbe anche risultare eccessivo. Se ci pensiamo un attimo, se agisco al buio non guardando dove vanno i nodi, se io ho ignorato ciò che quel nodo fa, evidentemente in prima battuta l'unica cosa che posso fare è cercarlo dappertutto. Il meglio che posso fare è inondare la rete di messaggi di ricerca per rintracciare il nodo, che potrebbe essere drammatico dal punto di vista della scalabilità di questi sistemi. Le cose non stanno così, questi tipi di approcci possono portare risparmi notevoli se utilizzati in modo corretto.

Tra questi due estremi poi c'è una zona intermedia. Sono due possibili valori quelli scritti il punto nel nostro spazio a quattro dimensioni, due possibili valori possibili per la coordinata della prima dimensione, ma ci sono due valori intermedi tra questi estremi. Il primo è quello degli approcci ibridi, parzialmente pro attivi parzialmente reattivi. L'idea generale, ma non sempre è così, nasce dall'idea: si utilizza in prima battuta un approccio pro attivo, ma con un livello di precisione ridotto. Invece di mantenere aggiornata l'informazione sulla posizione esatta del nodo mobile, al massimo livello di dettaglio possibile, io mantengo aggiornata un'informazione un po' più grossolana, meno precisa, su una regione. Finché il nodo si muove nella regione non la aggiorno, la aggiorno solo quando il nodo valica i confini di questa regione e passa in una regione adiacente. Questa



informazione viene mantenuta aggiornata più raramente, meno informazione circola sulla rete ed abbasso il costo. Quando qualcuno vorrà parlare con quel nodo, reagisco e mi occupo di localizzare quel nodo in modo esatto all'interno della regione. E' una soluzione di compromesso tra i due estremi, cerca di sfruttare i vantaggi di entrambi. Questa ibridazione, scelta tra pro attiva e reattiva, può essere effettuata in modo statico, determinato dall'infrastruttura. In altri tipi di soluzione è realizzare questa ibridazione in maniera dinamica, adattativa, per cui la linea di confine tra la parte adattiva e pro attiva non la definisco in modo fisso ma lo faccio variare a seconda delle esigenze che si possono creare.