

ALMA MATER STUDIORUM · UNIVERSITÀ DI BOLOGNA

FACOLTÀ DI SCIENZE MATEMATICHE, FISICHE E NATURALI
Corso di Laurea in Scienze e Tecnologie Informatiche

**MAPPING DI DATI SPAZIALI
PER L'INTEGRAZIONE DI
RETI DI SENSORI**

Tesi di Laurea in Basi di Dati

Relatore:
Prof. Dario Maio

Presentata da:
Francesco Degli Angeli

Correlatore:
Ing. Luca Calderoni

Sessione III
Anno Accademico 2011-2012

Indice

Introduzione	1
1 La Georeferenziazione	7
1.1 Metodi di Georeferenziazione	8
1.1.1 Nomi	8
1.1.2 Indirizzi e codici postali	8
1.1.3 Sistema di riferimento lineare	9
1.1.4 Catasto	9
1.1.5 Misurare la Terra: latitudine e longitudine	10
1.1.6 Proiezioni e coordinate	12
1.1.7 Autocorrelazione, Campionamento e Interpolazione	13
1.1.8 Geolocalizzazione	17
1.1.9 Geocoding e Reverse Geocoding	17
1.2 I Dati Spaziali	18
1.2.1 Livello Concettuale	19
1.2.2 Livello della Rappresentazione	20
1.2.3 Livello dell'Implementazione	20
1.2.4 Il Modello Logico dei Dati Spaziali	25
1.2.5 Metodi di Rappresentazione dei Dati Spaziali	26
1.2.6 Le Relazioni Spaziali	27
1.2.7 Relazioni basate sulla direzione	29
1.2.8 Relazioni basate sulla distanza	29
1.2.9 Interrogazione di una Base di Dati Spaziale	30

1.2.10	Join spaziali	32
1.2.11	Indici spaziali	34
1.3	GIS	39
1.3.1	Storia dei GIS	40
1.3.2	Componenti di un GIS	41
1.3.3	Come lavora un GIS	43
1.3.4	Referenze geografiche	43
1.3.5	Tipi di dato utilizzati	44
1.3.6	Compiti del GIS	44
1.3.7	Tecnologie connesse	47
2	Le Reti di Sensori	61
2.1	La Storia delle Reti di Sensori	63
2.2	Classificazione delle Reti di Sensori e Campi di Applicazione .	64
2.3	La Struttura delle Reti di Sensori	70
2.4	Scelte tecnologiche nella progettazione di una Rete di Sensori .	72
2.5	Esempio pratico di una Rete di Sensori: Minos System	78
2.5.1	SYRA	80
2.5.2	ANDROS	81
3	Parte Implementativa	83
3.1	Prima Fase: Modifiche Strutturali	84
3.2	Seconda Fase: Inserimento coordinate per ANDROS e SYRA .	86
3.3	Terza Fase: Mapping dei Sottosistemi	86
3.4	Quarta Fase: Interrogazioni di Dettaglio su Singolo Sensore . .	88
	Conclusioni	91
	Bibliografia	93
	Ringraziamenti	99

Elenco delle figure

1.1	Sistema di riferimento lineare	9
1.2	Definizione di latitudine	11
1.3	Esempio di autocorrelazione spaziale	16
1.4	Livelli di astrazione dei Dati Spaziali	19
1.5	Rappresentazione di spezzate in formato vettoriale	21
1.6	Rappresentazione di un poligono in formato vettoriale	21
1.7	Rappresentazione di una regione in formato vettoriale	22
1.8	Celle di dimensione variabile	24
1.9	Celle di dimensione fissa	24
1.10	Insieme di liste di coordinate	27
1.11	Struttura basata sulla topologia dei dati	28
1.12	Relazione basata sulla direzione	29
1.13	Esempio Range query	30
1.14	Esempio di selezione basata sulla direzione	31
1.15	Esempio di selezione basata sulla distanza	31
1.16	Join topologico	32
1.17	Join basato sulla direzione	32
1.18	Join basato sulla distanza	33
1.19	Esempio di Quad-Tree	35
1.20	Esempio di R-tree	37
1.21	Esempio di R-tree 2	38
1.22	Esempio di Directory Rectangle	38
1.23	Architettura di un GIS	39

1.24	Processo di trilaterazione di un ricevitore GPS	51
1.25	Trilazione 3D	53
2.1	Struttura tipica di una WSN	62
2.2	Classificazione delle reti in funzione dell'area coperta	64
2.3	Possibili topologie di rete nelle WSN	70
2.4	Sistema Centralizzato e Sistema Distribuito	71
2.5	Rappresentazione del multi-hop	77
2.6	Gerarchia <i>Minos System</i>	79
2.7	Tipologia di SYRA	80
2.8	Tipologia di ANDROS	81
3.1	Risultato popolamento campo GENITORE	85
3.2	Risultato verifica assegnamento coordinate ai sensori SYRA.	87
3.3	Tabella scelte assegnamenti coordinate	88

Introduzione

Il progresso in campo informatico è in continua evoluzione. Da tempo è possibile associare ad un dato informatico una coppia di coordinate geografiche; tale tecnica è detta georeferenziazione e presenta requisiti di unicità nel riferimento, in quanto a una data georeferenziazione è associato un solo luogo, e persistenza nel tempo, che permette di evitare problemi di inconsistenza dei dati. Negli anni si sono susseguiti svariati metodi di georeferenziazione: dal dare nomi alle località, la forma più semplice, antica e ancora oggi utilizzata, all'introduzione nel XIX secolo di indirizzi e codici postali, fino alla moderna geolocalizzazione, ovvero l'identificazione della posizione geografica nel mondo reale di un oggetto (come ad esempio smartphone, tablet, pc), in un dato momento, grazie a reti GSM, a sistemi WPS, a dispositivi GPS o alla stessa Rete Internet. In questo ambito si collocano i Dati Spaziali cioè dati che descrivono oggetti o fenomeni ai quali sia associata una dimensione spaziale, fornendo informazioni sulla posizione fisica e sulla forma geometrica. Per chiarire meglio il concetto di dato spaziale occorre effettuare un'astrazione dell'informazione al fine di eliminare i particolari irrilevanti e ridurre il costo di memorizzazione ed elaborazione dei dati. A livello concettuale si identificano essenzialmente due schemi per rappresentare i dati geografici: *discrete object view*, detto anche visione ad oggetti, secondo cui gli oggetti sono definiti attraverso la loro dimensione; *field view*, detto anche visione a campi, secondo cui il mondo geografico è descritto attraverso un insieme di variabili, o campi, misurabili all'interno della superficie terrestre, i cui valori, muovendosi, cambiano nello spazio. Questi modelli concettuali non

si riferiscono a come l'informazione geografica possa essere rappresentata in maniera digitale, ma sono alla base di due metodi che permettono a strumenti preposti di ridurre i fenomeni geografici in forme che possono essere codificate in un database: il modello vettoriale e il modello raster. A livello implementativo, nel modello vettoriale, lo spazio continuo viene approssimato costruendo strutture dati appropriate: le informazioni su punti, linee e poligoni sono codificate e memorizzate come una collezione di coordinate x , y . Questo modello risulta estremamente utile per descrivere fenomeni discreti, ma meno adatto per descrivere fenomeni continui, quali temperatura, precipitazioni, quota, pendenza, cioè fenomeni che rappresentano un'unica grandezza che varia continuamente nello spazio. Il modello raster si è sviluppato proprio per descrivere tali fenomeni: lo spazio geografico è suddiviso in celle, di dimensione fissa o variabile, a cui sono associati un insieme di dati e ogni variazione geografica è espressa assegnando proprietà o attributi a queste celle. Per trattare i Dati Spaziali a livello logico vengono introdotti nuovi tipi di dato, gli SDT (*Spatial Data Type*), e un modello, detto GEO-relazionale, che organizza le tabelle in database relazionali connettendole attraverso identificativi di oggetti spaziali. Come per il modello logico, occorre introdurre nuovi costrutti anche per le interrogazioni che riguardano questo tipo di dato, dette interrogazioni spaziali, utilizzando un'estensione dell'algebra relazionale, detta geoalgebra. Queste interrogazioni si possono dividere essenzialmente in 2 categorie: selezioni basate su relazioni spaziali rispetto a valori geometrici costanti, ovvero interrogazioni basate sulle relazioni di ogni valore geometrico con lo spazio di riferimento; e join spaziali, ovvero interrogazioni basate sulle relazioni dei valori geometrici contenuti nella base di dati. Un accesso efficiente ai dati è essenziale anche per le basi di dati spaziali. Le strutture d'accesso sviluppate per indicizzare i dati numerici o alfanumerici tradizionali si basano sull'ordinamento totale del dominio del campo chiave ma non si adattano ai dati spaziali. Quindi occorrono nuove strutture di indicizzazione per questa tipologia di dato: *space-driven* e *data driven*. Le prime sono strutture in cui lo spazio di riferimento è suddiviso

in celle rettangolari; fanno parte di questa categoria i Quad-Tree, ovvero strutture dati ad albero non bilanciate nella quali tutti i nodi interni hanno esattamente quattro nodi figli. Le seconde sono strutture organizzate partizionando l'insieme degli oggetti in base alla loro distribuzione nello spazio di riferimento. La struttura originaria di questa tipologia di indice è quella del R-tree: lo spazio viene diviso in descrittori innestati gerarchicamente e quando possibile sovrapposti, approssimando il dato spaziale con un rettangolo o con un poligono convesso. Un *Geographic Information System*, o più semplicemente GIS, è un sistema informativo computerizzato che permette l'acquisizione, la registrazione, l'analisi, la visualizzazione e la restituzione di informazioni derivanti da dati spaziali. Venne prodotto dalla *Canada Land Inventory* in Canada alla metà degli anni 60 come un sistema computerizzato di misurazione di mappe. Negli anni 70 i GIS raggiunsero discrete capacità di svolgere analisi ed elaborazioni di dati ma presentavano ancora carenze nella rappresentazione grafica. Tra gli anni 80 e 90, grazie anche ad un continuo miglioramento a livello hardware e software, ci fu una diffusione di questa tecnologia nelle piccole P.A. (Pubblica Amministrazione), nella sanità e a livello commerciale con l'avvento dei *personal computer*. Attualmente i GIS di ultima generazione sono in grado di gestire a livello implementativo due formati di dato spaziale: il formato *vettoriale* e il formato *raster*. Infatti si tende ad integrare le due tipologie nel miglior modo possibile, permettendo, in strutture omogenee di archivi, la convivenza di dati, l'elaborazione interconnessa e la conversione dei formati. I GIS sono strettamente connessi a molti altri tipi di tecnologie: Desktop GIS, un sistema *desktop mapping* che utilizza le mappe per organizzare i dati ed interagire con l'utente; CAD (*Computer Aided Design*), sistemi sviluppati per creare disegni, progetti architettonici ed infrastrutture; GNSS (*Global Navigation Satellite System*), un sistema di geo-radiolocalizzazione e navigazione, terrestre o aerea, che utilizza una rete di satelliti artificiali in orbita e pseudoliti. Fra i GNSS il più utilizzato è il GPS, un sistema di posizionamento e navigazione satellitare che consente ad un terminale mobile di ricavare la propria posizione sul globo terrestre, in

termini di coordinate geografiche.

Anche nel campo delle Reti di Sensori i progressi sono costanti. Dette WSN (*Wireless Sensor Networks*), sono reti informatiche formate da piccoli dispositivi autonomi, denominati nodi sensore (*Sensor Node*), in grado di prelevare dati dall'ambiente circostante e di comunicare tra loro. Questi dispositivi sono formati da componenti in grado di rilevare grandezze fisiche (come temperatura, pressione, intensità sonora, vibrazioni, umidità, pressione atmosferica, ecc.), di elaborare dati e di interagire tra loro. I sensori hanno le seguenti peculiarità: sono caratterizzati da dimensioni e pesi molto ridotti; sono prodotti in massa per via di un costo trascurabile; hanno una riserva d'energia limitata e non rinnovabile; mantengono costantemente i consumi molto bassi; riescono ad auto apprendere la loro posizione grazie ad algoritmi di geolocalizzazione. I nodi sensore sono sparpagliati in un'area chiamata *area di sensing* in cui ciascun nodo ha la capacità di accumulare e di instradare i dati fino al nodo *sink*, un nodo di raccolta informazioni, ed infine all'utente, o server, finale. Queste reti possono essere classificate generalmente in funzione dell'area coperta oppure in maniera più dettagliata in base alla finalità della rete. Le prime reti di sensori furono sviluppate negli Stati Uniti per scopi militari durante il periodo della guerra fredda. Ad oggi sono utilizzate per molteplici applicazioni grazie alla nascita di sistemi come *Bluetooth*, *WiFi*, *Wimax* e *ZigBee*, che consentono una connettività totale, e nuovi processori, caratterizzati da un basso costo e un minimo consumo energetico. Un esempio di Reti di Sensori è il "*Minos System*", il sistema più avanzato di telegestione dell'illuminazione pubblica ed esterna, progettato dall'azienda *UMPI Elettronica S.r.l.* Questo impianto rientra nella categoria delle "*Wired Sensor Network*", cioè reti di sensori dotate di un canale di comunicazione ad onde convogliate, in cui ogni lampada viene munita di un dispositivo di controllo e diagnostica, chiamato SYRA, facenti capo ad un dispositivo armadio, chiamato ANDROS. Normalmente il sistema è composto da più ANDROS, ciascuno dei quali è in grado di controllare fino a 1.022 SYRA. I vantaggi portati da questa tecnologia sono un massiccio risparmio

energetico, una notevole riduzione dei costi di manutenzione, una diminuzione dell'inquinamento atmosferico e la trasformazione del lampione in un punto intelligente in grado di erogare servizi aggiuntivi (videosorveglianza, info-mobilità e comunicazione WI-FI a basso costo).

L'obiettivo di questa tesi è l'analisi di database che contengono Dati Spaziali riferiti a Reti di Sensori e la creazione di un mapping dei vari sottosistemi di sensori per ottimizzare eventuali interrogazioni, in collaborazione con l'azienda *UMPI Elettronica S.r.l.*. In particolare questi tre database contengono rispettivamente: informazioni relative ai sensori a livello armadio (ANDROS) e a livello lampada (SYRA) del sistema di controllo Minos System, informazioni relative a sensori per il controllo del traffico, informazioni per il mapping. La memorizzazione e l'organizzazione dei dati è realizzata con *DBMS MySQL* mentre la fruizione dei dati e le operazioni sugli stessi sono garantiti da un software *PHP Server Side* sviluppato dalla *UMPI Elettronica S.r.l* e la cui discussione non rientra negli obiettivi di questa tesi.

La tesi è strutturata come segue.

Nel *primo capitolo* vengono descritti prima peculiarità e metodi della georeferenziazione, seguiti dalle prerogative dei Dati Spaziali. Inoltre viene presentato un sistema informativo computerizzato che lavora con questo tipo di dati geografici, il GIS.

Nel *secondo capitolo* vengono delineate le Reti di Sensori illustrando anche un esempio pratico di questa tipologia di reti, il "*Minos System*".

Nel *terzo capitolo* vengono presentate le quattro fasi di lavoro su un sottoinsieme della base dati della *UMPI Elettronica S.r.l* con relative soluzioni adottate.

Infine vengono illustrate le conclusioni scaturite dal lavoro di tesi analizzando alcuni possibili sviluppi futuri.

Capitolo 1

La Georeferenziazione

Si definisce georeferenziazione l'atto di assegnare una localizzazione ad atomi di informazione. Questa tecnica permette di associare ad un dato una coppia di coordinate che ne fissino la posizione sulla superficie terrestre, secondo un particolare sistema geodetico di riferimento [GEOREF].

La georeferenziazione presenta i seguenti requisiti: unicità nel riferimento e persistenza nel tempo. La prima peculiarità esige quindi un'accezione univoca, in modo da associare a una data georeferenziazione un solo luogo (geografico); inoltre il suo significato deve essere comune a tutte le persone che saranno interessate a questa informazione, inclusi i loro GIS (*Geographic Information System*) che verranno trattati nella terza sezione di questo capitolo. Per esempio, l'indirizzo "via Sacchi 3, 47521 Cesena (FC) Italia" si riferisce ad un solo edificio e ciò è sufficientemente chiaro a chiunque voglia spedire una raccomandata da ogni parte del mondo. La seconda peculiarità richiede un'accezione persistente nel tempo, in modo da evitare cambiamenti frequenti che causerebbero problemi di natura confusionale e, in taluni casi, obbligherebbero alla rettifica di record associati alla georeferenziazione modificata. Inoltre queste modifiche (ad esempio variazioni nei nomi delle vie, delle città o dei codici di avviamento postale) potrebbero creare problemi alle agenzie di mappatura e ad eventuali ricercatori.

1.1 Metodi di Georeferenziazione

1.1.1 Nomi

Dare nomi alle località è la forma più semplice di georeferenziare, e forse una delle più antiche ad essere utilizzata. Ogni caratteristica distintiva di un paesaggio, come un albero particolare, o una collina, può essere usato come punto di riferimento per due persone che vogliono scambiarsi un'informazione. Le zone abitate poi sono state rapidamente riempite di targhe che servivano ad orientarsi. Inoltre, il linguaggio può estendere le potenzialità di questo sistema attraverso parole come “tra” oppure “vicino”. Ma questo metodo è molto limitativo. Ad esempio dire che qualcosa si trova in Europa è molto generico, dato che l'Europa copre più di 10 milioni di km^2 . Inoltre ci sono molti nomi che sono conosciuti da piccole comunità e che non possono essere condivisi con il resto del mondo [1].

1.1.2 Indirizzi e codici postali

Introdotti con la diffusione della distribuzione della posta nel XIX secolo, gli indirizzi postali si basano su diverse considerazioni. Poiché ogni casa o ufficio è un potenziale destinatario di posta, vengono divisi e numerati in sentieri o strade secondo la seguente disposizione: nomi “localmente” unici per sentieri e strade; nomi unici dentro regioni più ampie per aree locali; infine nomi unici all'interno dei paesi per le regioni. Se valgono queste ipotesi, allora la posta fornisce un'univoca identificazione per ogni edificio terrestre. Chiaramente questo sistema è molto legato alle attività umane. Ad esempio, il monte Bianco non ha un indirizzo di posta. Inoltre ci sono delle culture che non prevedono la numerazione consecuzionale delle strade (ad esempio in Giappone, dove la numerazione riflette la data di costruzione, quindi è di tipo temporale e non spaziale). Molti GIS convertono le coordinate postali, in altri tipi di coordinate, come latitudine e longitudine. Così per semplificare

l'ordinamento e la distribuzione della posta sono stati introdotti nel XX secolo i codici postali [1].

1.1.3 Sistema di riferimento lineare

Il sistema di riferimento lineare definisce una posizione misurando la distanza da un punto di riferimento iniziale e seguendo un percorso definito sulla rete.

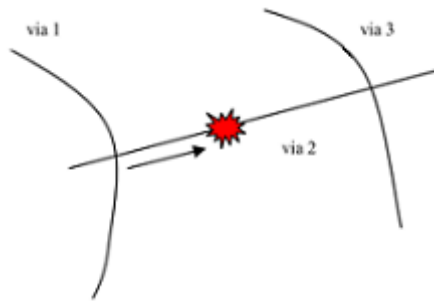


Figura 1.1: Sistema di riferimento lineare (tratto da [1])

Ad esempio, si può localizzare l'incidente nella Figura 1.1, misurando la distanza sulla via 2, a partire da un punto di riferimento noto, che è l'intersezione della via 2 con la via 1. Questo sistema è molto vicino a quello utilizzato nella numerazione delle strade, ma al contrario di quest'ultimo è prevista una misurazione delle distanze [1].

1.1.4 Catasto

Il catasto è definito come una mappa delle proprietà dei terreni in un'area, utilizzato allo scopo di tassare quei terreni, o per creare un registro pubblico delle proprietà. Il processo di suddivisione crea nuove particelle (catastali) suddividendo quelle esistenti. Le particelle in un catasto sono univocamente determinate, attraverso un numero o un codice, e sono anche ragionevolmente persistenti nel tempo, quindi soddisfano i requisiti di un sistema di

georeferenziazione. Tuttavia molto spesso i proprietari, o gli utenti in genere, non conoscono questa numerazione, e quindi l'utilizzo di questo sistema è confinato ad una utenza tecnica [1].

1.1.5 Misurare la Terra: latitudine e longitudine

I più potenti sistemi di georeferenziazione sono quelli che forniscono la possibilità di avere una risoluzione spaziale molto fine, che permette di calcolare le distanze fra coppie di punti, e supporta altre forme di analisi spaziali. Questo sistema è per molti versi il più esaustivo, ed è spesso chiamato sistema geografico di coordinate, basato sulla rotazione della Terra intorno al suo centro di massa. Per definire latitudine e longitudine dobbiamo per prima cosa identificare l'asse della rotazione terrestre. Il centro di massa della Terra cade proprio in questo asse, e il piano passante dal centro di massa e perpendicolare all'asse definisce l'Equatore. Le "fette" di Terra parallele all'asse e perpendicolari all'Equatore definiscono le linee di uguale longitudine, come i segmenti degli spicchi di arancia. Una di queste linee, che convenzionalmente passa dal *Royal Observatory di Greenwich* (in Inghilterra), definisce lo zero della longitudine, e l'angolo che questa linea forma con qualsiasi altra nel globo è la misura della longitudine. Ognuno dei 360 gradi di longitudine è suddiviso in 60 minuti e ognuno di questi in 60 secondi. Ma è molto più convenzionale usare la misura della longitudine con i gradi Est od Ovest, quindi la longitudine va da 180° ovest a 180° est. Inoltre, poiché i dati numerici vengono suddivisi equamente nei computer tra positivi e negativi, i gradi ovest vengono considerati negativi e quelli est positivi, usando i decimali piuttosto che i minuti e i secondi. Una linea di costante longitudine viene detta meridiano. La longitudine può essere definita in questo modo per qualsiasi solido di rotazione, al di là della sua forma, perché l'asse di rotazione e il centro di massa sono sempre definiti. Ma la definizione di latitudine presuppone che sappiamo qualcosa sulla forma. La Terra ha una forma complessa che è soltanto approssimativamente una sfera. Una rappresentazione più corretta è un ellissoide di rotazione, cioè la figura che si ottiene prendendo un'ellisse e

ruotandola attorno all'asse minore. La differenza tra un ellissoide e una sfera è misurata dal suo appiattimento, o la riduzione nell'asse minore dell'asse maggiore. L'appiattimento è definito dalla formula seguente:

$$f = (a - b) / a \quad (1.1)$$

dove a e b sono le lunghezze dell'asse maggiore e minore rispettivamente. La Terra è leggermente appiattita (1/300).

Affinchè le mappe fatte dai singoli paesi su ellissoidi diversi combaciassero nei loro confini e per evitare correzioni, l'ellissoide conosciuto come WGS84 (*World Geodetic System* del 1984) è adesso largamente accettato come unico standard. La Figura 1.2 mostra una linea disegnata lungo un punto di interesse perpendicolare all'ellissoide in quella posizione. L'angolo formato da questa linea con il piano dell'Equatore è definito come la latitudine del punto, e varia da 90° sud a 90° nord. Anche qui le latitudini sud sono usualmente memorizzate come numeri negativi e quelle nord come numeri positivi. Una linea di latitudine costante è detta parallelo.

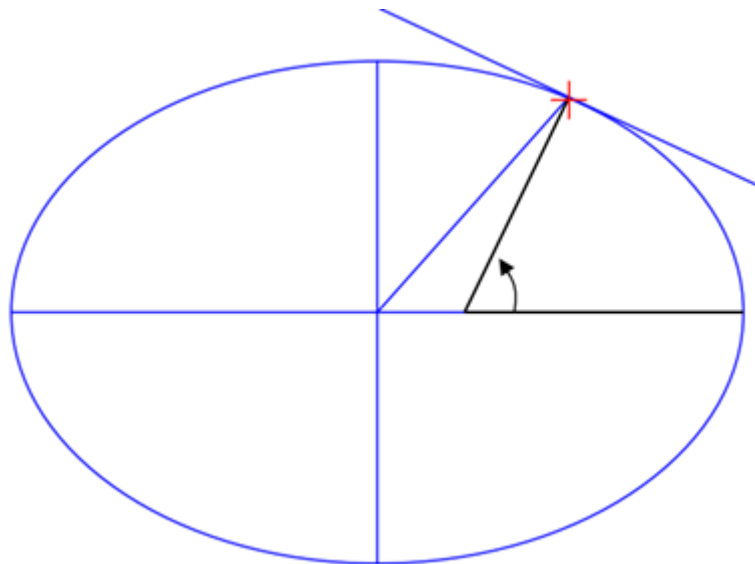


Figura 1.2: Definizione di latitudine del punto rosso, come l'angolo tra l'Equatore e la linea perpendicolare all'ellissoide (tratto da [1])

La latitudine è normalmente simbolizzata con la lettera greca *phi* (ϕ) e la longitudine con la lettera greca *lambda* (λ); i rispettivi range possono essere espressi matematicamente con $-180 \leq \lambda \leq 180$ e $-90 \leq \phi \leq 90$. Dati latitudine e longitudine è possibile calcolare la distanza fra una qualsiasi coppia di punti. Supponendo che la Terra sia sferica (per semplificare la formula), la distanza fra due punti sulla superficie è data dalla lunghezza dell'arco che li congiunge, ottenuto disegnando un cerchio che taglia la sfera e passa per i due punti e per il centro della sfera stessa. La lunghezza di quest'arco su una Terra sferica di raggio R è data da:

$$R \cdot \cos^{-1}[\sin\phi_1 \cdot \sin\phi_2 + \cos\phi_1 \cdot \cos\phi_2 \cdot \cos(\lambda_1 - \lambda_2)] \quad (1.2)$$

dove gli indici denotano i due punti. Per esempio, la distanza di un punto sull'Equatore alla longitudine 90° Est (nell'Oceano Indiano tra lo Srilanka e l'isola indonesiana di Sumatra) e il Polo Nord si trova calcolando la formula per $\phi_1 = 0$, $\lambda_1 = 90$, $\phi_2 = 90$ e $\lambda_2 = 90$. Spesso si usano i radianti (1 radiante è 57.30 gradi e 90 gradi sono $\pi/2$ radianti) [1].

1.1.6 Proiezioni e coordinate

Latitudine e longitudine definiscono il luoghi della terra in termini di angoli con riferimenti ben definiti: il Royal Observatory di Greenwich (UK), il centro di massa e l'asse di rotazione. Per questo costituiscono il sistema più completo di georeferenziazione, e supportano una certa quantità di possibilità di analisi, inclusa la possibilità di calcolare la distanza tra due punti sulla superficie curva della terra. Ma molte tecnologie per lavorare con i dati geografici si riferiscono a una geometria piana, come per esempio la carta e le stampe, e si sono evolute secoli prima della nascita dei GIS. Per varie ragioni, quindi, molto lavoro nei GIS ha a che vedere con una Terra piatta o proiettata, nonostante il prezzo che si paga nella distorsione conseguente all'appiattimento.

Specificamente, la Terra è spesso appiattita per le seguenti motivazioni:

- la carta è piatta, ed è usata come mezzo di input e di output dei dati geografici;
- i dati raster sono intrinsecamente piatti, poiché è impossibile ricoprire una superficie curva con quadrati eguali senza sovrapposizioni;
- la pellicola fotografica è piatta (ad es. le immagini aeree o da satellite);
- quando la Terra viene vista dallo spazio, la parte nel centro dell'immagine ha il maggior dettaglio, allontanandosi dal centro si perde dettagli, fino a perdere la possibilità di vedere la parte posteriore; per vedere la Terra completa con approssimativamente lo stesso dettaglio, deve essere distorta in un qualche modo, e il più conveniente è quello di appiattirla.

Le coordinate cartesiane assegnano due coordinate a tutto i punti di una superficie piana, misurando la distanza da un'origine [1].

1.1.7 Autocorrelazione, Campionamento e Interpolazione nei Dati Spaziali

“Everything is related to everything else, but near things are more related than distant things [18].”

La citazione precedente riporta la prima legge della geografia di Tobler e dice che tutto è correlato con qualsiasi altra cosa, ma le cose vicine sono più relazionate di quelle lontane. Formalmente, questa proprietà è conosciuta come autocorrelazione, che dà la misura di quanto le cose sono tra loro correlate. Nell'analisi dei dati spaziali, il processo che fa sorgere l'autocorrelazione spaziale può essere bidimensionale e tridimensionale. Grazie a questo processo si può avere una rappresentazione più precisa dei dati spaziali ad esempio generalizzando a partire da un campione; inoltre misura il tentativo di mettere assieme simultaneamente le similarità nella localizzazione degli oggetti spaziali e nei loro attributi. Se a una similarità spaziale corrisponde

una similarità negli attributi si parla di autocorrelazione positiva. Contrariamente, si dice che c'è una autocorrelazione negativa quando oggetti molto vicini nello spazio tendono ad essere molto distanti nei valori degli attributi. Un'autocorrelazione nulla c'è quando c'è l'indipendenza dagli attributi dalla localizzazione. Tuttavia spesso per favorire l'autocorrelazione spaziale è necessario avere una selezione accurata e efficace di un campione poiché la sua precisione è influenzata soprattutto dalla scala. Si può pensare al campionamento spaziale come al processo di selezione di punti da un campo continuo oppure, se il campo è rappresentato da un insieme di oggetti, selezionando alcuni di questi oggetti e tralasciando gli altri. Quindi si può dire che ogni rappresentazione spaziale viene fuori da un processo di campionamento. Ad esempio nelle immagini da satellite, dove i pixel rappresentano porzioni di territorio con un'altitudine media. Se lo schema di campionamento è dettagliato, i dati geografici saranno di elevata qualità, e sarà possibile inferire, con metodi statistici, le proprietà geografiche complessive dal campione.

Seguono diversi metodi di campionamento:

- campionamento per randomizzazione semplice (o campionamento casuale semplice) : utilizzato di frequente, può risultare inefficace per una serie “sfortunata” di numeri;
- campionamento per randomizzazione sistematica (o campionamento sistematico) : viene scelto un intervallo k , che è l'inverso del rapporto tra la popolazione totale N e la dimensione n del campione, in modo da creare una griglia uniforme sullo spazio di riferimento;
- campionamento per randomizzazione stratificata (o campionamento stratificato) : viene creata sempre una griglia, all'interno della quale il campionamento viene fatto random;
- campionamento a grappolo: si selezionano dei gruppi di unità dette *cluster* che sono dei dati individuati artificialmente sempre su base geografica.

Questi metodi vanno adattati alle varie situazioni e comunque un campione più grande genera una rappresentazione migliore della realtà in genere. Una volta ottenuto il campione, è necessario “riempire” i vuoti tra i suoi elementi per dare una rappresentazione più continua della realtà geografica. Questo richiede la conoscenza degli effetti della distanza dai campioni e si rende necessario l'utilizzo di funzioni di interpolazione. In molte applicazioni, la scelta di una funzione rispetto a un'altra è frutto dell'esperienza, dell'adattamento a un particolare dataset, e di una convenzione. In termini matematici, prendiamo b come parametro che condiziona la velocità alla quale il peso w_{ij} decresce con la distanza d_{ij} . In questi modelli lo spazio si suppone isotropo (non ci sono dimensioni privilegiate rispetto ad altre). Le funzioni più usate sono le seguenti:

$$\textit{Lineare} : w_{ij} = -b \cdot d_{ij} \quad (1.3)$$

$$\textit{Potenza} : w_{ij} = d_{ij}^{-b} \quad (1.4)$$

$$\textit{Esponenziale} : w_{ij} = \exp(-b \cdot d_{ij}) \quad (1.5)$$

L'autocorrelazione spaziale (Figura 1.3) riguarda due tipi di informazione: la similarità tra gli attributi e la similarità nella localizzazione. I modi in cui gli attributi possono essere misurati dipendono dai dati presenti, mentre il calcolo della prossimità spaziale dipende dal tipo degli oggetti. Grazie a queste funzioni si possono confrontare i valori degli attributi spaziali.

Si definisce una matrice W in cui ogni elemento w_{ij} misura la similarità¹ delle localizzazioni i e j (righe e colonne della matrice) e si pone $w_{ij} = 0$ per ogni i . Ponendo $w_{ij} = 1$ se le regioni i e j sono contigue e $w_{ij} = 0$ altrimenti, possiamo ottenere una matrice di pesi (denotata in genere con W).

Per semplificare le cose, usiamo la seguente notazione:

¹similarità: per oggetti con un'area, la contiguità, o l'adiacenza, viene spesso presa come la più tangibile misura di prossimità spaziale.

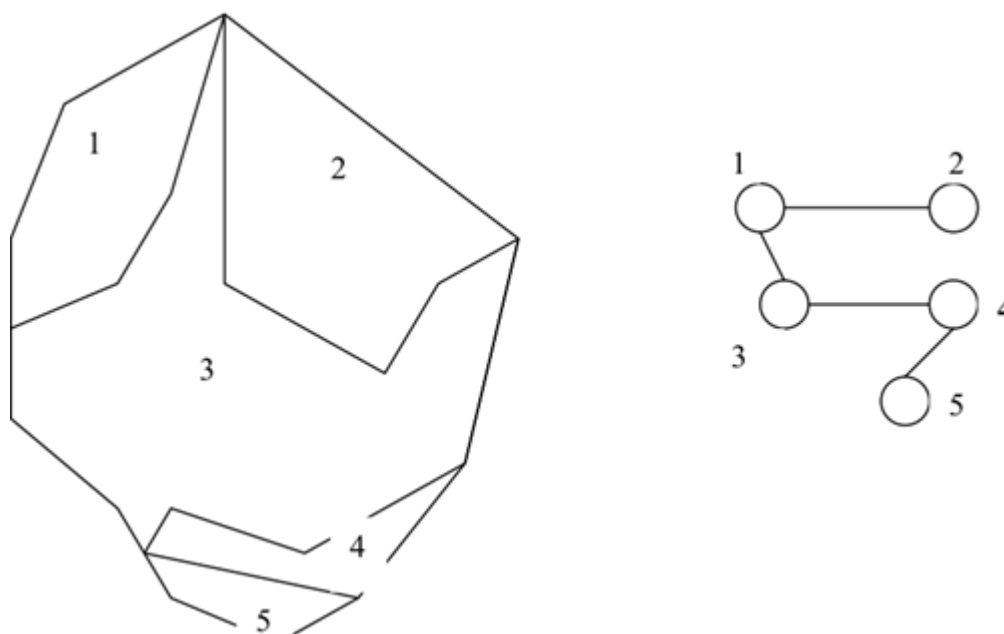


Figura 1.3: Esempio di autocorrelazione spaziale (tratto da [1])

- n il numero di oggetti nel campione
- i, j due diversi oggetti
- z_j il valore dell'attributo di interesse per l'oggetto i
- c_{ij} la similarità degli attributi di i e j
- w_{ij} la similarità delle localizzazioni di i e j , con $w_{ii} = 0$ per ogni i .

In generale, ogni misura di autocorrelazione spaziale è stata creata per comparare l'insieme di similarità degli attributi c_{ij} con l'insieme delle similarità posizionali w_{ij} , combinandole in un singolo indice sotto forma del seguente prodotto in croce:

$$\sum_i \sum_j (c_{ij} \times w_{ij}) \quad (1.6)$$

In altre parole, il totale è ottenuto moltiplicando ogni cella della matrice W con la sua corrispondente in C , e sommando [1].

1.1.8 Geolocalizzazione

La geolocalizzazione è l'identificazione della posizione geografica (latitudine e longitudine) nel mondo reale di un oggetto, come ad esempio smartphone, tablet, pc, in un dato momento. Questi dispositivi sono dotati di particolari strumenti che gli permettono di stabilire la loro posizione. Le tecnologie più utilizzate si suddividono tra sistemi di localizzazione basati su segnale radio (radiolocalizzazione), sistemi per via cablata o sistemi ibridi:

- Localizzazione tramite le celle della rete telefonica cellulare: si analizza la potenza del segnale radio di ogni cella telefonica in relazione alla rispettiva stazione radio base (che ha coordinate geografiche note) collegata con il dispositivo mobile o terminale in modo da determinarne, dopo operazioni di triangolazione matematica, la posizione sul territorio. Questo processo è alla base della rete GSM (*Global System for Mobile Communications*).
- Localizzazione tramite la rete *Wi-Fi* o *WLAN*: è basata sul segnale delle diverse fonti dette *access point Wi-Fi*, sparse nel territorio ma di cui si conosce l'esatta ubicazione, le quali vengono a loro volta localizzate tramite la Rete Internet. Questo sistema, detto anche WPS (*Wireless Positioning System*), permette di determinare la posizione del dispositivo con un'accuratezza di poche decine di metri.
- GPS trattato nella sottosezione 1.3.7 *Tecnologie connesse*.
- Rete Internet: tramite l'indirizzo IP della propria rete Internet si rileva l'area approssimativa di localizzazione; ciò dipende dalla registrazione del collegamento ad un provider che ha un range di indirizzi suddivisi in sottogruppi per determinare stati, regioni, province [GEOLOC].

1.1.9 Geocoding e Reverse Geocoding

Geocoding è il processo di conversione di dati geografici, come indirizzi stradali o codici di avviamento postale, in coordinate geografiche (spesso

esprese in latitudine e longitudine). Per esempio Google Maps è un sistema che utilizza questo processo per georeferenziare gli elementi. Infatti è possibile cercare negozi o località di interesse dei quali vengono fornite, non solo le tipiche informazioni che restituisce un motore di ricerca, ma viene evidenziato sulla mappa la posizione geografica ad essi riferita. Le API di Google Maps forniscono la classe geocoder che con facilità permette d'implementare questa funzionalità [GEOC].

Reverse Geocoding è il processo di ritorno (*reverse*) di una posizione del punto (latitudine, longitudine) ad un indirizzo leggibile. Ciò permette l'identificazione di indirizzi stradali nelle vicinanze, luoghi e / o le loro aree, come quartieri, contee, stati o paesi. In combinazione con geocoding e routing services, questo processo è alla base della applicazioni *mobile location-based services* in quanto permette di convertire le coordinate ottenute dal GPS in un indirizzo leggibile per l'utente finale [RGEOC].

1.2 I Dati Spaziali

Un dato spaziale è un qualunque dato che descrive oggetti o fenomeni ai quali sia associata una dimensione spaziale, fornendo informazioni sulla posizione fisica e sulla forma geometrica [3]. Questi oggetti possono rappresentare dei punti oppure forme più complesse come paesi, strade o laghi. I modelli dei dati spaziali hanno requisiti sostanzialmente diversi dalle applicazioni tradizionali anche perché estendono di fatto tali sistemi. Infatti sono dati con estensione e locazione spaziale che danno diverse possibili rappresentazioni della realtà (dualità della visione ad oggetti e della visione a campi) e delle relazioni spaziali fra oggetti. Poiché la rappresentazione dei dati spaziali fornisce una visione semplificata del mondo reale occorre effettuare un'astrazione dell'informazione (Figura 1.4) al fine di eliminare i particolari irrilevanti e ridurre il costo di memorizzazione ed elaborazione dei dati [11].



Figura 1.4: Livelli di astrazione dei Dati Spaziali (tratto da [4])

1.2.1 Livello Concettuale

Si specificano le classi di oggetti di base che verranno create nella base di dati e si descrivono le diverse rappresentazioni del mondo reale attraverso due schemi di natura concettuale per rappresentare i dati geografici: il *discrete object view* e il *field view*. Secondo il primo schema, detto anche visione ad oggetti, lo spazio è vuoto, a meno che non sia occupato da oggetti ben definiti, che sono istanze di categorie riconosciute. Gli oggetti geografici sono definiti attraverso la loro dimensione: sono detti bi-dimensionali gli oggetti che occupano un'area; sono detti uni-dimensionali oggetti più simili a linee come le strade o le ferrovie; infine sono detti zero-dimensionali oggetti rappresentati come punti nello spazio di riferimento. In diversi campi si riscontra l'utilità di questo schema come ad esempio nella rappresentazione degli oggetti che fanno parte e compongono una città (edifici, strade, ferrovie ecc.). Ma ci sono altri fenomeni geografici che non vengono bene rappresentati da questo schema; infatti un problema della *discrete object view* è che per la rappresentazione degli oggetti è necessario definire bene le loro proprietà caratterizzanti. In questi è necessario basarsi su un altro schema, detto

field view o visione a campi, che descrive il mondo geografico attraverso un insieme di variabili, ognuna delle quali misurabile all'interno della superficie terrestre, vista come una superficie continua, e i cui valori, muovendosi, cambiano nello spazio. Gli oggetti sono definiti dalla loro dimensione, mentre i campi sono definiti da cosa varia e con quale rapidità [11].

1.2.2 Livello della Rappresentazione

Questi modelli concettuali non si riferiscono a come l'informazione geografica possa essere rappresentata in maniera digitale, ma sono alla base di due metodi che permettono a strumenti, come i moderni GIS, di ridurre i fenomeni geografici in forme che possono essere codificate in un database: il modello vettoriale e il modello raster.

1.2.3 Livello dell'Implementazione

Modello vettoriale

Nel modello vettoriale, lo spazio continuo viene approssimato costruendo strutture dati appropriate: le informazioni su punti, linee e poligoni sono codificate e memorizzate come una collezione di coordinate x, y . Secondo la visione ad oggetti i dati vettoriali si distinguono in:

- punti (pozzi, pali, alberi, località, città ...) rappresentati da una singola coppia di coordinate;
- spezzate (curve di livello, fiumi, strade ...), rappresentate da una lista di punti e memorizzate come una stringa di coordinate x, y (Fig 1.5);
- poligoni (confini amministrativi, zone di vendita, bacini idrici ...) rappresentati da una spezzata chiusa (ogni poligono con n vertici ammette $2n$ possibili rappresentazioni) e memorizzati come aree omogenee perimetrale in cui la prima e l'ultima coordinata hanno lo stesso valore (Fig 1.6);

- regioni formate da un insieme di poligoni (Fig 1.7).

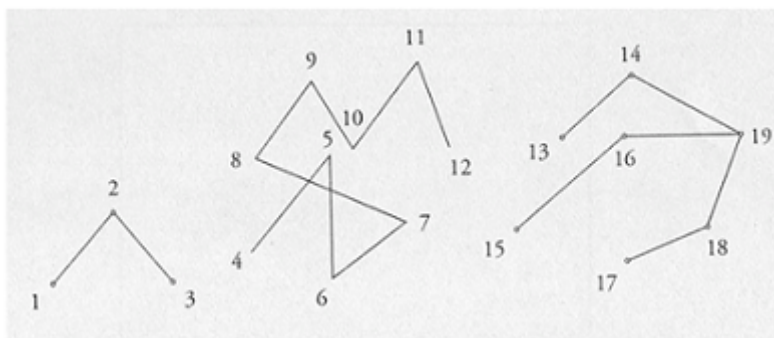


Figura 1.5: Rappresentazione di spezzate in formato vettoriale (spezzata semplice, spezzata non semplice, insieme di spezzate) (tratto da [4])

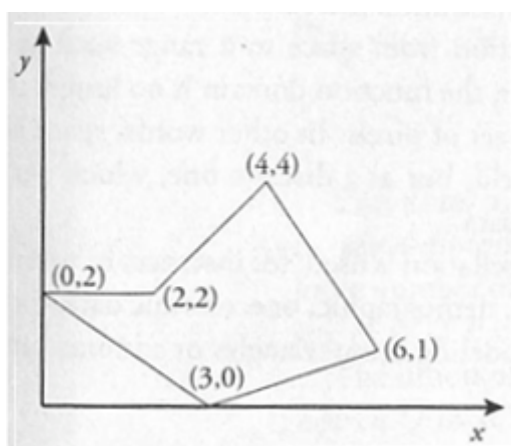


Figura 1.6: Rappresentazione di un poligono in formato vettoriale (tratto da [4])

Questi dati possono essere ricavati in vari modi: dalla digitalizzazione manuale, da vettorializzazioni di dati raster, da rilievi topografici con appositi strumenti, dai CAD o dai GPS. Il formato vettoriale, essendo definito attraverso equazioni matematiche, è indipendente dalla risoluzione e permette di esprimere i dati in un formato che occupi (molto) meno spazio rispetto

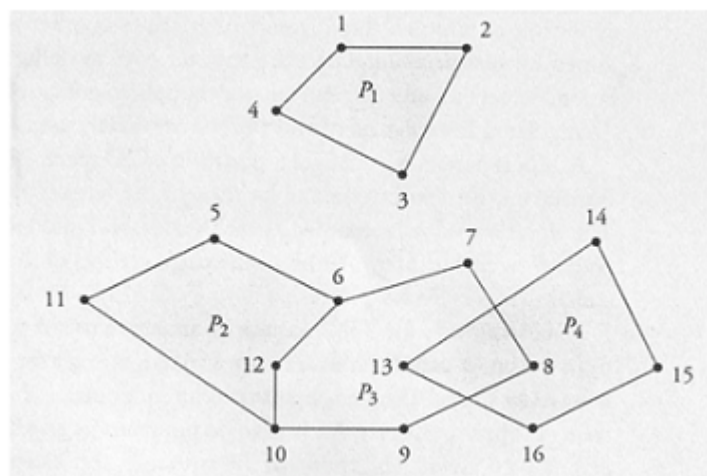


Figura 1.7: Ogni poligono è descritto da una lista di vertici e una regione consiste in un insieme di poligoni (tratto da [4])

all'equivalente raster, con una riduzione dell'occupazione di RAM e memoria di massa. Risulta, quindi, più facile da gestire e da modificare, essendo minore la quantità di dati coinvolti in ogni singola operazione di aggiornamento. Questo rende il vettoriale particolarmente adatto per gestire grandi quantità di dati come quelli cartografici che sono tipicamente gestiti con questa modalità; infine l'ingrandimento o la riduzione delle misure e proporzioni del soggetto prodotto in vettoriale non incidono in maniera significativa sul peso dell'immagine stessa, il riempimento di forme con tinte piatte è generato da semplici funzioni matematiche e risulta, quindi, estremamente leggero in termini di memoria utilizzata. Le entità geografiche che usano il modello vettoriale sono chiamate spesso feature. Le feature in un database geografico sono raggruppate in classi, oppure in una tabella, dove ogni riga si riferisce ad una singola feature, e le colonne rappresentano le proprietà delle feature. In genere i GIS gestiscono due tipi di feature: semplici e topologiche. La struttura dei dataset formati da linee e poligoni nel caso di feature semplici è a volte chiamata spaghetti, perché proprio come gli spaghetti tutti gli oggetti possono sovrapporsi e incrociarsi e non c'è alcun tipo di relazione tra

di loro. Il modello vettoriale è estremamente utile per descrivere fenomeni discreti, ma risulta meno adatto per descrivere fenomeni continui, quali temperatura, precipitazioni, quota, pendenza, cioè fenomeni che rappresentano un'unica grandezza che varia continuamente nello spazio. Il modello raster si è sviluppato proprio per descrivere tali fenomeni [4, 10, 11].

Modello raster

In una rappresentazione raster, lo spazio geografico è suddiviso in celle, di dimensione fissa o variabile, a cui sono associati un insieme di dati. Ogni variazione geografica è quindi espressa assegnando proprietà o attributi ad esse. Nella soluzione con celle di dimensione variabile (Fig 1.8) si usano unità di decomposizione di diverse dimensioni. Nella soluzione con celle di dimensione fissa (Fig 1.9) di solito lo spazio è decomposto in una griglia (raster) bidimensionale regolare di $N \times M$ celle rettangolari, chiamate *pixel*. Ogni pixel è identificato dalla coppia (x, y) , dove $x \leq N$ è la colonna e $y \leq M$ la riga della cella. In base alla dimensione delle celle, che è inversamente proporzionale alla precisione della carta, si ha una determinata risoluzione spaziale. Questa risoluzione è caratterizzata dal numero di pixel contenuti nell'unità di misura considerata (in genere il pollice inglese, equivalente a 2,54 cm) ed è ottenuta moltiplicando il numero delle righe di pixel per quello delle colonne di pixel; si misura in PPI (*Pixel Per Inch*) oppure in DPI (*Dot Per Inch*, punti per pollice).

I dati raster possono essere ricavati in vari modi: da programmi per interpretazione di immagini, da una macchina fotografica digitale, da uno scanner, da un satellite o attraverso l'elaborazione di dati, raster o vettoriali, già acquisiti. Sono abitualmente memorizzati come array di valori di una griglia, con dei metadati sull'array contenuti in un'intestazione. Tipicamente i metadati includono le coordinate geografiche dell'angolo in alto a sinistra della griglia, la dimensione delle celle e il numero di righe e colonne. L'array stesso è normalmente memorizzato in un file compresso o come record in un

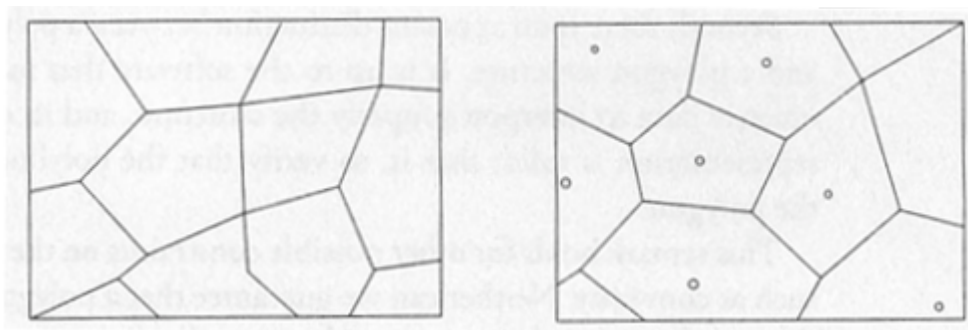


Figura 1.8: Celle di dimensione variabile (tratto da [4])

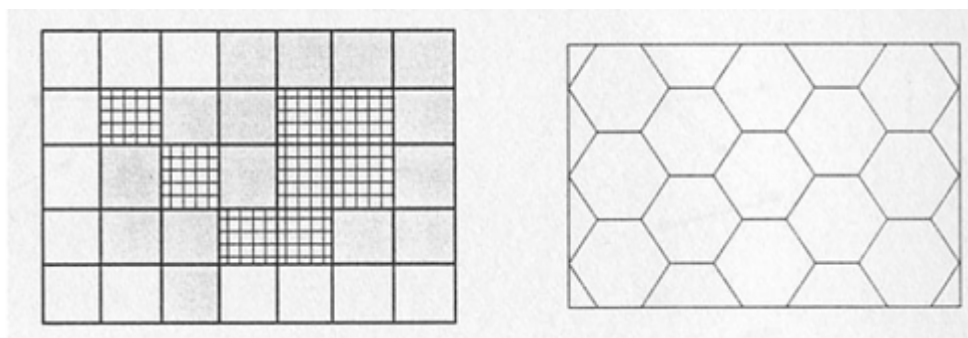


Figura 1.9: Celle di dimensione fissa (tratto da [4])

DBMS. Poiché occupano molta memoria, vengono adottate varie tecniche di compressione, ad esempio la codifica run-length o la codifica a blocchi. Chiaramente per rappresentare la superficie della terra col modello raster è necessario un processo di proiezione o appiattimento, visto che si tratta di una superficie curva ridotta ad una piatta. Il valore associato ad una data cella può essere assegnato in vari modi. Ad esempio usando la regola largest share (in cui la cella assume un valore in base al valore più diffuso al suo interno) oppure quella central point (in cui la cella assume il valore del suo punto centrale). Il primo metodo è preferito al secondo ma spesso il secondo viene usato per la facilità di calcolo [4, 10, 11].

1.2.4 Il Modello Logico dei Dati Spaziali

Un modello logico permette una rappresentazione astratta dei dati. La maggior parte dei sistemi di basi di dati oggi sul mercato si basa sul modello relazionale che è appunto un modello logico di rappresentazione o strutturazione dei dati di un database implementato su sistemi di gestione di basi di dati (DBMS), detti relazionali (RDBMS). Questo modello si basa sulla teoria degli insiemi e sulla logica del primo ordine² ed è strutturato intorno al concetto matematico di relazione. Per il suo trattamento ci si avvale di strumenti quali il calcolo relazionale e l'algebra relazionale [5].

Il modello relazionale prevede una serie di tipi di dato di base:

- INTEGER: numeri interi
- REAL: numeri reali approssimati
- STRING: stringhe di caratteri
- BOOLEAN: valori booleani (vero, falso)

Per trattare i dati spaziali occorre però introdurre uno o più tipi di dato detto SDT (*Spatial Data Type*). Si parla quindi di modello GEO-relazionale. Questo modello è organizzato in una serie di tabelle archiviate in un database relazionale e connesse attraverso comuni identificativi di oggetti. Per esempio in MySQL, noto RDBMS prodotto da Oracle, gli SDT, detti dati geometrici, sono stati introdotti con la versione 4.1 e si basano sulle specifiche dell'*Open GIS Consortium*. Qui vediamo quali sono i tipi di dati geometrici previsti da MySQL [MYSQLM]:

- GEOMETRY: può contenere un valore geometrico generico
- POINT: contiene un punto
- LINESTRING: contiene una linea

²linguaggio formale che serve per gestire meccanicamente enunciati e ragionamenti che coinvolgono i connettivi logici, le relazioni e i quantificatori

- POLYGON: contiene un poligono
- MULTIPOINT: rappresenta un insieme di punti
- MULTILINESTRING: rappresenta un insieme di linee
- MULTIPOLYGON: rappresenta un insieme di poligoni
- GEOMETRYCOLLECTION: rappresenta un insieme di dati geometrici di qualsiasi

1.2.5 Metodi di Rappresentazione dei Dati Spaziali

La rappresentazione dei valori geometrici in un GEO-DBMS presenta alcuni problemi: l'occupazione di memoria di un valore geometrico; l'imprecisione nelle coordinate (rappresentazione approssimata dei numeri reali); i valori geometrici sono fortemente correlati. Per questo vengono adottati essenzialmente due metodi per rappresentarli: un insieme di liste di coordinate (Fig 1.10) e una struttura basata sulla topologia dei dati (Fig 1.11). Secondo il primo metodo ogni tipo di dato spaziale è definito in modo analogo: poligoni come insieme di coordinate; linee come insieme di coordinate; punti come un singolo di coordinate. Ciò porta all'indipendenza tra i valori geometrici e a buone prestazioni dell'operazione di visualizzazione grafica della geometria di un sottoinsieme limitato di valori. Ma comporta anche degli svantaggi. Infatti tutte le selezioni spaziali e le operazioni di manipolazione richiedono l'applicazione di algoritmi di geometria computazionale e si crea una ridondanza nella rappresentazione della geometria.

Per il secondo metodo ogni tipo di dato spaziale è definito in modo diverso a seconda della sua struttura topologica: poligoni come insieme di linee; linee come insieme di punti; punti come coppia di coordinate. Questo comporta: eliminazione di ridondanza; riduzione di alcune operazioni geometriche a selezione di un insieme di identificatori o puntatori a valori geometrici elementari precalcolati; verifica dei vincoli spaziali immediata e supportata

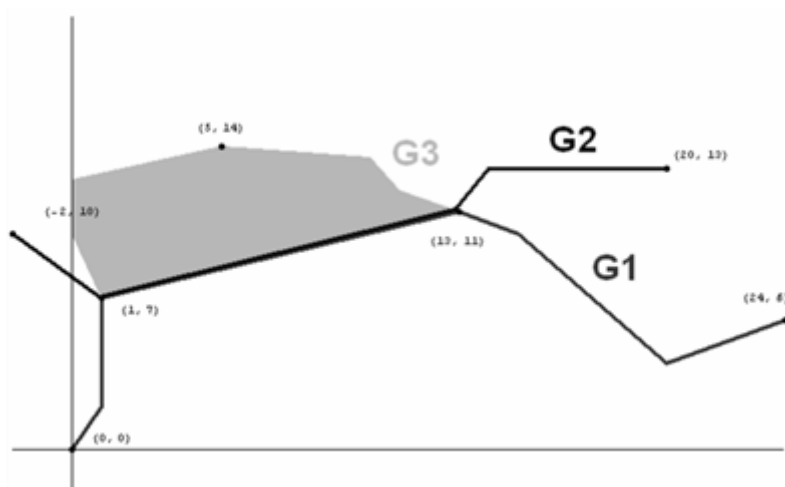


Figura 1.10: Rappresentazione dei dati spaziali come un insieme di liste di coordinate (tratto da [5])

dalla struttura topologica. Gli svantaggi che porta sono: il ripristino della topologia corretta a valle di ogni inserimento di un nuovo valore geometrico; la pura visualizzazione di un valore geometrico richiede il ritrovamento della geometria attraverso una catena di puntatori [5, 11].

1.2.6 Le Relazioni Spaziali

Per poter studiare e classificare le interrogazioni spaziali occorre analizzare le relazioni spaziali. Queste relazioni sono binarie poiché definiscono una proprietà relativa alla loro posizione nello spazio legando due valori geometrici di un determinato insieme. Sono presentate tre classi di relazioni spaziali binarie.

Relazioni topologiche

Esplicitano il legame topologico esistente fra due entità spaziali (esempio: la Senna attraversa Parigi). Queste relazioni fanno riferimento a concetti di

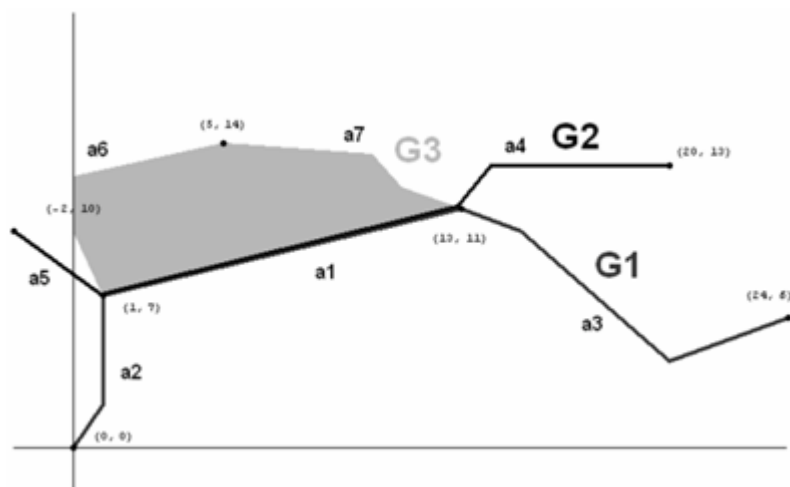


Figura 1.11: Rappresentazione dei dati spaziali come una struttura basata sulla topologia dei dati (tratto da [5])

alto livello sono di tipo qualitativo e non dipendono dalla distanza e dall'estensione. Dati due valori geometrici λ_1 e λ_2 si definiscono le seguenti relazioni topologiche:

- λ_1 **TOUCH** $\lambda_2 \Leftrightarrow (\lambda_1^\circ \cap \lambda_2^\circ = \phi) \wedge (\lambda_1 \cap \lambda_2 \neq \phi)$
- λ_1 **IN** $\lambda_2 \Leftrightarrow (\lambda_1^\circ \cap \lambda_2^\circ \neq \phi) \wedge (\lambda_1 \cap \lambda_2 = \lambda_1) \wedge (\lambda_1 \cap \lambda_2 \neq \lambda_2)$
- λ_1 **CROSS** $\lambda_2 \Leftrightarrow \dim(\lambda_1^\circ \cap \lambda_2^\circ) \leq (\max(\dim(\lambda_1), \dim(\lambda_2)) - 1) \wedge (\lambda_1 \cap \lambda_2 \neq \lambda_1) \wedge (\lambda_1 \cap \lambda_2 \neq \lambda_2)$
- λ_1 **OVERLAP** $\lambda_2 \Leftrightarrow \dim(\lambda_1^\circ \cap \lambda_2^\circ \neq \phi) = \dim(\lambda_1) = \dim(\lambda_2) \wedge (\lambda_1 \cap \lambda_2 \neq \lambda_1) \wedge (\lambda_1 \cap \lambda_2 \neq \lambda_2)$
- λ_1 **DISJOINT** $\lambda_2 \Leftrightarrow (\lambda_1 \cap \lambda_2 = \phi)$

dim ci restituisce la dimensione del risultato (0 punto, 1 linea, 2 poligono, o empty se l'intersezione è vuota).

1.2.7 Relazioni basate sulla direzione

Specificano la posizione in cui un oggetto si trova rispetto ad un altro (Fig 1.12). Queste relazioni sono definite sulla base di un sistema di riferimento direzionale che partiziona lo spazio in almeno quattro quadranti : l'origine del sistema viene posto su uno dei due valori geometrici considerati e intersecando i quadranti con l'altro valore geometrico si deriva la relazione tra i due valori.

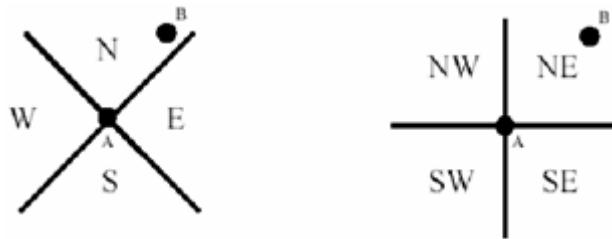


Figura 1.12: Relazione basata sulla direzione: Brescia si trova a nord-est di Alessandria (tratto da [6])

1.2.8 Relazioni basate sulla distanza

Specificano la distanza fra due oggetti (per esempio la palestra è a 2 km dalla scuola). È necessaria la definizione nello spazio di riferimento di una funzione che calcola la distanza come quella euclidea fra due punti $P = (a, b)$ e $Q = (c, d)$:

$$D(P, Q) = \sqrt{(a - c)^2 + (b - d)^2} \quad (1.7)$$

Se si considerano due valori con estensione significativa occorre ridefinire il concetto di distanza, ad esempio:

$$Dist(X, Y) = \min(D(P, Q) : P \in X, Q \in Y) \quad (1.8)$$

Quindi fissata una coppia di distanze (d_1, d_2) è possibile definire una relazione metrica tra due valori geometrici come segue:

$$R_{(d_1, d_2)}(X, Y) \Leftrightarrow (d_1 < Dist(X, Y) < d_2) \quad (1.9)$$

1.2.9 Interrogazione di una Base di Dati Spaziale

Come per il modello logico, occorre introdurre dei costrutti per le interrogazioni spaziali. A tal fine viene utilizzata un'estensione dell'algebra relazionale, detta geoalgebra. Queste interrogazioni si possono dividere essenzialmente in 2 categorie: selezioni basate su relazioni spaziali rispetto a valori geometrici costanti e join spaziali.

Selezioni basate su relazioni spaziali rispetto a valori geometrici costanti

La prima categoria si riferisce alle interrogazioni basate sulle relazioni di ogni valore geometrico con lo spazio di riferimento o con valori geometrici costanti. Di questa categoria fanno parte le range query, le selezioni basate sulla direzione e le selezioni basate sulla distanza.

Range query. Si selezionano tutti i valori geometrici della base di dati che sono in una particolare relazione con una regione dello spazio di riferimento. Le relazioni spaziali in questione sono quelle topologiche fra cui: **DISJOINT**, **OVERLAP**, **CROSS**, **IN/CONTAIN**, **TOUCH**, **EQUAL**. Nella Figura 1.13 è illustrato un esempio.

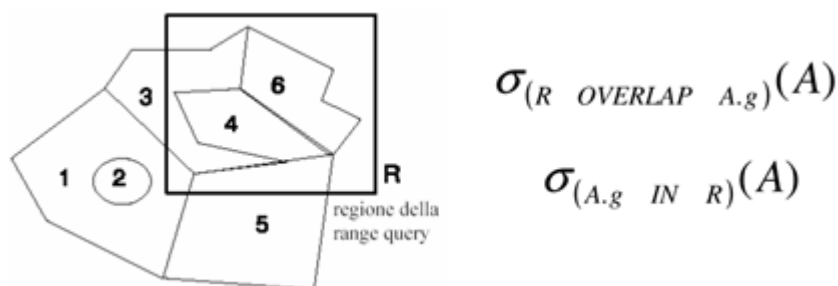


Figura 1.13: Esempio Range query. Dati una tabella A con attributo geometrico g e una regione R (tratto da [6])

Selezioni basate sulla direzione. Si selezionano tutti i valori geometrici della base di dati che sono in una particolare relazione topologica con una delle suddivisioni dello spazio, derivanti dal riferimento direzionale posizionato su un particolare valore geografico. Nella Figura 1.14 è illustrato un esempio.

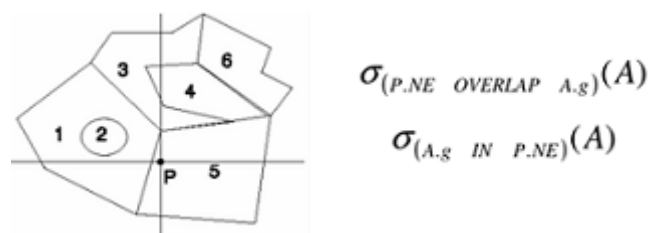


Figura 1.14: Esempio di selezione basata sulla direzione. Dati una tabella **A** con attributo geometrico **g** e un valore geometrico **P** (punto) (tratto da [6])

Selezioni basate sulla distanza. Si selezionano tutti i valori geometrici della base di dati che sono ad una distanza compresa tra un valore minimo d_1 e un valore massimo d_2 da un particolare valore geografico. Nella Figura 1.15 è illustrato un esempio.

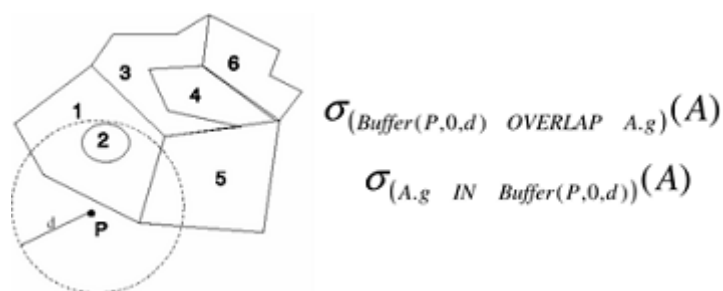


Figura 1.15: Esempio di selezione basata sulla distanza. Dati una tabella **A** con attributo geometrico **g**, un valore geometrico **P** e due distanze **0** e **d** (tratto da [6])

1.2.10 Join spaziali

La seconda categoria si riferisce alle interrogazioni basate sulle relazioni dei valori geometrici contenuti nella base di dati. Queste operazioni vengono utilizzate per confrontare i valori geometrici di tuple diverse di una stessa tabella duplicandola. Di questa categoria fanno parte:

join topologici (Fig 1.16) *join basati sulla direzione* (Fig 1.17) e *join basati sulla distanza* (Fig 1.18) [6, 11].

Esempio Join topologico. Date le tabelle FIUME e COMUNE identificare tutti i comuni il cui territorio è attraversata da un fiume.

FIUME (nome, tracciato:LN)

COMUNE (nome, estensione:PG)



Figura 1.16: Join topologico (tratto da [6])

Esempio Join basato sulla direzione. Date le tabelle CIMA e RIFUGIO identificare per ogni cima tutti i rifugi che stanno a nord-est.

CIMA (nome, posizioneC:PT)

RIFUGIO (nome, posizioneR:PT)



Figura 1.17: Join basato sulla direzione (tratto da [6])

Esempio Join basato sulla distanza. Date le tabelle OSPEDALE e SCUOLA identificare per ogni scuola gli ospedali in un raggio di 2 Km.

OSPEDALE (nome, posizioneO:PT)

SCUOLA (nome, posizioneS:PT)

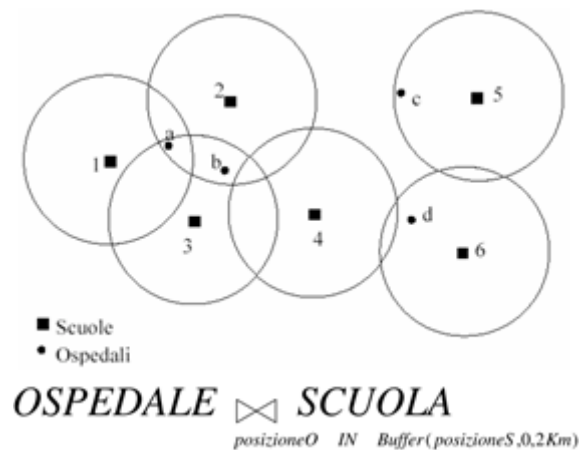


Figura 1.18: Join basato sulla distanza (tratto da [6])

Di seguito sono elencati alcuni esempi di interrogazioni spaziali [7].

- Selezionare le nazioni confinanti con la (adiacenti alla) Francia.

```

SELECT N1.Nome
FROM NAZIONE N1, NAZIONE N2
WHERE N2.Nome = "Francia"
and Touch(N1.Geometria, N2.Geometria)

```

- Determinare i fiumi che attraversano la Germania.

```

SELECT F.Nome
FROM FIUME F, NAZIONE N
WHERE N.Nome = "Germania"
and Cross(F.Geometria, N.Geometria)

```


- Determinare la lunghezza totale dei fiumi che attraversano l'Italia.

```
SELECT Sum(Length(F.Geometria))
FROM FIUME F, NAZIONE N
WHERE N.Nome = "Italia"
and Cross(F.Geometria, N.Geometria)
```

- Determinare le nazioni contenute nel rettangolo *@rectangle*.

```
SELECT Nome
FROM NAZIONE
WHERE In(Geometria, @rectangle)
```

1.2.11 Indici spaziali

Un accesso efficiente ai dati è essenziale anche per le basi di dati spaziali. Le strutture d'accesso sviluppate per indicizzare i dati numerici o alfanumerici tradizionali si basano sull'ordinamento totale del dominio del campo chiave ma non si adattano ai dati spaziali. Quindi occorrono nuove strutture di indicizzazione per questa tipologia di dato: *space-driven* e *data driven* [8].

1. Strutture basate sullo spazio (space-driven)

Lo spazio di riferimento è suddiviso in celle rettangolari, indipendentemente dalla distribuzione degli oggetti. Gli oggetti sono riferiti alle celle in base a qualche criterio geometrico. Di seguito sono elencati alcuni esempi.

- **fixed grid** : struttura proposta per indicizzare insiemi di punti;
- **grid file** : struttura proposta per indicizzare oggetti rispetto al valore di più attributi (indice multichiave);
- **strutture lineari**: usate nelle estensioni spaziali dei DBMS relazionali e si integrano facilmente con i B^+ -Tree (Oracle Spatial utilizza i Quad-Tree).

Quad-Tree

Un Quad-Tree (Fig 1.19) è una struttura dati ad albero non bilanciata nella quale tutti i nodi interni hanno esattamente quattro nodi figli. Questi indici sono spesso usati per partizionare uno spazio bidimensionale suddividendolo ricorsivamente in quattro quadranti, comunemente denotati come NE (Nord-Est), NW (Nord-Ovest), SE (Sud-Est), SW (Sud-Ovest). Quindi fissato un numero massimo di elementi contenuti in un quadrante, ogni volta che si verifica un *overflow*, il quadrante interessato viene suddiviso. La dimensione dei sottoquadranti, ottenuti dalla suddivisione di un quadrante, non deve essere necessariamente la stessa, anche se di solito è ciò che accade.

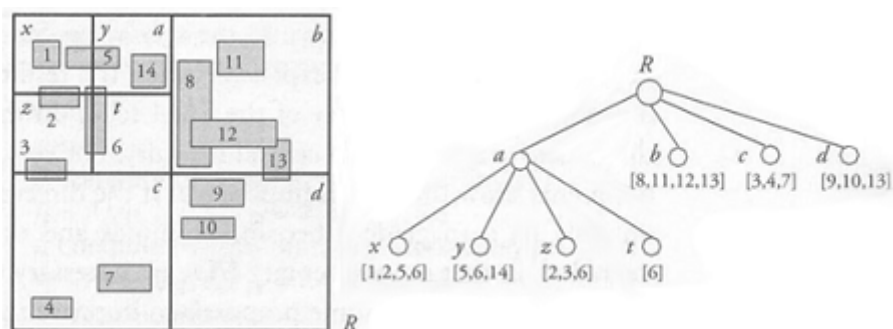


Figura 1.19: Esempio di Quad-Tree semplice (tratto da [8])

I Quad-Tree presentano diversi limiti: i 4 figli di un nodo occupano una piccola parte di pagina quindi è difficile mapparli nelle pagine di memoria secondaria; il tempo di accesso è legato alla profondità dell'albero quindi, nel caso peggiore, il numero di operazioni di I/O è pari a questa profondità poiché ogni nodo interno si trova in una pagina diversa; infine vi è il problema della duplicazione degli oggetti che causa un aumento della dimensione dell'indice [8, QTREE].

2. Strutture basate sui dati (data-driven)

Le strutture sono organizzate partizionando l'insieme degli oggetti, tenendo conto della loro distribuzione nello spazio di riferimento. Ad ogni nodo è associato un rettangolo, detto DR (*Directory Rectangle*, (Fig 1.22), che rappresenta l'MBB³ dei rettangoli dei nodi figli o degli oggetti ad esso associati. Per accedere a un rettangolo della collezione indicizzata, si percorre l'albero dalla radice verso una o più foglie, verificando per ogni DR incontrato, le relazioni di sovrapposizione e inclusione. La struttura originaria di questa tipologia di indice è quella dell' R-tree poi ci sono diverse varianti come R^* -tree e R^+ -tree [9].

R-tree

Gli R-tree (Fig 1.20, 1.21) sono un'estensione multidimensionale dei B^+ -Tree (struttura ad albero bilanciata e paginata), e sono usati per indicizzare ad esempio le coordinate spaziali (x, y) per i dati geografici. La struttura dati divide lo spazio in MBR (*Minimum Bounding Rectangles*, infatti R-tree deriva proprio da *Rectangle*) innestati gerarchicamente e quando possibile sovrapposti. Ogni nodo dell' R-tree ha un numero variabile di entry (fino ad un massimo predeterminato). Ogni entry che non sia un nodo foglia contiene due entità: una identifica il nodo figlio, l'altra l'MBR che contiene tutte le entry del nodo figlio. L'algoritmo di inserimento e cancellazione di entry dagli MBR assicura che elementi vicini siano posizionati nello stesso nodo foglia: un nuovo elemento andrà nel nodo foglia che richiede il minor numero di estensioni delle dimensioni dell'MBR. Gli algoritmi di ricerca usano gli MBR per decidere se cercare o meno nel nodo figlio del nodo corrente. In questo modo la maggior parte dei nodi non viene esplorata dagli algoritmi. Per questo motivo, come per i B-tree, ciò rende gli R-tree

³*Minimal Bounding Box*: è un descrittore che approssima il dato spaziale con un rettangolo (MBR) o con un poligono convesso

adatti ai database, dove i nodi possono essere copiati in memoria solo quando necessario. Diversi algoritmi possono essere usati per dividere i nodi quando diventano troppo estesi, ovvero quando vengono aggiunti in un nodo un numero di elementi che supera il limite prestabilito [9, RTREE].

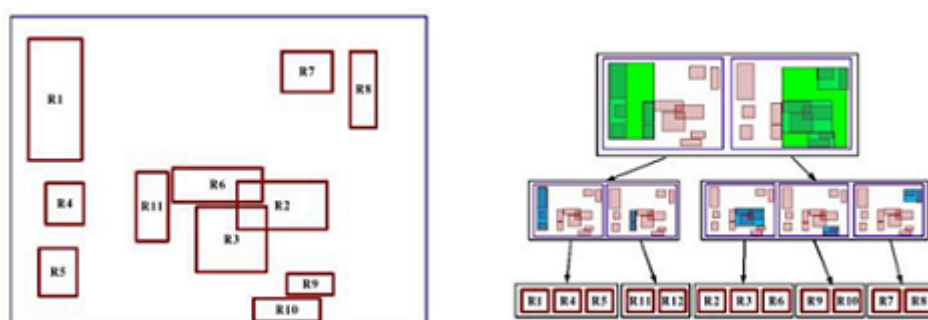


Figura 1.20: Indice R-tree di una base di dati bidimensionale (tratto da [12])

Nella Figura 1.20, il rettangolo esterno a sinistra rappresenta il dominio spaziale di una base di dati bidimensionale. Al suo interno sono presenti 12 oggetti. A destra è rappresentato l'indice R-tree relativo agli oggetti della figura di sinistra. Poiché ogni record rappresenta una porzione di spazio, essi sono stati rappresentati graficamente. Ad esempio, la radice contiene due record, e per ognuno è stata rappresentata l'intera base di dati con evidenziata in l'area da essi indicizzata. La stessa cosa è ripetuta nei due nodi interni. Si può notare come l'area indicizzata da ogni record comprenda l'unione delle aree dei figli. Per compattezza, i record nelle foglie sono stati rappresentati con il solo nome.

Un R-tree è efficiente se sia la *coverage* (l'area totale di tutti i rettangoli associati ai nodi del livello) che l'*overlap* (l'area totale coperta da due o più nodi) sono minimizzati. In particolare l'overlap comporta problemi in fase di ricerca. Infatti per cercare R12 nella Figura 1.20, dobbiamo attraversare i nodi interni R1, R2, R3, R4, R5 ed R6 a causa del fatto che R1 ed R2 hanno una sovrapposizione. Inoltre è cruciale l'ordine di

inserimento per ottenere un albero maggiormente bilanciato [12].

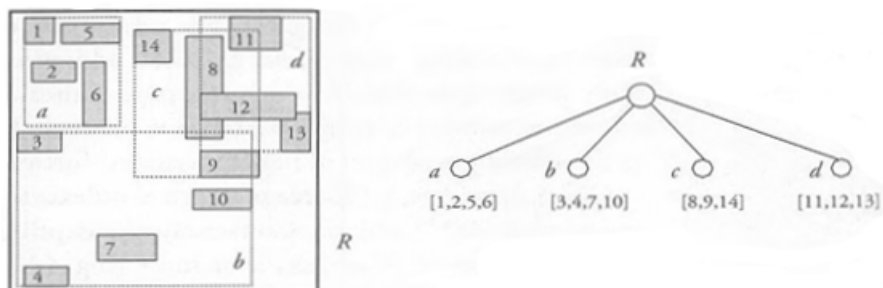


Figura 1.21: Indice R-tree dove i rettangoli delle foglie sono rappresentati da linee punteggiate (tratto da [9])

Gli R-tree non garantiscono una performance ottima di caso peggiore: i rettangoli si possono sovrapporre tra di loro, dando vita a percorsi differenti che portano allo stesso rettangolo. Questo può provocare un accesso multiplo al disco ma in generale si comportano molto bene con dati reali. Infatti, poiché riducono notevolmente l'altezza dell'albero, sono molto utilizzati per l'ottimizzazione degli accessi al disco.

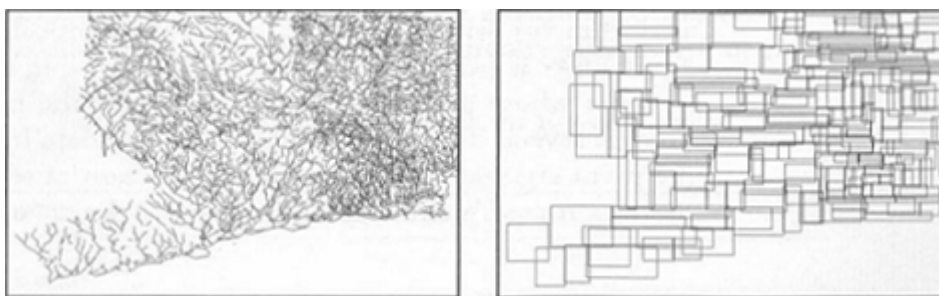


Figura 1.22: DR di un R-tree costruito per raccogliere i dati idrografici dello Stato del Connecticut negli USA (tratto da [9])

1.3 GIS

Un *Geographic Information System*, o più semplicemente GIS (Fig 1.23), è un sistema informativo computerizzato che permette l'acquisizione, la registrazione, l'analisi, la visualizzazione e la restituzione di informazioni derivanti da dati geografici (georiferiti) [GIS]. Trattasi quindi di un sistema informatico in grado di produrre, gestire e analizzare dati spaziali associando a ciascun elemento geografico una o più descrizioni alfanumeriche. In pratica il GIS è l'ambiente Software che permette l'elaborazione e la manipolazione dei dati geometrici georeferenziati, che sono memorizzati in apposite strutture dati dei DBMS (*Database Management System*) che gestiscono anche la spazialità, o in singoli file. Questa tecnologia consente di interpretare fenomeni del mondo reale attraverso l'elemento unificante del territorio e permette di effettuare interrogazioni, analisi spaziali e statistiche in aggiunta a quelle tradizionali. Ciò che contraddistingue un GIS è la possibilità di rappresentare relazioni che nessun collegamento di tipo logico (ad esempio in un database) potrebbe rappresentare, cioè quelle legate alla posizione.

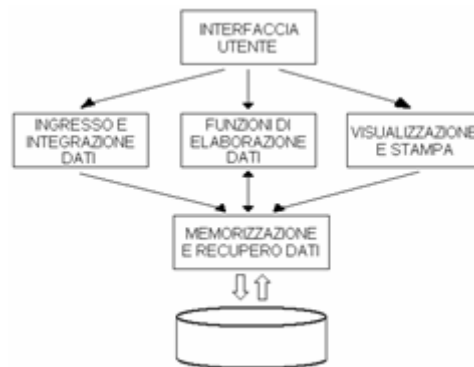


Figura 1.23: Architettura di un GIS (tratto da [3])

Seguono alcune definizioni.

“Il GIS è composto da una serie di strumenti software per acquisire, memorizzare, estrarre, trasformare e visualizzare dati spaziali dal mondo reale [13]”.

“I sistemi informativi geografici sono sistemi informatizzati per l’acquisizione, la memorizzazione, il controllo, l’integrazione, l’elaborazione e la rappresentazione di dati che sono spazialmente riferiti alla superficie terrestre [14]”.

“Un insieme organizzato di hardware, software e persone finalizzato all’acquisizione, gestione, analisi e visualizzazione di informazioni alfanumeriche e geografiche riferite ad un determinato territorio [15]”.

“Un insieme organizzato di strumenti hardware e software, dati geografici e procedure orientati ad una efficiente raccolta, memorizzazione, modifica, manipolazione, analisi e restituzione di dati riferiti spazialmente, per la risoluzione di problemi complessi di gestione e pianificazione del territorio [16]”.

I GIS vengono utilizzati in diversi campi applicativi: nell’analisi territoriale di sistema per rappresentare le varie componenti del territorio quali la morfologia, la geologia, l’indice demografico e l’idrologia; nella progettazione del territorio come piani regolatori, piani di sviluppo e zonizzazioni; nell’analisi dei rischi ambientali o sismici; nell’analisi dei percorsi e delle infrastrutture quali reti di servizi e traffico; nell’analisi dei flussi e dei bacini, come la rappresentazione delle aree esposte a esondazioni o la diffusione di contagi epidemici; nel monitoraggio e controllo del territorio per confrontare, attraverso foto aeree, il modificarsi del territorio o la localizzazione di mezzi mobili; nella gestione tattica e strategica delle forze armate; infine nel marketing per individuare bacini potenziali d’utenza [19, 20].

1.3.1 Storia dei GIS

Il GIS nasce in Canada, prodotto dalla *Canada Land Inventory*, alla metà degli anni 60 come un sistema computerizzato di misurazione di mappe (*CGIS: Canada GIS*), che integra informazione di tipo tabulare con quella grafica per la rappresentazione di mappe. Alla fine degli anni 60 l’*US Bureau of Census* pianifica la realizzazione di uno strumento utile alla realizzazione

del censo della popolazione del 1970. Viene data vita al progetto DIME (*Dual Independent Map Encoding*) che associa le informazioni dettagliate su tutte le strade degli USA ai record del censo. I due progetti (CGIS e DIME) danno quindi vita ad un programma unitario presso il laboratorio di *Computer Graphics and Spatial Analysis* dell'università di Harvard, teso allo sviluppo di GIS *general-purpose* che potesse essere utile ad entrambe le esigenze. Tale programma portò alla realizzazione di *ODISSEY GIS* alla fine degli anni 70. D'altra parte, già dagli anni 60 i cartografi e le agenzie che si occupavano di mappe avevano cominciato a chiedersi come i computer potessero aiutare il loro lavoro e alla fine degli anni 70 la maggior parte delle principali agenzie di cartografia erano in larga parte computerizzate. I GIS raggiunsero discrete capacità di svolgere analisi ed elaborazioni di dati ma presentavano ancora carenze nella rappresentazione grafica. Negli anni 80 ci fu una lenta diffusione nelle grandi P.A. (Pubblica Amministrazione): si svilupparono miglioramenti tecnologici hardware che coinvolgevano sia la resa grafica a livello video che quella su supporto cartaceo per le stampanti grazie anche all'unione della metodologia di gestione informatica dei database con la grafica computerizzata. Agli inizi degli anni 90 grazie anche ad un continuo miglioramento a livello hardware e software, ci fu una diffusione anche nelle piccole P.A., nella sanità e a livello commerciale con l'avvento dei *personal computer*. Attualmente i GIS sono largamente diffusi nella P.A. e anche in ambiente privato. Si ricorda che un importante ruolo nello sviluppo dei GIS l'hanno avuto gli strumenti di misura remoti, come i satelliti che si sono cominciati a diffondere già dagli anni 50; in particolare durante la guerra fredda sono state sviluppate molte applicazioni e innovazioni sui GIS, per opera dei militari (ad esempio il GPS) [20, 2].

1.3.2 Componenti di un GIS

Uno strumento GIS operativo integra cinque componenti chiave: hardware, software, dati, utenti e metodi.

1. L'*hardware* è rappresentato dal computer su cui opera il GIS. Oggi un software GIS è in grado di essere supportato su un'ampia gamma di piattaforme hardware, dal server al desktop usato sia in configurazione stand-alone sia in rete.
2. Il *software* fornisce le funzioni e gli strumenti per memorizzare, analizzare e visualizzare informazioni geografiche. Le componenti chiave del software sono:
 - Strumenti per l'inserimento e la manipolazione delle informazioni geografiche
 - Sistema per la gestione del database (DBMS) con dimensioni da pochi MB a molti TB
 - Strumenti per la ricerca, l'analisi e la visualizzazione
 - Interfaccia grafica di facile accesso agli strumenti
3. La componente più importante di un GIS è costituita dai *dati*. I dati geografici e le informazioni alfanumeriche a loro associate possono essere acquisiti direttamente dall'utente o acquistati da un fornitore. Un GIS è in grado di integrare i dati spaziali con altri tipi di sorgenti di dati e può usare un DBMS, comune alla maggior parte delle strutture o società, per organizzare e mantenere i propri dati, perfino per la gestione dei dati spaziali.
4. La tecnologia GIS avrebbe un valore limitato senza le persone che gestiscono il sistema e sviluppano progetti per la risoluzione di problemi del mondo reale. Gli *utenti* GIS spaziano dagli specialisti tecnici che disegnano e gestiscono il sistema a coloro che lo utilizzano per adempiere al lavoro quotidiano.
5. Uno strumento GIS di successo rende compatibile un progetto ben disegnato con le regole dell'impresa, cioè con i *metodi* e le procedure operative specifiche di ciascuna organizzazione. Quindi sono tutti quei

meccanismi che assicurano che le attività di un GIS stiano nel budget programmato, abbiano una buona qualità e intercettino i bisogni dell'organizzazione [20, 2].

1.3.3 Come lavora un GIS

Il GIS memorizza le informazioni geografiche come una collezione di *layers* (strati) tematici che possono essere tra loro relazionati, tramite collegamento e sovrapposizione geografica, e rappresentati sullo schermo geometricamente da punti, linee o aree. Per esempio uno strato di punti può rappresentare pozzi; uno strato di linee può rappresentare corsi d'acqua o strade; uno strato di aree può rappresentare zone di uguale uso del suolo, o costruzioni. Ad ogni elemento geografico corrisponde un attributo o elemento descrittivo che indica cosa rappresenta l'elemento spaziale, e la sua esatta posizione geografica espressa in coordinate. Tale concetto, semplice ma estremamente potente e versatile, si è rivelato di incalcolabile valore per la risoluzione di molti problemi del mondo reale quali ottimizzazione di percorsi, applicazioni di pianificazione urbanistica, modelli di circolazione atmosferica, localizzazione di veicoli di consegna [20, 2].

1.3.4 Referenze geografiche

L'informazione geografica contiene un riferimento spaziale esplicito (quale latitudine e longitudine o un sistema di coordinate nazionali) o implicito come un indirizzo, un codice postale, una sezione di censimento, una denominazione stradale. Un processo automatico chiamato *geocodifica* viene utilizzato per creare un esplicito riferimento geografico (localizzazione multipla) da un implicito riferimento (descrizione per indirizzi). Questi riferimenti geografici aiutano a localizzare gli elementi nei diversi applicativi di analisi [20, 2].

1.3.5 Tipi di dato utilizzati

I GIS di ultima generazione sono in grado di gestire a livello implementativo due formati di dato spaziale: il formato *vettoriale* e il formato *raster*. Infatti si tende ad integrare le due tipologie nel miglior modo possibile, permettendo, in strutture omogenee di archivi, la convivenza di dati, l'elaborazione interconnessa di questi ultimi e la conversione dei formati. Oltre ai dati spaziali, in un GIS sono utilizzati anche gli *attributi*: dati non grafici, detti descrittivi, che hanno la funzione di esprimere le caratteristiche degli elementi grafici. Sono i dati alfanumerici (ma anche le descrizioni, le foto, i disegni, ecc.) associati all'elemento geografico. Per i dati geografici di tipo raster gli attributi sono associati alla singola cella. Per i dati geografici di tipo vettoriale gli attributi sono associati agli elementi punto, linea, poligono. Vengono generalmente gestiti attraverso un sistema di database relazionale. Per esempio in un formato tipico per i database, il *DBF*, gli attributi alfanumerici sono riuniti in tabelle tipizzate per colonna: ad ogni riga (record) della tabella è associato un diverso elemento geografico; ogni colonna, o campo (*field*), della tabella rappresenta una diversa classe di attributi [1, 2].

1.3.6 Compiti del GIS

L'obiettivo generale dei sistemi informativi geografici e' essenzialmente quello di svolgere cinque compiti: inserimento, manipolazione, gestione, ricerca o analisi e visualizzazione.

Inserimento. Prima che possano essere utilizzabili in un GIS, i dati geografici devono essere convertiti in un idoneo formato digitale. Il processo di conversione dei dati dalle mappe cartacee ai file è denominato digitalizzazione. La moderna tecnologia GIS è in grado di automatizzare questo processo per grandi progetti utilizzando il sistema della scansione; per attività di minore entità è possibile ricorrere alla digitalizzazione manuale. Oggi molti tipi di dati geografici sono già disponibili in for-

mato digitale e possono essere acquisiti da fornitori specifici e caricati direttamente nel GIS.

Manipolazione. E' verosimile che i dati richiesti da uno specifico progetto GIS necessitino di essere trasformati o manipolati per renderli compatibili con il sistema. Per esempio, l'informazione geografica è disponibile a scale differenti (mezzerie stradali dettagliate, limiti censuari con un livello inferiore di dettaglio, delimitazione delle aree postali). Prima che tali informazioni possano essere integrate, devono essere ricondotte ad una medesima scala (livello di dettaglio o accuratezza). Tale trasformazione può essere solo temporanea per scopi di visualizzazione o permanente per eseguire analisi. La tecnologia GIS offre molti strumenti per manipolare i dati spaziali e per eliminare le informazioni non necessarie.

Gestione. Per piccoli progetti GIS può essere sufficiente memorizzare le informazioni geografiche come semplici file. Naturalmente, quando volume dei dati e numero degli utenti aumentano, è preferibile utilizzare un DBMS per memorizzare, organizzare e gestire i dati. Un DBMS non è altro che un software per la gestione del database. Esistono molti differenti modelli di DBMS, ma nel GIS il più utilizzato è il modello relazionale, che memorizza i dati come una collezione di tabelle. Campi comuni in differenti tabelle ne consentono il collegamento. Questo modello sorprendentemente semplice è stato usato ampiamente per la sua flessibilità e ha trovato un vasto impiego in molte applicazioni GIS e non.

Ricerca ed analisi. Avendo a disposizione un GIS funzionante e contenente informazioni geografiche, è possibile cominciare a rivolgere al sistema semplici domande quali:

- Qual è la distanza tra due posti?
- Quali sono le zone ad uso industriale?

Oppure si possono eseguire analisi più complesse quali:

- Quali sono i siti più indicati per edificare nuove case?
- Costruendo una nuova autostrada, quali saranno le conseguenze sul traffico?
- Quale è il tipo di suolo prevalente in provincia?
- Quali sono i boschi siti in luoghi con pendenza rilevante?

Il GIS consente sia di effettuare delle semplici query interattive (*point-and-click*), sia di compiere sofisticate analisi legate ai contenuti propri di specifiche discipline territoriali per fornire informazioni tempestive. La tecnologia GIS dimostra tutte le sue potenzialità quando è utilizzata nell'analisi dei dati geografici allo scopo di creare modelli, individuare tendenze e predisporre scenari. I moderni GIS possiedono molti strumenti analitici potenti, ma due sono particolarmente importanti.

1. *Proximity Analysis* (analisi superficiale)

- Quante sono le case situate nei pressi di 1000 metri dalla condotta principale dell'acqua?
- Quanti sono gli abitanti serviti nel raggio di 600 metri dalla fermata dell'autobus?
- In quale proporzione la coltivazione di erba medica è entro 500 metri dal pozzo?

Per rispondere a queste domande la tecnologia GIS utilizza un processo di calcolo chiamato *buffering*, che consente di stabilire relazioni di prossimità tra gli elementi geografici.

2. *Overlay analysis* (analisi delle sovrapposizioni)

L'integrazione di livelli di dati differenti implica l'uso di un processo chiamato *overlay*. Questa può sembrare una semplice operazione di visualizzazione ma in realtà, dal punto di vista analitico,

si tratta di collegare tra loro informazioni utilizzando sia il calcolo geometrico che quantitativo. Tale *overlay*, o *join spaziale*, può integrare per esempio i dati su suoli, pendenze e vegetazione, oppure può incrociare la proprietà del terreno con gli accertamenti fiscali. Include la costruzione di opportuni modelli deterministici e stocastici interpretativi.

Visualizzazione. Per molti tipi di operazioni geografiche il risultato finale è rappresentato nel modo ottimale come mappa o grafico. Le mappe sono molto efficaci per memorizzare e comunicare informazioni geografiche: mentre i cartografi hanno creato mappe per millenni, il GIS fornisce nuovi strumenti per estendere l'arte e la scienza della cartografia. La rappresentazione su mappe può essere integrata con report, viste tridimensionali, immagini fotografiche ed altre rappresentazioni multimediali [2].

1.3.7 Tecnologie connesse

I GIS sono strettamente connessi a molti altri tipi di sistemi informativi, ma è la capacità di manipolare ed analizzare dati geografici che distingue la loro tecnologia. Benchè non esistano regole ferree su come classificare i sistemi informativi, la descrizione che segue potrà aiutare a differenziare i GIS da altre tecnologie, quali *Desktop GIS*, CAD (*Computer-Aided Design*), telerilevamento, GNSS e DBMS, pur essendo spesso strettamente correlato ed integrato con molte di loro.

Desktop GIS. Un sistema *desktop mapping* usa le mappe per organizzare i dati ed interagire con l'utente. Il punto focale di tali sistemi è la creazione della mappa cioè il database. La maggior parte dei sistemi *desktop mapping* hanno funzioni molto limitate di gestione dati, analisi spaziale e personalizzazione. I sistemi desktop mapping girano su computer desktop quali PC, Macintosh e piccole workstation UNIX. Segue una descrizione di 2 Desktop GIS.

GRASS

GRASS (*Geographic Resources Analysis Support System*) è un Desktop GIS utilizzato per la gestione, elaborazione, modellamento spaziale e visualizzazione di molti tipi di dati, analisi dell'immagine e produzione di grafici. Comprende oltre 350 moduli per l'elaborazione di dati vettoriali (2D/3D), raster e voxel. Ha diverse interfacce per l'integrazione con altri programmi di geostatistica, basi di dati, applicazioni geografiche su internet e altri pacchetti GIS. E' il più grande progetto GIS nell'ambito degli Open Source e può essere utilizzato sia come Desktop GIS che come elemento principale di una più completa infrastruttura GIS. Originariamente scritto, sviluppato, gestito e diffuso dall'USA-CERL (*U.S. Army Construction Engineering Research Laboratories*), sezione dell'*US Army Corp of Engineers*, come strumento per la gestione del territorio e la pianificazione ambientale per scopi militari, GRASS si è evoluto divenendo rapidamente un potente strumento in un ampio campo di applicazioni in molte differenti aree della ricerca scientifica. GRASS è attualmente utilizzato in tutto il mondo in ambienti accademici e commerciali, in molti settori governativi, che includono NASA, NOAA, USDA, il *National Park Service*, l'*U.S. Census Bureau*, USGS, e in molte compagnie di consulenza ambientale. GRASS è sviluppato da una rete mondiale di operatori che continuano a rilasciare con regolarità nuove versioni di GRASS. Le piattaforme supportate includono Linux, MacOS X, MS Windows (NT, 2000 Professional, XP, Vista, 7), Sun Solaris, Silicon Graphics Irix, HP-UX, DEC-Alpha etc. [17, GRASSORG, GRASSIT]

gvSIG

gvSIG è frutto dell'iniziativa intrapresa in Spagna dal Dipartimento delle Infrastrutture e dei Trasporti della Comunità Autonoma di Valencia (*Generalitat Valenciana*) a partire dal 2003. Il progetto è finanziato

mediante il Fondo Europeo di Sviluppo Regionale (FERS/FEDER). Questo software presenta un'interfaccia user friendly ed un rapido accesso ai formati raster e vettoriali più comuni. È in grado di integrare dati locali e remoti all'interno della stessa vista utilizzando servizi WMS, WFS, WCS, ArcIMS e sorgenti di dati JDBC, PostGIS, Oracle Spatial, ecc. e, inoltre, include una serie di funzioni fra cui l'editing avanzato, il geoprocessing, la gestione dei CRS, il supporto a diversi linguaggi di scripting, il georeferencing, il network analyst, ecc. È anche un client SDI (Spatial Data Infrastructure), poiché prevede tra le varie funzionalità anche strumenti di ricerca di geodati quali Gazetteer e Catalog. Attualmente gvSIG si è arricchito dell'estensione Sextante, che raccoglie più di 150 strumenti di analisi dati geografici (raster e vettoriali).[GVPONTIS, SEXTANTE]

CAD. I sistemi *Computer Aided Design*, noti come CAD, si sono sviluppati per creare disegni, progetti architettonici ed infrastrutture [CAD]. Tale attività richiedeva che componenti con caratteristiche specifiche fossero assemblate per creare l'intera struttura. Questi strumenti prevedono poche regole per specificare come le componenti possono essere assemblate grazie a funzioni analitiche molto limitate. I sistemi CAD sono stati estesi per supportare dati cartografici ma hanno generalmente un'utilità limitata per la gestione e l'analisi di grandi database geografici. Infatti in un sistema GIS hanno la funzione di arricchirne la cartografia.

Telerilevamento e GNSS. Il telerilevamento, in inglese *Remote Sensing*, è la scienza che permette di acquisire dati e di eseguire misurazioni della terra utilizzando strumenti detti sensori: macchine fotografiche posizionate su aerei, ricevitori GPS, termocamere, dispositivi a scansione. Tali sensori raccolgono dati sotto forma di immagini e segnali in grado di fornire le capacità specializzate per manipolare, analizzare e visualizzare tutte le informazioni ad esse relative. Mancando di

funzioni di gestione ed analisi dei dati geografici, non possono essere considerati dei veri GIS. Un GNSS, da *Global Navigation Satellite System*, è un sistema di geo-radiolocalizzazione e navigazione, terrestre o aerea, che utilizza una rete di satelliti artificiali in orbita e pseudoliti. Detto sistema satellitare globale di navigazione, è dedicato alla fornitura di un servizio di posizionamento geospaziale a copertura globale che permette a piccoli ed appositi ricevitori elettronici di determinare le loro coordinate geografiche (longitudine, latitudine ed altitudine) su un qualunque punto della superficie terrestre o dell'atmosfera, con un minimo errore in metri, elaborando segnali a radiofrequenza trasmessi in linea ottica da tali satelliti [GNSS]. Segue una descrizione del GPS, il GNSS più utilizzato, e vengono accennati il GLONASS, un GNSS russo attuale, e il *Galileo Positioning System*, un GNSS europeo futuro.

GPS

Il GPS è un sistema di posizionamento e navigazione satellitare che consente ad un terminale mobile o ricevitore GPS di ricavare la propria posizione sul globo terrestre, in termini di coordinate geografiche. Questo avviene mediante l'elaborazione del segnale proveniente da una rete di 32 satelliti NAVSTAR (*Navigation Satellite Timing and Ranging*) attivi in orbita attorno alla Terra, disposti in modo che vi sia un contatto privo di ostacoli con almeno quattro di essi. È il sistema satellitare di navigazione più usato nel mondo ed è controllato dal Dipartimento della Difesa degli Stati Uniti. Il suo principio di funzionamento si basa sul calcolo della distanza tra il ricevitore, ovvero dell'unità portatile, e i satelliti rilevati in orbita. Tale calcolo consiste nel determinare il tempo necessario affinché il segnale radio trasmesso da ciascun satellite venga ricevuto dal GPS mobile. Il GPS contiene al suo interno un orologio al quarzo che si sincronizza con gli orologi al cesio dei satelliti, i quali si sincronizzano a loro volta fra loro. poiché la precisione di un orologio al cesio è molto maggiore rispetto a quella dell'orologio interno

del GPS, la precisione richiesta si ottiene confrontando il segnale ed il tempo di percorrenza di più satelliti. Per il calcolo della posizione, lo strumento deve ricevere il segnale di almeno tre satelliti, anche se si necessiterebbe di un quarto satellite per la verifica ed il calcolo dell'altitudine. Il processo di trilaterazione è utilizzato da un ricevitore GPS per determinare la propria posizione sulla superficie terrestre mediante la ricezione di segnali orari provenienti da tre satelliti NAVSTAR. Oltre al segnale orario, ogni satellite della costellazione GPS emette segnali a intervalli regolari: questi segnali vengono ricevuti da dispositivi GPS che successivamente calcolano la distanza tra il dispositivo e ciascun satellite in base al ritardo tra il movimento di invio e quello di ricezione del segnale. I segnali, simili a onde radio, viaggiano alla velocità della luce, ma con un ritardo dovuto all'altitudine dei satelliti. Dopo la ricezione delle distanze relative ad almeno tre satelliti, il dispositivo GPS può eseguire i calcoli in base al processo di trilaterazione. Esso consente l'individuazione di una posizione geografica su una mappa in base alla distanza precisa da tre punti di riferimento diversi [GPS].

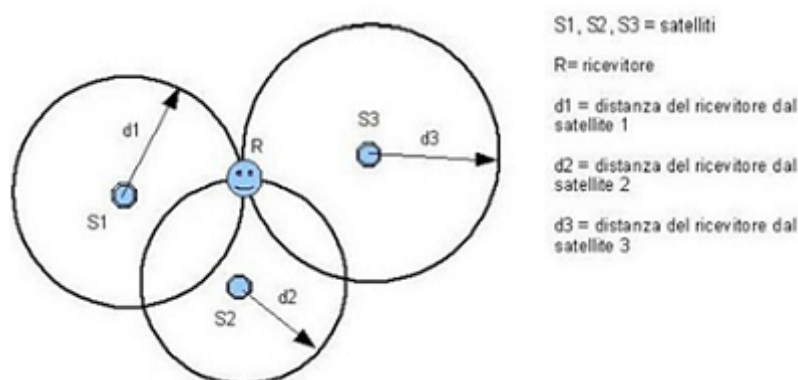


Figura 1.24: Processo di trilaterazione di un ricevitore GPS

Come si può vedere nella *Figura 1.24*, la posizione da individuare è rappresentata dall'intersezione dei tre cerchi con al centro ciascuno dei

punti di riferimento, considerando che il raggio di ogni cerchio corrisponde alla distanza da questi ultimi. Nella versione GPS, i calcoli vengono eseguiti nello spazio tridimensionale: il punto da individuare sarà quindi dato dall'intersezione di tre sfere il cui raggio è costituito dalla distanza da ciascuno dei tre satelliti. Se il dispositivo GPS è in grado di ricevere il segnale di un quarto satellite, è possibile ricontrollare le misurazioni. Alcuni ricevitori GPS utilizzano la trilaterazione 3D. Si considera uno spazio euclideo tridimensionale, centrato al centro della Terra. Si vuole trovare la posizione della ricevente, ossia $T(t_1; t_2; t_3)$, partendo dai dati trasmessi da tre satelliti, ossia la loro posizione $C_i(c_{i,1}; c_{i,2}; c_{i,3})$, e la loro distanza dalla ricevente, ossia ρ_i . La trilaterazione tridimensionale necessita di quattro punti di riferimento per trovare un quinto punto. La condizione che il punto cercato deve trovarsi sulla Terra permette però di determinare la posizione usando solo tre punti di riferimento, e l'altro può essere scartato con una semplice condizione.

Si procede quindi nel modo seguente:

1. si trova un primo piano radicale delle sfere di centro C_1 e C_2 , notato π_1 ;
2. si trova un secondo piano radicale delle sfere di centro C_1 e C_3 , notato π_2 ;
3. si intersecano i piani π_1 e π_2 per trovare la retta $r = \pi_1 \cap \pi_2$;
4. si interseca la retta r con una delle tre sfere. Si trovano così due punti;
5. si tiene presente che il punto cercato deve trovarsi sulla Terra o nella sua atmosfera: in questo modo si elimina il secondo punto, che si trova nello spazio.

Il procedimento è illustrato in *Figura 1.25*.

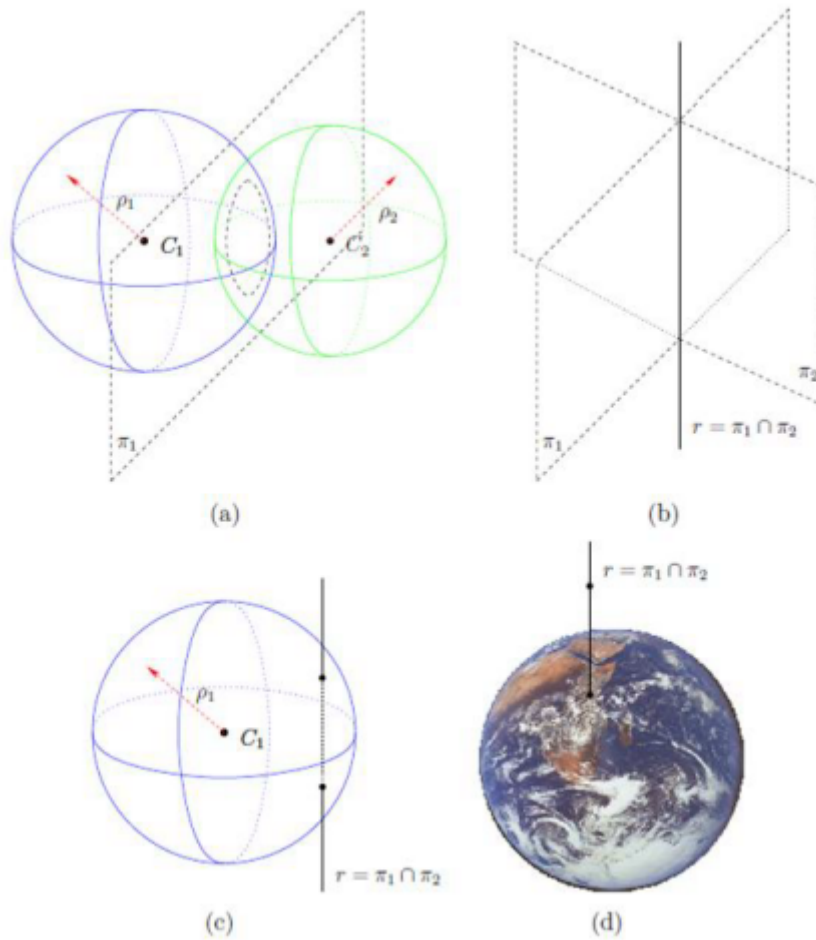


Figura 1.25: Trilazione 3D

- (a) Piano radicale che contiene i punti di intersezione delle due sfere. La stessa cosa viene fatta con altre due sfere in modo da ottenere il secondo piano radicale.
- (b) Intersezione dei due piani radicali trovati, si ottiene una retta.
- (c) Intersezione della retta con una delle tre sfere, si ottengono due punti.
- (d) Il punto più vicino alla Terra è il punto cercato.

Il segnale che riceviamo dai satelliti deve giungere in condizioni ottima-

li, ovvero la trasmissione del segnale attraverso gli strati dell'atmosfera non deve subire variazioni non compensabili. La correzione dell'errore dovuto ad effetti di rifrazione attraverso gli strati più alti dell'atmosfera richiede però l'elaborazione di due segnali distinti, sfasati tra di loro: questo procedimento è effettuabile solamente con apparati militari che ricevono entrambi i due segnali trasmessi dai satelliti. Altri errori sono quelli dovuti al calcolo della posizione dei satelliti, compensati per lo più dai satelliti stessi che aggiornano di continuo fra loro le informazioni sulla loro posizione. Errori più comuni sono, ad esempio, quelli dovuti al segnale riflesso da pareti rocciose, da costruzioni elevate, ovvero quando il percorso fatto dal segnale non è diretto ed è maggiore di quello che dovrebbe essere (il tempo di percorrenza è più lungo); oppure quelli dovuti ad ostacoli naturali e non, che impediscono la ricezione del segnale o consentono la ricezione di un numero insufficiente di satelliti. Per tutti questi motivi è naturale che la precisione del rilevamento della posizione mediante GPS non sia costante, ma vari in base alle diverse situazioni. In generale si può quindi dire che la precisione aumenta sempre con l'aumentare dei satelliti osservati, numero che può variare considerevolmente spostandosi anche di poco, o in momenti successivi della giornata. Il grande vantaggio del GPS è la precisione che può arrivare al metro. Lo svantaggio di questa tecnologia è che non può funzionare dentro agli edifici.

L'A-GPS, da *Assisted GPS*, è un sistema che consente di diminuire i tempi necessari alla prima localizzazione di un terminale GPS [AGPS]. La sua utilità subentra soprattutto nei cosiddetti "canyon" urbani, quali vie strette o viali alberati, in cui è difficile stabilire con precisione la lista di satelliti in vista. Uno dei principali problemi dei terminali GPS classici è relativo al cosiddetto *fixing*, ovvero il rilevamento della posizione di un terminale in fase di accensione del dispositivo: essi devono infatti ricavare la lista dei satelliti in vista in quel momento al fine di potervisi agganciare e poter risalire alla propria locazione. Tale pro-

cesso è alquanto dispendioso in termini di tempo e risorse ed il sistema A-GPS è stato studiato al fine di abbattere tali costi, anche in previsione di un utilizzo su terminali con basse capacità di elaborazione o con risorse energetiche limitate. Lo scopo principale dell'A-GPS è quindi quello di assistere il ricevitore GPS nel calcolo della disposizione, fornendogli informazioni sui satelliti in vista con il supporto dell'operatore di telefonia mobile. Questo sistema sta mostrando una notevole diffusione, ed è normalmente associato ai sistemi di localizzazione (LBS, *Location Based Service*) basati su telefonia cellulare. Si illustra brevemente il funzionamento di un A-GPS. Il sistema si basa sul fatto che ogni cella di telefonia mobile presente sul territorio ha una posizione fissa: sarà quindi la cella stessa a ricavare quali sono i satelliti GPS in vista, istante per istante. Quando un terminale A-GPS vuole conoscere la propria posizione, si collega tramite la rete cellulare ad un *Assistance Server* (che può essere gestito anche dall'operatore stesso), al quale viene inviata anche l'informazione sulla cella alla quale l'utente risulta essere agganciato. Dato che i satelliti in vista dalla cella risultano essere noti, si può assumere che anche il terminale A-GPS veda i medesimi satelliti; pertanto il server elaborerà una lista di satelliti in vista e la invierà attraverso la rete cellulare al terminale, che potrà ricavare immediatamente la propria posizione. La ridotta quantità di dati da inviare ha suggerito l'impiego di messaggi di tipo cell broadcast, già ampiamente utilizzati, ad esempio, per effettuare tariffazioni diverse a seconda dell'area in cui l'utente si trova, e che compaiono sullo schermo dell'utente sotto forma di indicazione della provincia. Il protocollo standard ideato per ottemperare alle funzionalità proprie di A-GPS è il SUPL (*Secure User Plane Location*).

GLONASS

Il sistema sovietico *Global'naya Navigatsionnaya Sputnikovaya Sistema*, o GLONASS, è un sistema satellitare globale di navigazione russo

gestito dalle Forze Spaziali Russe (VKS). Similmente al GPS, la flotta ufficiale GLONASS comprende 31 satelliti Uragan, 24 operativi e quattro di scorta, disposti su tre piani orbitali. Alcuni satelliti sono fuori uso o di emergenza. Ogni piano ospita otto satelliti identificati da un numero di slot che identifica il satellite ed il piano: 1-8, 9-16, 17-24. I tre piani sono separati da 120° ed i satelliti di uno stesso piano distanziati 45° tra loro. Le orbite sono pressoché circolari con una inclinazione di $64,8^\circ$ e un semiasse maggiore di 25440 km. I satelliti orbitano ad una quota di 19100 km, leggermente inferiore ai satelliti GPS e compiono una rotazione ogni 11 ore e un quarto. In questo modo sono visibili in ogni momento almeno cinque satelliti. Dopo anni in cui il sistema è stato penalizzato dal costo dei ricevitori GLONASS e dalla loro difficile reperibilità, a fine 2011 cominciano ad apparire sul mercato dispositivi elettronici di consumo che integrano ricevitori GPS+Glonass (dei quali i più famosi sono il Samsung Galaxy S3, l'iPhone 4S, il Samsung Galaxy Ace 2 e il Nokia Lumia 920). La collaborazione con Android e Apple da parte del Governo Russo ha permesso di integrare ricevitori GLONASS in telefoni molto venduti come il Galaxy S3 [GLONASS].

Galileo Positioning System

Il sistema di posizionamento Galileo, in inglese *Galileo Positioning System*, è un GNSS civile sviluppato in Europa come alternativa al NAVSTAR GPS. La sua entrata in servizio è prevista per il 2014 e conterà 30 satelliti orbitanti su 3 piani inclinati rispetto al piano equatoriale terrestre e ad una quota di circa 24.000 km.

I principali scopi del sistema Galileo sono:

- una maggior precisione nella geolocalizzazione degli utenti rispetto a quella attualmente fornita dal NAVSTAR GPS;

- un aumento della copertura globale dei segnali inviati dai satelliti, soprattutto per le regioni a più alte latitudini (maggiori di 75°);
- un'alta disponibilità del segnale nelle aree urbane;
- una certificata affidabilità, supportata anche dall'invio del messaggio di integrità che avverte immediatamente l'utente di eventuali perdite di integrità nel segnale della costellazione o, viceversa, ne conferma l'esattezza del segnale ricevuto;
- una elevata continuità di servizio che, essendo indipendente dagli USA, potrà funzionare sempre e non verrà disattivato senza preavviso (come successo in tempi di guerra per il GPS).

Sebbene il sistema Galileo sia stato pensato per essere completamente indipendente ed autosufficiente, sarà compatibile ed interoperabile con il sistema GPS; cioè, le caratteristiche di Galileo saranno tali da non interferire con il funzionamento del GPS (principio di compatibilità) e potranno anche essere utilizzati congiuntamente (principio di interoperabilità) [GALILEOS].

DBMS per la gestione dei Dati Spaziali. I DBMS, essendo ottimizzati per memorizzare e recuperare informazioni, sono specializzati nell'archiviazione e gestione di tutti i tipi di dati, inclusi i dati geografici. Infatti vengono utilizzati nei Gis per gestire le tre tipologie di informazioni:

- Geometriche: relative alla rappresentazione cartografica degli oggetti rappresentati; quali la forma (punto, linea, poligono), la dimensione e la posizione geografica;
- Topologiche: riferite alle relazioni reciproche tra gli oggetti (connessione, adiacenza, inclusione ecc ...);
- Informative: riguardanti i dati (numerici, testuali ecc ...) associati ad ogni oggetto.

Vengono adottati diversi diversi approcci:

1) Uso di un DBMS relazionale standard (inadeguato)

Il tema stato può essere rappresentato da una tabella stato contenente gli attributi descrittivi più un attributo (codice) che ne specifica la frontiera. Dato che uno stato può essere costituito da più parti separate (è questo il caso dell'Italia), una tabella frontiera specifica, per ogni parte, i contorni che la definiscono. Una tabella contorno associa ad ogni contorno (una poligonale) un insieme di punti (i vertici della poligonale), specificando l'ordine di tali punti nella poligonale. Infine, una tabella punto fornisce le coordinate di ciascun punto.

Questa soluzione ha diversi limiti tra cui: violazione del principio di indipendenza dei dati (per interrogare i temi occorre conoscere la struttura dei dati spaziali); basse prestazioni (vedi esempio); poca chiarezza (non è user-friendly); difficoltà o impossibilità di esprimere computazioni geometriche (test di adiacenza, window query, point query).

2) Accoppiamento debole

Viene adottata una gestione separata dei dati descrittivi, usualmente con un DBMS relazionale, e dei dati geometrici (è la soluzione adottata da sistemi GIS quali ArcInfo-ESRI, MGE, TiGRis-Intergraph) Questa soluzione ha diversi limiti tra cui coesistenza di modelli dei dati eterogenei e perdita di funzionalità di base di un DBMS (interrogazione, ottimizzazione, ripristino).

3) Approccio integrato

Viene sfruttata l'estendibilità dei DBMS in quanto si aggiungono nuovi tipi di dato e nuove operazioni ad un DBMS relazionale. Tra queste estensioni introdotte si evidenziano: il linguaggio di interrogazione SQL viene arricchito per poter manipolare sia dati tradizionali (descrittivi) sia dati spaziali; vengono introdotti nuo-

vi tipi di dato spaziale (punto, linea, regione); vengono adattate diverse altre funzionalità dei DBMS (ad esempio, l'ottimizzazione delle interrogazioni) per consentire una gestione efficiente dei dati geospaziali. La soluzione dell'approccio integrato viene adottata da Oracle, Postgres e MySQL.

Capitolo 2

Le Reti di Sensori

Una “Rete di Sensori Wireless”, spesso indicata con l’acronimo WSN (*Wireless Sensor Networks*, Fig 2.1), è una rete formata da piccoli dispositivi autonomi, denominati nodi sensore (*Sensor Node*), in grado di prelevare dati dall’ambiente circostante e di comunicare tra loro. Questi dispositivi sono formati da componenti in grado di rilevare grandezze fisiche (come temperatura, pressione, intensità sonora, vibrazioni, umidità, pressione atmosferica, ecc.), di elaborare dati e di interagire tra loro. Sono prodotti e distribuiti in massa, hanno un costo di produzione trascurabile e sono caratterizzati da dimensioni e pesi molto ridotti. Ogni sensore ha una riserva d’energia limitata e non rinnovabile e, una volta messo in opera, deve lavorare autonomamente; per questo motivo tali dispositivi devono mantenere costantemente i consumi molto bassi, in modo da avere un maggior ciclo di vita. Per ottenere la maggior quantità possibile di dati occorre effettuare una massiccia distribuzione di sensori (nell’ordine delle migliaia o decine di migliaia) in modo da avere un’alta densità (fino a 20 nodi/m³) e far sì che i nodi siano tutti vicini tra loro, condizione necessaria affinché possano comunicare. I nodi sensore sono sparpagliati in un’area chiamata *area di sensing*. Ciascun nodo, all’interno di questa area, ha la capacità di accumulare e di instradare i dati fino al nodo *sink* ed infine all’utente finale. Il nodo sink consiste in un’antenna capace di illuminare tutto il dominio occupato dai nodi sensori, ovvero l’area

di sensing. La posizione dei nodi all'interno della rete non deve essere predefinita in quanto questo consente di utilizzare questa tecnologia in posti difficilmente accessibili o in operazioni di soccorso in luoghi disastrati per i quali è necessaria una disposizione random dei nodi. Ciò significa che gli algoritmi e i protocolli utilizzati nelle reti di sensori devono possedere capacità auto-organizzative. I sensori, conoscendo le caratteristiche di trasmissione del sink, e sfruttando il suo segnale di illuminazione (che può essere, o meno, sentito), hanno la capacità rilevare la loro posizione, permettendo quindi la distribuzione random di questi nella rete. L'autoapprendimento di posizione dei sensori è quindi una delle loro caratteristiche più importanti ma, vista la scarsa quantità di energia di cui è dotato un sensore, è opportuno utilizzare algoritmi di geolocalizzazione in grado di ottimizzare le tempistiche di questo procedimento [21, WSN].

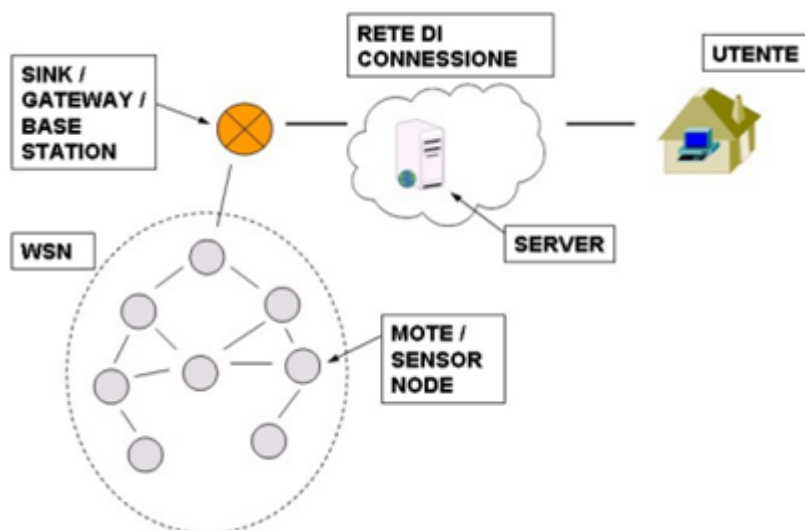


Figura 2.1: Struttura tipica di una WSN

2.1 La Storia delle Reti di Sensori

Si cominciò a parlare di Reti di Sensori dall'epoca della guerra fredda. In questo periodo negli Stati Uniti furono sviluppate delle reti acustiche di sorveglianza sottomarina, utilizzate ancora oggi per rilevare attività sismiche sottomarine, e furono create nel Nord America delle reti radar per la difesa aerea, i cui sensori erano costituiti da aerei AWACS (*Airborn Warning and Control System*). Successivamente nei primi anni 80 il DARPA (*Defence Advanced Research Projects Agency*) sponsorizzò la creazione di DSN (*Distributed Sensor Network*) per determinare se protocolli TCP/IP, insieme alla rete ARPA, potessero essere usati come reti di sensori. Le DSN prevedevano diversi nodi autonomi equamente spazati e in grado di collaborare tra loro, il cui obiettivo era quello di raccogliere i dati in nodi con lo scopo di aumentarne l'efficienza. Le applicazioni delle DSN erano il calcolo distribuito, il signal processing e il tracking. I componenti di queste reti erano sensori acustici, protocolli di comunicazione di alto livello, algoritmi di calcolo e software distribuito. Dagli anni 80 ai 90 furono sviluppati i primi prodotti commerciali derivanti dalla tecnologia delle DSN che, proprio grazie alla commercializzazione, erano caratterizzati da un costo minore e dall'implementazione dei primi standard. Queste reti di sensori venivano utilizzate in situazioni di *network-centric warfare*, ovvero in ambienti in cui i sistemi d'arma devono cooperare tra di loro scambiandosi informazioni sull'obiettivo. I sensori infatti erano in grado di tracciare obiettivi multipli, anche molto distanti, con ottimi tempi di risposta; vennero sfruttati per creare reti di radar per rilevare obiettivi aerei, reti di sensori acustici negli oceani per rilevare sottomarini e reti di sensori schierate sul campo di battaglia. Dopo l'evoluzione tecnologica degli anni 90 e dei primi anni del 2000, si è giunti ad una nuova generazione di sensori definita come seconda generazione di prodotti commerciali. Il progresso ha consentito lo sviluppo di tecnologie ad alta densità: i MEMS (*Micro Electro-Mechanical Systems*), cioè sistemi "intelligenti" che, integrando la tecnologia dei sensori, abbinano funzioni elettroniche, ottiche, biologiche, chimiche meccaniche e di gestione dei fluidi in uno spazio ridotto

al micrometro; e i recenti NEMS (*Nanoscale Electromechanical Systems*), ovvero dispositivi che integrano funzionalità elettriche e meccaniche nell'ambito della nanoscala. La diffusione delle WSN, favorita da una standardizzazione, la nascita di sistemi come *Bluetooth*, *WiFi*, *Wimax* e *ZigBee*, che consentono una connettività totale, e nuovi processori, caratterizzati da un basso costo e un minimo consumo energetico, ci consentono oggi di adoperare le Reti di Sensori per molteplici applicazioni [23, 24].

2.2 Classificazione delle Reti di Sensori e Campi di Applicazione

Una prima classificazione delle reti informatiche in generale può essere fatta in funzione dell'area coperta, cioè della dimensione fisica della rete. Nella Figura 2.2 vengono riportate le dimensioni caratteristiche di queste reti.

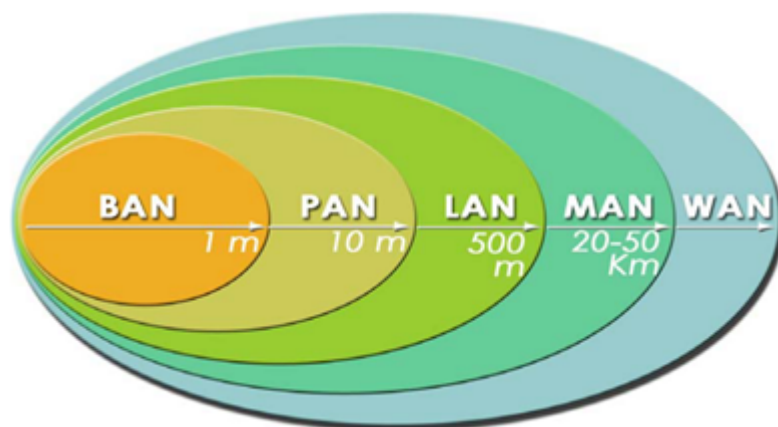


Figura 2.2: Classificazione delle reti in funzione dell'area coperta (tratto da [22])

- Le BAN (*Body Area Network*) sono reti che interconnettono dispositivi indossabili, il cui raggio di copertura è inferiore ad un metro. Questa rete permette il passaggio di comunicazione attraverso segnale elettrico o luminoso a seconda del link di comunicazione utilizzato. Le applicazioni più tipiche sono le Reti di Sensori corporei in campo biomedico [BAN].
- La PAN (*Personal Area Network*) è una rete informatica personale che può essere utilizzata per collegare i vari dispositivi vicini a un singolo utente (telefoni, palmari, ecc.) tra di loro in modo da consentire scambio di informazioni o per consentire la connessione a reti di livello più alto (come per esempio Internet). I singoli dispositivi possono anche non appartenere all'utente in questione [PAN].
- La LAN (*Local Area Network*) è una tipologia di rete informatica contraddistinta da un'estensione territoriale non superiore a qualche chilometro. Si tratta storicamente delle prime tipologie di reti informatiche realizzate al mondo per semplicità di realizzazione e costi sostenibili anche da piccoli privati. L'implementazione classica di LAN è quella che serve ad esempio un'abitazione o un'azienda all'interno di un edificio, o al massimo più edifici adiacenti fra loro (campus). Le reti LAN rappresentano attualmente l'unità minima di base della rete Internet essendo queste interconnesse tra loro a formare reti metropolitane (MAN) [LAN].
- La MAN (*Metropolitan Area Network*) è una rete di trasporto storicamente nata per fornire servizi di tv via cavo alle città dove c'era una cattiva ricezione terrestre. In pratica un'antenna posta su una posizione favorevole distribuiva poi il segnale alle case mediante cavo. Attualmente utilizza la fibra ottica come mezzo trasmissivo di collegamento e topologie di rete ad anello, detti appunto anelli metropolitani, con funzionalità di collegamento metropolitano da e verso le sottoreti e

le utenze interne all'area metropolitana stessa. L'interconnessione di più MAN dà vita a reti WAN [MAN].

- La WAN (*Wide Area Network*) è una rete di trasporto che può connettere più reti locali e/o metropolitane collegate tra loro da backbone¹. Molte WAN sono costruite per una particolare organizzazione e sono private, come ad esempio la rete GARR² (*Gruppo per l'Armonizzazione delle Reti della Ricerca*). Invece Internet, la più grande WAN mai realizzata con una rete di computer che copre l'intero pianeta, è ad accesso pubblico. La funzionalità di questa rete è generalmente la connessione tra host [WAN].

Le WSN rientrano usualmente nel panorama delle PAN, in cui sono diffuse molte tecnologie wireless, e sono caratterizzate dallo standard IEEE 802.15.4. Questo tipo particolare di protocollo è stato progettato per offrire, rispetto agli altri della stessa famiglia IEEE 802.15, la possibilità di poter realizzare reti costituite da un gran numero di nodi, quindi con una densità spaziale anche molto elevata, con una quantità di banda richiesta piuttosto limitata (low-data-rate protocol) e un basso consumo energetico (low-power protocol). Una classificazione specifica delle WSN può essere fatta in base alla finalità della rete e di seguito vengono riportate le varie categorie.

Monitoraggio. Questo tipo di rete viene utilizzata per tracciare in maniera continuativa una certa grandezza. La sorgente da monitorare può essere un singolo sensore o una sottorete da cui proviene un aggregato di dati. Una rete di questo genere richiede un campionamento solitamente isocrono e fisso con un consumo energetico considerevole.

Riconoscimento di eventi. Un altro importante scenario applicativo risulta essere l'*event detection*: la rete deve accorgersi di situazioni di

¹un collegamento ad alta velocità di trasmissione e capacità tra due server o router di smistamento informazioni

²una rete telematica nazionale a banda ultralarga dedicata al mondo dell'Università e della Ricerca

“allarme”, ossia di quando una determinata grandezza esce dai livelli stabiliti. A livello energetico questo scenario è potenzialmente meno pesante del precedente, in quanto il nodo entra in uno stato attivo solo in casi eccezionali.

Classificazione di oggetti. La finalità di queste reti è il riconoscimento di alcune grandezze tra un insieme di prototipi noti. Questo implica un carico computazionale superiore agli altri tipi di rete.

Tracciamento di oggetti. In questo caso la rete funge da sistema di sorveglianza, riconoscendo e stimando la posizione di alcuni oggetti in una determinata area geografica.

Dopo queste classificazioni, le WSN si distinguono ulteriormente in base ai diversi campi di applicazione come illustrato di seguito.

Monitoraggio ambientale (*Environmental and habitat monitoring*).

In campo ambientale le applicazioni possono essere molteplici: monitorare movimenti di animali, studiare particolari habitat e valutare l'impatto ambientale di alcune scelte umane con sistemi di monitoraggio di agenti inquinanti. Appartengono a questo settore anche il rilevamento di eventi naturali catastrofici quali incendi, tornado, terremoti ed eruzioni vulcaniche.

Monitoraggio Agricolo (*Agriculture sensing*). Nel settore agricolo le

WSN vengono utilizzate per il monitoraggio di particolari situazioni ambientali; in particolar modo la cosiddetta “agricoltura di precisione” utilizza le reti di sensori per rilevare in tempo reale il livello di pesticidi nell'acqua o nei terreni agricoli, per debellare insetti, con opportuni trattamenti, o per intervenire in situazioni critiche.

Monitoraggio di strutture (*Structural Health Monitoring*). Le Re-

ti di Sensori posizionate sulle strutture rilevano lo stato di salute di edifici, ponti, case sottoposte a sollecitazioni esterne; in alternativa,

potrebbero essere utilizzate anche per misurare difetti strutturali di componenti.

Monitoraggio di apparecchiature industriali (*Industrial sensing*).

I sensori wireless possono essere applicati a macchinari industriali al fine di analizzare il comportamento dei componenti sottoposti a stress meccanico, di migliorare le performance o di prevenire rotture e guasti. Altre tipiche applicazioni industriali possono essere la realizzazione di bridge wireless verso altre reti già esistenti come *DeviceNet* o *FieldBus*, creando un'interfaccia che possa consentire il monitoraggio remoto e la modifica dei parametri di funzionamento di dispositivi connessi alle reti preesistenti.

Controllo del traffico veicolare (*Traffic control*). Un sistema di sensori finalizzato al monitoraggio del traffico; hanno il compito di controllare il passaggio di automobili, di analizzare la velocità e l'intensità del traffico o di individuare eventuali blocchi o situazioni anomale.

Sorveglianza di edifici (*Infrastructure control*). Questo tipo di reti può essere utilizzato come ausilio per la sorveglianza di centri commerciali o di luoghi a rischio, come stazioni ferroviarie o aeroporti.

Sorveglianza militare (*Militar control*). Le Reti di Sensori sono state utilizzate in principio per questo scopo. Le loro applicazioni spaziano dal monitoraggio sullo stato o la posizione delle forze sul campo di battaglia alla sorveglianza di luoghi strategici. Possono inoltre essere utilizzate per rilevare attacchi nucleari o biochimici oppure dispiegate in luoghi ostili al fine di raccogliere informazioni sugli spostamenti del nemico, essendo caratterizzate da un elevato numero di nodi dal costo contenuto. Infatti l'eventuale distruzione di uno o più nodi non influenza l'efficienza della rete, cosa che invece potrebbe accadere utilizzando reti cablate.

Applicazioni Mediche (*Personal Health Care*). Nel campo della medicina, le WSN sono oggi impiegate per monitorare pazienti, eseguire valutazioni diagnostiche, amministrare farmaci in ospedale e monitorare a distanza dati fisiologici. Ma questo scenario è solo il primo passo verso una Rete di Sensori molto piccoli e distribuiti all'interno dei tessuti in grado di misurare la temperatura, la pressione sanguigna e molti altri parametri di rilievo. Le prime applicazioni in questo senso si hanno nei dispositivi di *personal healthcare*, integrati in orologi da polso o bracciali appositi, utilizzati per rilevare il battito cardiaco o le calorie bruciate durante l'attività fisica.

Applicazioni Personali (*Personal sensing*). Nel settore della domotica³, le Reti di Sensori riescono a fornire servizi all'utente all'interno della propria abitazione, ad esempio informandolo tempestivamente di eventuali guasti. I sensori possono essere introdotti negli apparecchi elettrici e questi, interagendo tra loro e con una rete esterna o Internet stesso, permettono all'utente di comandare facilmente gli elettrodomestici a distanza. Un progetto di una casa intelligente di questo tipo può prevedere due diversi approcci progettuali: un sistema *human centered*, che prevede che la tecnologia sia in grado di rispondere alle esigenze dell'utente finale in termini di interazione input/output; oppure un sistema *technology centered*, che vuole creare un cosiddetto *smart environment*⁴. Un'altra applicazione importante è la gestione del riscaldamento, della ventilazione e dell'illuminazione. L'uso di sensori wireless riduce i costi di installazione e rende molto meno invasiva la modifica di sistemi esistenti (si pensi alla necessità di forare le pareti: con la tecnologia wireless questo non è più necessario) [23, 24].

³la scienza che si occupa dello studio delle tecnologie atte a migliorare la qualità della vita nella casa

⁴ambiente domestico intelligente dove ogni dispositivo integra uno smart device per comunicare con un server di stanza a sua volta collegato con vari i server delle sale della casa

2.3 La Struttura delle Reti di Sensori

Con il termine topologia si indica la modalità in cui i diversi dispositivi della rete vengono disposti e, in particolare, i collegamenti fisici e logici che li interconnettono. Una Rete di Sensori può essere realizzata in tre differenti topologie, riassunte anche nella Figura 2.3: *stella (star)*, *mesh* e *albero (tree)*.

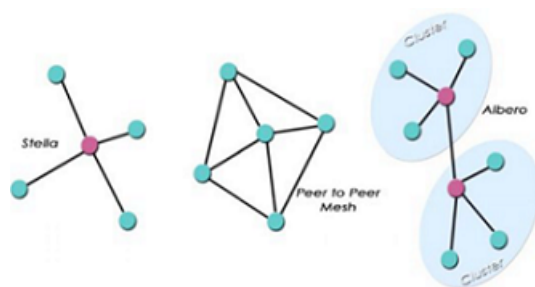


Figura 2.3: Possibili topologie di rete nelle WSN (tratto da [21])

La topologia di *rete a stella* prevede che tutti i nodi siano connessi direttamente a un unico nodo coordinatore tramite un tratto dedicato. Per cui ogni nodo della stella per poter comunicare con l'altro deve inviare il segnale attraverso il nodo centrale che lo smisterà al destinatario. I vantaggi di questa rete sono l'alta affidabilità e l'alta scalabilità: l'unico punto debole della rete è il coordinatore, detto anche *hub*; se si guasta tutta la rete cessa di funzionare. La funzionalità non risulta però dipendente dai singoli nodi o dai singoli tratti, inoltre, è possibile aggiungere o togliere nodi e connessioni senza modificare la rete, fino al massimo previsto dall'hub. Un'ulteriore vantaggio che presenta questa topologia è rappresentato dal fatto che è possibile collegare tra di loro due reti stella, collegando quindi due hub. Questo aumenta il livello di scalabilità e in caso di guasto di un nodo centrale una parte della rete rimarrà attiva. La *struttura a mesh*, detta anche *rete magliata*, è formata da nodi sensori interconnessi tra loro tramite tanti canali quanti sono i nodi nella rete, ma in alcune tipologie, dette *magliate non completamente connesse*, non sono presenti tutti i rami teoricamente possibili. In questa rete è necessario definire chiare regole di routing per lo scambio di

messaggi tra i vari nodi. Può esistere anche un coordinatore e trovarsi al di fuori della struttura mesh; in questo caso sarà collegato ad uno solo dei nodi ed avrà funzioni di controllo limitate. Se non esiste un coordinatore tutti i nodi avranno funzioni di coordinamento e raccolta dati in modo paritetico. Questo tipo di rete richiede costi elevati ma presenta un'alta affidabilità dato che, a fronte del guasto di uno o più nodi o rami, risulta possibile identificare percorsi alternativi che aggirano i punti di guasto. Infine, la *struttura ad albero* è una topologia di rete che combina la struttura a stella e quella a mesh. Nodi e canali sono disposti in modo ramificato: più nodi, che fanno capo ad un unico nodo, formano un gruppo chiamato cluster. I nodi coordinatori possono far riferimento a loro volta ad un nodo coordinatore di livello superiore costruendo, così, vari livelli di gerarchia. Questa topologia è la più adatta alle reti WAN anche perché comporta costi contenuti; tuttavia se si guasta un canale o un nodo d'origine viene compromessa la funzionalità di parte della rete o dell'intera rete. Un'ulteriore distinzione va fatta per la scelta di WSN *centralizzate* o *distribuite* (Fig 2.4). Nel primo caso i compiti verranno assegnati a particolari nodi (ad esempio, l'elaborazione dei dati potrebbe venir completamente demandata alla stazione base), nel secondo sarà la rete stessa a provvedere, distribuendo i compiti a tutti i nodi che la compongono. Nonostante il fatto che una struttura centralizzata sembri più facile ed efficiente, è importante notare che una Rete di Sensori è spesso composta da dispositivi omogenei sia dal punto di vista delle caratteristiche hardware sia da quello del Sistema Operativo, e che pertanto caricare solo alcuni nodi di compiti specifici introdurrebbe forti disomogeneità nella struttura della rete, compromettendone la longevità [21, 22, 24].

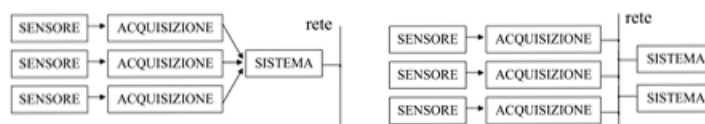


Figura 2.4: Sistema Centralizzato e Sistema Distribuito (tratto da [22])

2.4 Scelte tecnologiche nella progettazione di una Rete di Sensori

Il progetto di una Rete di Sensori richiede la valutazione di diversi fattori tecnologici che possono guidare il progettista ad effettuare una scelta fra le diverse tecnologie disponibili. Alcuni di questi aspetti si possono riassumere nelle caratteristiche illustrate di seguito.

Tolleranza ai guasti. Un nodo sensore potrebbe cessare di funzionare a causa della mancanza di energia, essere danneggiato (o distrutto) o perdere il collegamento wireless con gli altri nodi per via della presenza di un ostacolo di disturbo non previsto. In una simile situazione non deve però essere pregiudicato il corretto funzionamento della rete. Con tolleranza ai guasti, in inglese *fault tolerance*, viene indicata la capacità di una Rete di Sensori di mantenere le sue funzionalità in corrispondenza del malfunzionamento dei suoi nodi. Possiamo definire un modello per questa tolleranza come una funzione del tasso di decadenza dei nodi, cioè l'indice del malfunzionamento dei nodi, e del tempo. Ogni qualvolta, ci si appresta al progetto di una WSN, si dovrà tenere presente quale sia la tolleranza richiesta, e rispettarla implementando opportuni algoritmi.

Lifetime. Nella maggior parte degli scenari applicativi, i nodi sensore sono alimentati da batterie che non è possibile o non è conveniente sostituire. In ogni caso, l'obiettivo di una WSN è quello di rimanere attiva più a lungo possibile o, almeno, per la durata della sua funzione. Quindi, il risparmio energetico assume un ruolo fondamentale in una Rete di Sensori. Anche nelle applicazioni in cui è previsto l'ausilio di un piccolo pannello solare per ricaricare la batteria dei nodi, il risparmio energetico rimane un fattore critico e in queste situazioni l'obiettivo è quello di mantenere la rete attiva per un tempo indeterminato. Tuttavia, l'implementazione di meccanismi di risparmio energetico richiede

dei compromessi con la qualità del servizio: la soluzione è ovviamente quella di trovare un giusto bilanciamento tra le due caratteristiche. La definizione di *lifetime* non è univoca, nel senso che dipende dall'applicazione che si vuole misurare: a volte infatti si indica con *lifetime* il tempo entro il quale il primo nodo della rete finisce la propria energia o smette di funzionare; altre volte invece corrisponde al momento in cui il 50% dei nodi vengono persi; oppure indica la prima volta in cui una regione sotto controllo non è più monitorata da alcun nodo.

Scalabilità. Una WSN deve essere progettata in modo da realizzare un'architettura in grado di supportare molti nodi. Infatti ci si aspetta di individuare anche centinaia di sensori disposti molto vicini fra loro che possano raggiungere in generale densità nell'ordine dei 20 dispositivi per metro quadrato. Una densità così elevata deve essere tenuta in considerazione in sede di progetto per garantire il corretto funzionamento della rete stessa. Inoltre si dovrà considerare il fatto che i singoli nodi possono evolvere nel tempo sia in termini di posizione che di funzionamento. Si pensi per esempio ad una abitazione che contenga diversi dispositivi in ogni elettrodomestico; oppure si consideri la densità di dispositivi che si può raggiungere quando una serie di persone dotate di più *smart sensor* si trovano in unico luogo come una lezione universitaria o uno spettacolo teatrale. In questi particolari contesti si dovrà garantire che i diversi dispositivi siano in grado di interagire fra loro senza generare mutue interferenze.

Programmabilità. I nodi devono essere riprogrammabili, cioè essere in grado, in qualunque momento, di poter cambiare i propri compiti. In questo ambito si riscontrano diversi studi relativi a reti che riescono a riprogrammarsi autonomamente senza dover programmare manualmente ogni singolo nodo.

Mezzo trasmissivo. La comunicazione senza fili fra i diversi dispositivi in una WSN, può essere fatta impiegando diversi sistemi: induttivi, ottici

(o ad infrarossi), a ultrasuoni ed a radio-frequenza. Ciascuno di questi mezzi presenta dei pregi e dei difetti. Però è importante sottolineare che questi sistemi, affinché siano presi in considerazione all'interno di un progetto, devono essere autorizzati in ogni parte del mondo, evitando problemi burocratici che obbligherebbero a una re-ingegnerizzazione continua da uno stato all'altro. Di seguito si analizzano brevemente i mezzi trasmissivi citati.

- *Sistemi induttivi.* La tecnologia che sfrutta questo sistema è la *Radio Frequency Identification* (RFID), e consiste nell'utilizzare dei tag attivi o passivi che vengono letti da apposite porte di lettura. Queste porte generano un forte campo elettromagnetico che crea una corrente indotta all'interno del tag, permettendo la lettura/scrittura di informazioni su di esso. I principali svantaggi di questo tipo di tecnologia risultano essere le distanze, per cui è possibile effettuare correttamente operazioni di *lettura/scrittura*, e l'elevato consumo di potenza per generare il campo elettromagnetico.
- *Sistemi ottici.* In questo gruppo includiamo le tecnologie come l'*Infrared Data Association* (IrDA) ed altri standard che prevedono l'uso di tecnologie di tipo ottico. Il principale svantaggio risulta essere la necessità di un allineamento dei dispositivi, cioè ricevitore e trasmettitore devono essere correttamente collegati da una linea di vista diretta (*line of sight*). Al contempo tecnologie di questo tipo garantiscono bassi costi ed alta affidabilità.
- *Ultrasuoni.* Questi sistemi consentono di superare la limitazione della linea di vista tipica dei sistemi ottici, ma lo svantaggio è l'elevata energia richiesta dal coordinatore della rete per la generazione di ultrasuoni. Ciò comporta elevate dimensioni e alti costi per i dispositivi che utilizzano queste tecnologie.
- *Radio-frequenza (RF).* L'utilizzo di sistemi in radio frequenza con-

sente di superare i problemi legati alla linea di vista e garantisce dimensioni contenute dei dispositivi e consumi molto bassi. Però la disponibilità del mezzo trasmissivo risulta un problema rilevante, poiché le bande di frequenza vengono assegnate attraverso legislazioni locali. Infatti queste limitazioni, inerenti la potenza di uscita delle trasmissioni radio e dovute a ragioni legislative, hanno lo scopo di minimizzare l'inquinamento elettromagnetico ed in secondo luogo di ridurre i costi. Le bande ISM (*Industrial, Scientific and Medical*) rappresentano una delle migliori scelte effettuabili, visto che sono disponibili in molti stati. Per rispettare il principio dei costi ridotti e delle dimensioni contenute dei dispositivi delle WSN, vengono considerate solo alcune di queste bande anche perché hanno il vantaggio di essere svincolate dall'utilizzo di un particolare standard di comunicazione, consentendo quindi al progettista ampia libertà nell'implementazione di tecniche di *power saving*, fondamentali nelle Reti di Sensori.

Auto-mantenimento. Dato che sia la WSN sia l'ambiente in cui si trova sono in continuo mutamento, la rete deve essere in grado di adattarsi, monitorando il proprio stato di salute, aggiornando i propri parametri e decidendo tra nuovi compromessi (come diminuire la qualità del servizio quando l'energia sta per terminare).

Densità di nodi non uniforme. In una Rete di Sensori possono esserci zone con molti nodi e zone con pochissimi nodi sparsi. La densità dei nodi può variare nello spazio e nel tempo (perché i sensori possono esaurire l'energia) e la rete deve essere in grado di adattarsi a queste variazioni.

Type of Service. Le WSN non sono delle reti tradizionali che si limitano a comunicare dei bit, bensì devono provvedere informazioni utili o agire ad eventi particolari. Per questo motivo servono nuovi paradigmi, nuove interfacce utente e nuovi modi di pensare ai servizi della rete.

Quality of Service (QoS). I parametri QoS adottati tradizionalmente come *ritardo*, *banda minima* e *jitter*, non servono se si stanno utilizzando software che tollerano i ritardi e se i nodi scambiano pochi pacchetti alla volta. Inoltre, in alcuni casi basta ricevere qualche dato ogni tanto, mentre in altri tutti i dati devono essere ricevuti oppure devono essere ricevuti entro un certo tempo. Ciò che conta è quindi la quantità e la qualità dei dati ricevuti da un nodo di una certa area. Ad esempio, dei parametri validi possono essere l'affidabilità nel rilevamento di certi eventi o l'approssimazione di certe misure. Per includere tutti questi aspetti in un progetto di una WSN, sono stati sviluppati nel corso degli anni diversi meccanismi per la comunicazione come architetture di sistema e protocolli di varie tipologie. Con tali sviluppi si riescono ad ottenere le seguenti importanti funzionalità.

Autoconfigurazione. La rete deve essere in grado di configurare automaticamente tutti i suoi parametri vitali come la gestione autonoma dell'ingresso di un nuovo nodo o l'aggiornamento delle tabelle di routing dopo la perdita di alcuni nodi.

Collaborazione e in-network processing. In alcune applicazioni, affinché ogni singolo nodo sia in grado di capire se si è verificato un evento, è necessaria una collaborazione fra i nodi anche attraverso *in-network processing*. Grazie a questo processo si riduce la quantità di dati trasmessa nella rete eseguendo calcoli sui dati, tipicamente *data-aggregation* (come il calcolo della temperatura media di una zona), oppure sfruttando la correlazione tra misurazioni di più sensori.

Connessioni wireless multihop. La comunicazione diretta tra due nodi non è sempre possibile (per via di ostacoli fisici o di una distanza elevata tra i nodi) e l'utilizzo di una potenza trasmittiva elevata comporterebbe un rapido esaurimento della batteria. Quindi la soluzione è quella di adoperare dei nodi che fungano da *relay* verso altri nodi (Fig. 2.5)).

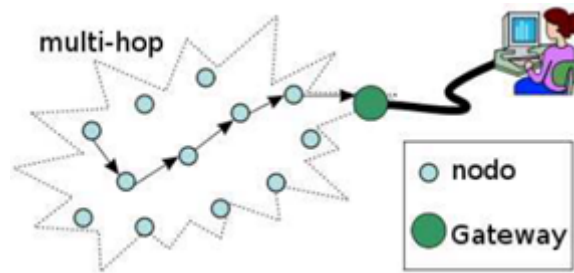


Figura 2.5: Rappresentazione del multi-hop

Operazioni energeticamente efficienti. Nel progetto di una WSN, risulta importante tenere conto del risparmio energetico per tutte le operazioni compiute al fine di renderle efficienti; durante questa fase si cerca di evitare la formazione di *hotspot*, ovvero regioni o gruppi di nodi che esauriscono la propria energia molto più rapidamente degli altri.

Data-centric. In una rete di comunicazione tradizionale, lo scambio di dati avviene tra entità aventi ognuna un indirizzo di rete specifico (architettura *address-centric*) mentre in una WSN ci si focalizza piuttosto sulla regione di provenienza del dato. Infatti ciò che interessa è richiedere una certa informazione ad una certa area monitorata e non richiedere una certa informazione ad un certo nodo.

Località. Per risparmiare risorse hardware, si preferisce far memorizzare localmente al nodo informazioni di routing solo verso i nodi vicini a lui. Così facendo le risorse hardware occupate rimarrebbero inalterate anche se la rete dovesse crescere esponenzialmente. Chiaramente, conciliare località e protocolli di routing efficienti è una delle sfide da affrontare.

Bilanciamento dei trade-off. Durante la fase di progettazione della WSN e anche durante il suo runtime, occorre bilanciare i diversi *trade-off* che si possono presentare. Fra questi “compromessi” si ricordano: lifetime

e qualità del servizio; lifetime della rete e lifetime del singolo nodo; densità della rete ed efficienza del routing [21, 23, 24].

2.5 Esempio pratico di una Rete di Sensori: Minos System

Dopo aver introdotto le Reti di Sensori in maniera generica, in questo paragrafo viene presentato “*Minos System*”: il sistema più avanzato di telegestione dell’illuminazione pubblica ed esterna, progettato dall’azienda *UMPI Elettronica S.r.l.* Non appartiene alla categoria WSN, inteso come *Wireless Sensor Network*, ma a quella delle *Wired Sensor Network*, ovvero reti in cui i nodi sensore comunicano attraverso un canale di trasmissione *Powerline* (PLC⁵). Inoltre questa tecnologia innovativa, che rientra nell’ambito della Smart City⁶, funge da sistema di controllo e telegestione degli impianti di pubblica illuminazione: ogni lampada viene dotata di un dispositivo di controllo e diagnostica, chiamato SYRA (Fig. 2.7), facenti capo ad un dispositivo PLC⁷ proprietario, chiamato ANDROS (Fig. 2.8). Un impianto è normalmente composto da più ANDROS, ciascuno dei quali è in grado di controllare fino a 1.022 SYRA. Tali elementi vengono suddivisi gerarchicamente come è indicato in Figura 2.6.

⁵*Power Line Communication*, in italiano onde convogliate, è una tecnologia per la trasmissione di voce o dati che utilizza la rete di alimentazione elettrica come mezzo trasmissivo.

⁶Dall’inglese “città intelligente”, indica un ambiente urbano in grado di agire attivamente per migliorare la qualità della vita dei propri cittadini.

⁷*Programmable Logic Controller*, in italiano controllore logico programmabile, è un controllore per industria specializzato in origine nella gestione o controllo dei processi industriali. Il PLC esegue un programma ed elabora i segnali digitali ed analogici provenienti da sensori e diretti agli attuatori presenti in un impianto dotato di questa tecnologia.

Figura 2.6: Gerarchia *Minos System*

Minos System per come è stato progettato porta diversi vantaggi; fra questi si evidenziano:

- un risparmio energetico tra il 30 e il 45% dei consumi attuali evitando sprechi e programmando i comandi di spegnimento, accensione o riduzione del flusso luminoso di ogni punto luce;
- una riduzione fino al 55% dei costi generali di manutenzione fornendo in tempo reale tutte le informazioni sullo stato degli impianti;
- una diminuzione dell'inquinamento atmosferico, attraverso la riduzione dei consumi energetici e la programmazione degli interventi di manutenzione, e dell'inquinamento luminoso, attraverso l'ottimizzazione dell'utilizzo della luce aumentando così la vita media delle lampade;
- la trasformazione del lampione in un punto intelligente in grado di erogare servizi aggiuntivi quali videosorveglianza, info-mobilità e comunicazione WI-FI a basso costo, senza necessità di cablaggi aggiuntivi.

A livello applicativo questo sistema viene impiegato, grazie ai suoi innumerevoli vantaggi e funzionalità, in diversi ambiti come l'illuminazione pubblica e di opere artistiche, in stazioni ferroviarie, in autostrade, in aree industriali, in tunnel e in parcheggi. Di seguito verranno presentate le peculiarità dei dispositivi *Minos System* (SYRA e ANDROS) per la telegestione dell'illuminazione sempre prodotti dalla *UMPI Elettronica S.r.l.*

2.5.1 SYRA

Questo sensore funge da modulo di diagnostica e comando di accensione o spegnimento del punto luce per lampada; ne esistono vari modelli che si differenziano più che altro per la tipologia di lampada gestita. Come già indicato in precedenza, il mezzo trasmissivo utilizzato fra i vari dispositivi è una linea di alimentazione con tecnologia *Powerline* (PLC). Per quando riguarda le funzionalità, SYRA ha il compito di rilevare gli stati di funzionamento della lampada, quali accensione e corretto funzionamento, e di segnalare diverse anomalie quali: condensatore guasto, mancante, o insufficiente; bassa potenza della lampada; lampada interrotta o accenditore difettoso; lampada non collegata, lampeggiante o in esaurimento; fusibile guasto o collegamento al dispositivo SYRA interrotto.



Figura 2.7: Tipologia di SYRA (tratto da [MINOSX])

2.5.2 ANDROS

Questo sensore funge da modulo di controllo quadro per la gestione di linee e di lampade su *Powerline*; in pratica ANDROS raccoglie, elabora e memorizza le informazioni e le segnalazioni del quadro e della lampada provenienti dai vari dispositivi e dai propri ingressi. Gestisce fino a 1.022 moduli SYRA ed è configurabile sia localmente che da remoto. Le sue varie funzionalità comprendono: il controllo di stato della rete di alimentazione per rilevare l'assenza di rete o la presenza di una batteria scarica e/o insufficiente; il controllo degli eventi a livello armadio come l'attivazione o la disattivazione degli ingressi, la presenza o l'assenza di tensione, un'anomalia nei moduli collegati e il malfunzionamento di più lampade; l'invio di comandi al modulo SYRA di accensione o spegnimento e una riduzione di flusso su singolo punto luce; l'attivazione di chiamate d'emergenza sui principali eventi; una programmazione dei cicli di comando attraverso orologio astronomico e timer settimanali; infine la visualizzazione sul display dello stato dei dispositivi di controllo lampada (SYRA) [MINOSX].



Figura 2.8: Tipologia di ANDROS (tratto da [MINOSX])

Capitolo 3

Parte Implementativa

Dopo aver introdotto concettualmente nei capitoli precedenti i Dati Spaziali (capitolo 1) e le Reti di Sensori (capitolo 2), vengono presi in esame alcuni database che contengono appunto questi tipi di dati riferiti a sensori georeferenziati. Per svolgere queste attività sono stati utilizzati i seguenti strumenti: *DBMS MySQL* (versione: 5.1.63), il più popolare RDBMS (*Relational database management system*) *open-source* prodotto dalla Oracle Corporation e composto da un client con interfaccia a riga di comando e un server [MYSQL]; *phpMyAdmin* (versione: 3.3.7), un'applicazione PHP che consente di amministrare in modo semplificato database MySQL tramite un qualsiasi browser [PMAD]. Questi strumenti sono stati messi a disposizione dall'azienda *UMPI Elettronica S.r.l.* e pertanto verranno utilizzati dei nomi fittizi e riferimenti non specifici per rispettare le politiche di riservatezza aziendale.

In questo progetto verranno utilizzate le seguenti basi di dati: il database "DB_1", che contiene le informazioni relative ai sensori a livello armadio (ANDROS) e a livello lampada (SYRA) del sistema di controllo Minos System (descritto precedentemente nella sezione 2.5); il database "DB_2", che contiene le informazioni relative a sensori per il controllo del traffico (WTC¹);

¹Wireless Traffic Controller: sensori che hanno il compito di controllare il passaggio di automobili, di analizzare la velocità e l'intensità del traffico.

infine il database “DB_MAPPING” che, come verrà spiegato in seguito, è stato utilizzato per “mappare” i tre sottosistemi di sensori (sottosistema 1 per WTC, sottosistema 2 per ANDROS e sottosistema 3 per SYRA).

Il progetto viene articolato in quattro fasi: nella prima si effettuano modifiche strutturali in alcune tabelle dei database in modo da agevolare le fasi successive; nella seconda vengono inserite delle informazioni nel DB_1 relative a sensori ANDROS e SYRA ipoteticamente situati in alcune città europee; nella terza viene creata una tabella, detta *Map_tab*, nel database DB_MAPPING che contiene le informazioni basilari per mappare i tre sottosistemi; nella quarta vengono aggiunti dei campi a *Map_tab* per consentire interrogazioni di dettaglio su un singolo oggetto.

3.1 Prima Fase: Modifiche Strutturali

Per rendere il lavoro successivo di mapping più agevole, si apportano dei cambiamenti strutturali alla tabella del DB_1 che contiene le informazioni anagrafiche di ANDROS e SYRA (quindi dei sottosistemi 2 e 3). In questo contesto questa tabella viene detta *Minos_tab* e gli elementi vengono divisi gerarchicamente (Fig. 2.6) in base al campo CATEGORIA: 0 per TUTTI, 1 per CITTÀ, 2 per ANDROS e 3 per SYRA. Si procede illustrando le modifiche effettuate su *Minos_tab*. Per prima cosa viene inserito un campo in formato *integer*, detto ELEMENTO e popolato in maniera incrementale, come chiave primaria che identifica ogni elemento del *Minos System*; questo per esemplificare le relazioni nella tabella visto che il precedente campo della chiave primaria, detto EL_SRT, era una stringa di caratteri e numeri di lunghezza fissa, generata dal sistema. Successivamente, sempre per facilitare il lavoro sulla tabella, viene aggiunto un campo in formato *integer*, detto GENITORE, che indica l'identificativo di ogni dispositivo a livello gerarchico superiore (SYRA → ANDROS, ANDROS → CITTÀ); come in precedenza l'intero sostituisce il campo con la stringa, detto GEN_SRT. Di seguito è indicata la query utilizzata per popolare il campo GENITORE sfruttando il

join fra i campi EL_SRT e GEN_SRT.

```
UPDATE Minos_tab M1
INNER JOIN Minos_tab M2 on M1.GEN_SRT = M2.EL_SRT
SET M1.GENITORE = M2.ELEMENTO
```

Come risultato vengono aggiornate tutte righe.

Per testare la correttezza del passaggio precedente si utilizza la seguente query di verifica settando GENITORE = 62; vengono restituiti i sensori SYRA (CATEGORIA = 3) collegate al dispositivo armadio che chiameremo ANDROS_62. Una parte del risultato è illustrato in Figura 3.1.

```
SELECT ELEMENTO, GENITORE, CATEGORIA
FROM Minos_tab
WHERE GENITORE = 62
```

58	62	3
59	62	3
60	62	3
63	62	3
256	62	3
257	62	3
258	62	3
259	62	3
260	62	3
261	62	3
262	62	3
263	62	3
264	62	3

Figura 3.1: Risultato della query di verifica popolamento campo GENITORE. La prima colonna è riferita all'identificativo del sensore (ELEMENTO), la seconda all'identificativo del GENITORE e la terza a CATEGORIA

Oltre ai campi ELEMENTO e GENITORE, si introducono i seguenti campi per esprimere le coordinate geografiche degli elementi presenti nella tabella *Minos_tab*: il primo, in formato *float*, indica la latitudine in gradi decimali dell'elemento e viene detto LAT; il secondo, sempre in formato *float*, indica la longitudine in gradi decimali dell'elemento e viene detto LONG. Questi due campi ci serviranno per georeferenziare i sensori ANDROS e SYRA come vedremo nella fase successiva.

3.2 Seconda Fase: Inserimento coordinate per ANDROS e SYRA

In questa fase si procede con la scelta di alcuni dispositivi armadio ANDROS (CATEGORIA = 2) nella tabella *Minos_tab* e si associa ad essi coordinate vicine, ma non coincidenti, a centri di alcune città europee (Fig 3.3); in seguito si aggiornano anche quelle delle SYRA (CATEGORIA = 3) collegate al dispositivo ANDROS. Le città vengono scelte dalla tabella *Mappa*, sempre del database DB_1, in cui sono riportate le rispettive coordinate geografiche in gradi decimali.

Successivamente si popolano i campi LAT e LONG dei sensori Syra associati. Con la seguente query di verifica si visualizzano i sensori SYRA (CATEGORIA = 3), con relativi dati spaziali (LAT e LONG), associate all'ANDROS_62 (GENITORE = 62). Una parte del risultato è illustrato in Figura 3.2.

```
SELECT ELEMENTO, GENITORE, CATEGORIA, LAT, LONG
FROM Minos_tab
WHERE GENITORE = 62
```

3.3 Terza Fase: Mapping dei Sottosistemi

Con il termine *mapping* si intende il processo che identifica la localizzazione delle tabelle e dei file all'interno di un database. In questa fase del

58	62	3	52.536545	13.453872
59	62	3	52.557838	13.473213
60	62	3	52.526535	13.381212
63	62	3	52.542534	13.429812
256	62	3	52.552734	13.459012
257	62	3	52.552784	13.449212
258	62	3	52.553753	13.478692
259	62	3	52.533752	13.434222
260	62	3	52.571751	13.324222
261	62	3	52.513210	13.483222
262	62	3	52.502213	13.383222
263	62	3	52.542210	13.443233
264	62	3	52.598209	13.493763

Figura 3.2: Risultato della query di verifica dei sensori SYRA. La prima colonna è riferita all'identificativo del sensore (ELEMENTO), la seconda all'identificativo del GENITORE, la terza a CATEGORIA, la quarta alla latitudine (LAT) e la quinta alla longitudine (LONG) del sensore

progetto si crea una tabella, detta *Map_tab*, nel database DB_MAPPING che contiene le informazioni basilari per mappare i tre sottosistemi di sensori (1 per WTC, 2 per ANDROS e 3 per SYRA). Tutto questo perché le informazioni di interesse non si trovano nello stesso database ma rispettivamente nel DB_1 per ANDROS e SYRA e nel DB_2 per WTC. Di seguito sono illustrati alcuni campi appartenenti alla tabella *Map_tab*. In particolare questi campi contengono i nomi dei campi di riferimento tra cui: il campo che identifica il dispositivo (come ELEMENTO per ANDROS e SYRA); il campo che definisce la categoria del sensore (come CATEGORIA per ANDROS e SYRA); il campo che definisce l'identificativo di ogni dispositivo genitore a livello gerarchico superiore (come GENITORE per ANDROS e SYRA); il campo che definisce la longitudine del sensore (come LONG per ANDROS e SYRA); il campo che definisce la latitudine del sensore (come LAT per ANDROS e SYRA); il nome del database del sottosistema (come DB_1 per ANDROS e SYRA e DB_2 per WTC); il nome della tabella del database da cui si ricavano i campi per il mapping (come *Minos_tab* per

	LATITUDINE	LONGITUDINE	UBICAZIONE
Berlin	52.32	13.25	Germania
ANDROS_62	52.543211	13.423222	Senefelderstraße 21 10437 Berlin, Germania
Barcellona	41.25	2.1	Spagna
ANDROS_274	41.392135	2.17186	Gran Via Corts Catalanes 08010 Barcelona, Spagna
Le Mans	48.01	0.11	Francia
ANDROS_290	48.006678	0.199749	6 Avenue Pierre Mendès France 72100 Le Mans, Francia
Milano	45.28	9.12	Italia
ANDROS_313	45.464616	9.185514	Via Armorari 20123 Milano, Italia

Figura 3.3: ANDROS e Città scelti con relative coordinate e ubicazione

ANDROS e SYRA). Infine viene inserito un campo di tipo *longtext*, detto DESCRIZIONE, che contiene una spiegazione dei campi precedentemente inseriti. In particolare per DESCRIZIONE si utilizza la seguente sintassi: *tabella#campo#spiegazione#-#tabella#campo#spiegazione#...*

3.4 Quarta Fase: Interrogazioni di Dettaglio su Singolo Sensore

Queste interrogazioni restituiscono tutte le informazioni interessanti di un determinato oggetto richiesto; ma questi dati non si trovano necessariamente nella stessa tabella. Per cui saranno necessari diversi join fra le tabelle del database relativo al sottosistema dell'oggetto, in modo da restituire il risultato desiderato. Una soluzione possibile è quella di creare un campo di tipo *longtext*, che chiameremo JOIN_FIELD, nella tabella *Map.tab*. Questo

campo è inizializzato a NULL se non ci sono join, altrimenti viene inserita la seguente sintassi:

```
tabella1#tabella2#campo join 1#campo join 2#-#tabella1#tabella2#campo  
join 1#campo join 2#-#...
```

Dopo il join iniziale si effettuano join a catena in modo da ottenere i campi che contengono le informazioni interessanti per un singolo oggetto. Una spiegazione di tali campi viene inserita nel campo DESCRIZIONE rispettando la semantica precedentemente indicata.

Esempio di query di dettaglio:

```
SELECT *  
FROM $FROM_CLAUSE  
WHERE $WHERE_CLAUSE
```

Dove *\$FROM_CLAUSE* e *\$WHERE_CLAUSE* sono costruite dinamicamente in base ai dati estrapolati dalla tabella di mapping ed eventualmente effettuando il parsing del campo JOIN_FIELD.

Conclusioni

In questa tesi è stato affrontato uno studio teorico sulla georeferenziazione, sui Dati Spaziali e sulle Reti di Sensori, seguito da un lavoro pratico su un sottoinsieme dei database dell'azienda *UMPI Elettronica S.r.l* e concluso con il mapping dei vari sottosistemi. Durante le attività pratiche sulle basi di dati sono sorti alcuni problemi. In primo luogo è stato opportuno modificare la struttura di alcune tabelle, introducendo campi e modificando la chiave primaria, per facilitare il lavoro successivo di mapping. Inoltre è stato risolto il problema delle interrogazioni di dettaglio garantendo il recupero di informazioni interessanti anche per un singolo sensore.

Il lavoro svolto è stato stimolante e istruttivo poiché mi ha permesso di approfondire e consolidare le mie conoscenze teoriche e di comprendere meglio le problematiche e le metodiche dell'amministrazione di basi di dati aziendali.

Alcuni sviluppi futuri potrebbero riguardare la progettazione di un applicativo mobile per Android o iOS: l'idea è quella di sviluppare un meccanismo che permetta di accedere ai dati archiviati dentro la tabella di mapping, *Map_tab*, grazie alla quale verranno poi prelevate le informazioni da visualizzare sullo strumento end-user. Tale meccanismo deve essere realizzato tenendo presente che le informazioni sorgente, prelevate dalle tabelle di origine, possono variare ed inoltre il formato delle coordinate archiviate in esse può essere differente dal formato di destinazione. Potrebbe essere implementato in due modalità differenti: mediante web service da invocare ogni qualvolta un utente effettua un login allo strumento end-user, permettendo

l'interoperabilità tra diverse applicazioni software su piattaforme hardware differenti; mediante stored procedure di MySQL, invocata da trigger, ovvero una procedura che viene eseguita in maniera automatica nella base di dati in coincidenza di un determinato evento. Indipendentemente dalla scelta implementativa, l'applicativo deve operare in maniera trasparente rispetto ai database e ai dispositivi gestiti; per questo si potrebbe cercare di rendere parametrici i nomi delle tabelle e dei campi utilizzati.

Bibliografia

- [1] Lotti Nevio, “Appunti di Gis”, Paper: GIS Georeferenziazione, 2004, <http://www.lvproject.com/images/GIS%20Georeferenziazione.pdf> (10/01/2013)
- [2] Sedita Silvia Rita, “Concetti sui Gis”, 2006, http://sid.decon.unipd.it/materiale3/bel_GIS.pdf (21/12/2012)
- [3] Gubiani Donatella, Montanari Angelo, “Introduzione ai Dati Spaziali e alle Basi di Dati Spaziali”, 2010, http://users.dimi.uniud.it/~donatella.gubiani/universita/corsi/bds/bds/slide1_10.pdf (12/01/2013)
- [4] Gubiani Donatella, Montanari Angelo, “Rappresentazione di oggetti spaziali”, Corso di Basi di Dati Spaziali, 2010, http://users.dimi.uniud.it/~donatella.gubiani/universita/corsi/bds/bds/slide2_10.pdf (12/01/2013)
- [5] Gubiani Donatella, Montanari Angelo, “Modello logico per dati spaziali”, Corso di Basi di Dati Spaziali, 2010, http://users.dimi.uniud.it/~donatella.gubiani/universita/corsi/bds/bds/slide3_10.pdf (12/01/2013)
- [6] Gubiani Donatella, Montanari Angelo, “Linguaggi: (Geo) Algebra relazionale”, Corso di Basi di Dati Spaziali, 2010, http://users.dimi.uniud.it/~donatella.gubiani/universita/corsi/bds/bds/slide4_08.pdf (12/01/2013)

-
- [7] Gubiani Donatella, Montanari Angelo, “Estensione spaziale di SQL (i sistemi PostgreSQL e Oracle Spatial)”, Corso di Basi di Dati Spaziali, 2010, http://users.dimi.uniud.it/~donatella.gubiani/universita/corsi/bds/bds/slide5_10.pdf (12/01/2013)
- [8] Gubiani Donatella, Montanari Angelo, “La progettazione fisica: gli indici spaziali - 1 ”, Corso di Basi di Dati Spaziali, 2010, http://users.dimi.uniud.it/~donatella.gubiani/universita/corsi/bds/bds/slide5_10.pdf (13/01/2013)
- [9] Gubiani Donatella, Montanari Angelo, “La progettazione fisica: gli indici spaziali - 2 ”, Corso di Basi di Dati Spaziali, 2010, http://users.dimi.uniud.it/~donatella.gubiani/universita/corsi/bds/bds/slide5_10.pdf (13/01/2013)
- [10] Rigaux P. , Scholl M. , Voiosard A. , *Spatial Databases (with application to GIS)*, Morgan Kaufmann, 2002.
- [11] Shekhar S. , Chawla S. , *Spatial Databases: A Tour*, Prentice Hall, 2002.
- [12] Lattanzi Emanuele *Strutture dati efficienti per la ricerca della similarità*, Corso di Sistemi Informativi Multimediali, 2005, <http://www.sti.uniurb.it/lattanzi/SIM/lezioni/Lez6.pdf> (13/01/2013)
- [13] Burrough P.A. , *Principles of geographical information systems for land resource assessment*, Clarendon Press, Oxford, U.K, pp. 194, 1986.
- [14] Arnaud A. , Masser I. , Salgè F. , Scholten H. , *GISDATA Research Programme*, European Science Foundation GISDATA Newsletter, n. I., 1993.
- [15] Cortellessa C. M. , *Breve introduzione al GIS*, Supplemento a Mondo Autocad n.5, 1995.

- [16] Basile Paolo, *I Sistemi di navigazione geografica come strumento di valorizzazione delle basi statistiche territoriali*, Sistemi Territoriali Pisa, 2002.
- [17] Neteler M. , Bowman M.H. , Landa M. , Metz M. , *GRASS GIS: A multi-purpose open source GIS*, Environmental Modelling I& Software, pp. 124-130, 2012.
- [18] Tobler W. , *A computer movie simulating urban growth in the Detroit region*, Economic Geography, 46(2): pp. 234-240, 1970.
- [19] Worboys M. F. , Duckham M. , *GIS- A Computing Perspective*, Second Edition, CRC Press, 2001.
- [20] Gerbino Cesare, “Geographic Information Systems: Principi, Strumenti e Architetture”, 2002, <http://www.cesaregerbino.it/Pdf/GIS-PrincipiStrumentiArchitetture.pdf> (10/01/2013)
- [21] Flammini A. , Sisinni E. , Ferrari P. , Depari A. , *Reti di sensori wireless, confronto tra le varie tecnologie*, Wireless Sensor Networks Laboratory, Università di Brescia, 2004, http://www.ing.unibs.it/~wsnlab/download/WSNlab_tutorial04_gen2.pdf (16/01/2013)
- [22] Flammini Alessandra, “Reti di sensori: generalità e tendenze”, Presentazione Seminario di Eccellenza “Italo Gorini”, 2005, <http://www.ing.unibs.it/~wsnlab/download/Flammini-gorini.pdf> (16/01/2013)
- [23] Sohraby K. , Minoli D. , Znati T. , *Wireless Sensor Networks: Technology, Protocols, and Applications*, WileyInterscience, 2007.
- [24] Nitaigour P. Mahalik, *Sensor Networks and Configuration: Fundamentals, Standards, Platforms, and Applications*, Springer, 2006.

SITOGRAFIA

- [GRASSORG] <http://grass.osgeo.org> - sito ufficiale di GRASS (27/12/2012)
- [GRASSIT] <http://grass-italia.como.polimi.it> - sito utenti italiani di GRASS (27/12/2012)
- [GVPONTIS] <http://www.gvpontis.gva.es> - sito ufficiale di gvSIG (27/12/2012)
- [SEXTANTE] <http://sextantegis.com> - sito ufficiale di Sextante (27/12/2012)
- [MYSQLM] <http://dev.mysql.com/doc/refman/5.6/en/spatial-extensions.html> - *MySQL 5.6 Manual*, sezione *12.18 Spatial Extensions* (8/01/2013)
- [GEOREF] <http://en.wikipedia.org/wiki/Georeference> (27/12/2012)
- [GIS] http://it.wikipedia.org/wiki/Geographical_Information_System (18/12/2012)
- [CAD] <http://it.wikipedia.org/wiki/CAD> (20/12/2012)
- [GNSS] <http://it.wikipedia.org/wiki/GNSS> (20/12/2012)
- [GPS] <http://it.wikipedia.org/wiki/GPS> (20/12/2012)
- [AGPS] <http://it.wikipedia.org/wiki/A-GPS> (20/12/2012)
- [GLONASS] <http://it.wikipedia.org/wiki/GLONASS> (20/12/2012)
- [GALILEOS] http://it.wikipedia.org/wiki/Sistema_di_posizionamento_Galileo (20/12/2012)
- [GEOC] <http://en.wikipedia.org/wiki/Geocoding> (12/01/2013)
- [RGEOC] http://en.wikipedia.org/wiki/Reverse_geocoding (12/01/2013)

- [QTREE] <http://it.wikipedia.org/wiki/Quadtree> (15/01/2013)
- [RTREE] <http://it.wikipedia.org/wiki/R-tree> (15/01/2013)
- [GEOLOC] <http://it.wikipedia.org/wiki/Geolocalizzazione> (12/01/2013)
- [WSN] http://it.wikipedia.org/wiki/Wireless_sensor_network (14/01/2013)
- [BAN] http://it.wikipedia.org/wiki/Wireless_Body_Area_Network
(20/01/2013)
- [PAN] http://it.wikipedia.org/wiki/Personal_Area_Network (20/01/2013)
- [LAN] http://it.wikipedia.org/wiki/Local_Area_Network (20/01/2013)
- [MAN] http://it.wikipedia.org/wiki/Metropolitan_Area_Network
(20/01/2013)
- [WAN] http://it.wikipedia.org/wiki/Wide_Area_Network (20/01/2013)
- [MINOSX] <http://www.minos-system.com> (24/01/2013)
- [MYSQL] <http://it.wikipedia.org/wiki/MySQL> (21/02/2013)
- [PMAD] <http://it.wikipedia.org/wiki/PhpMyAdmin> (21/02/2013)

NOTA: tutte le citazioni da testi in lingua straniera sono state effettuate traducendo o riassumendo nel modo semanticamente più fedele possibile. Per ogni url viene indicata la data di ultima consultazione.

Ringraziamenti

Si ringraziano il *Prof. Dario Maio* e l'*Ing. Luca Calderoni* per il tempo dedicato alla realizzazione di questa tesi. Si ringrazia inoltre la *UMPI Elettronica S.r.l.*, con particolare attenzione alla grande disponibilità dell'*Ing. Stefano Rovis*, per la collaborazione e gli strumenti messi a disposizione.