

UNIVERSITA' DEGLI STUDI DI PADOVA



FACOLTA' DI INGEGNERIA

Tesi di Laurea in Ingegneria Informatica

Generalizzazione cartografica della rete ferroviaria nel progetto CARGEN

Relatore: Prof. Massimo Rumor

Laureando: Daniele Ronzani

Matricola: 545269 / IF

Anno Accademico 2011 - 2012

INTRODUZIONE	4
LA CARTOGRAFIA	5
1.1 Cenni Storici	5
1.2 La Carta Geografica e i GIS	6
1.2.1 Le Carte Geografiche.....	6
1.2.2 I GIS.....	8
1.3 Il processo cartografico	9
1.3.1 Differenze tra processo cartografico tradizionale e digitale	10
LA GENERALIZZAZIONE	11
2.1 Definizione di generalizzazione	11
2.2 Il Processo di generalizzazione	13
2.3 La generalizzazione automatica	13
IL PROGETTO CARGEN	15
3.1 I modelli utilizzati nel progetto CARGEN	16
3.1.1 Carta tecnica regionale numerica	16
3.1.2 DBT	16
3.1.3 DB25	17
3.2 Le fasi del progetto CARGEN	17
3.2.1 Creazione e popolamento del DBT	17
3.2.2 Analisi dei modelli	18
3.2.3 Creazione del DB25	18
3.3 Strumenti e tecniche	18
3.3.1 JTS Topology Suite.....	19
3.3.2 OpenJump	19
LA GENERALIZZAZIONE DELLE FERROVIE	21
4.1 Dataset utilizzato per i test	22
4.2 Pulizia del grafo	24

4.2.1 KdTree	24
4.2.2 Snapping	27
4.2.3 Dissolve	27
4.2.4 Noding	27
4.2.5 Eliminazione linea	28
4.2.6 Collapse	28
4.3 Creazione e scelta delle coppie.....	29
4.3.1 Fase 1: Creazione delle linee ortogonali.....	29
4.3.2 Fase 2: Creazione delle coppie candidate.....	30
4.3.3 Fase 3: Scelta delle coppie.....	32
4.4 Creazione delle <i>mid-line</i>	34
4.5 Gestione delle connessioni.....	35
CONCLUSIONI.....	37
BIBLIOGRAFIA.....	38

INTRODUZIONE

La generalizzazione cartografica è un processo essenziale nell'ambito della cartografia. Il suo obiettivo è generare una mappa ad una scala minore, partendo da una mappa di scala maggiore. I problemi che nascono in questo processo non sono trascurabili, poiché le principali difficoltà derivano dallo spazio per contenere tutti gli elementi rilevanti della carta geografica; infatti questo diminuisce al diminuire della scala.

Solitamente la produzione di una carta a bassa scala, parte da informazioni geografiche derivanti da una mappa preesistente a scala superiore. Queste astrazioni della realtà, anni fa, erano svolti dai cartografi, che ridisegnavano a mano ogni dettaglio all'interno della cartografia.

Per ovviare ai lunghi tempi da cui derivava questo lavoro a mano, è nata l'idea di automatizzare questo processo: è da qui che parte il progetto di ricerca CARGEN.

Il progetto CARGEN ha lo scopo di sviluppare strumenti in grado di automatizzare la produzione di un database topografico in scala 1:25000 e 1:50000 partendo da carte in scala 1:5000.

Questa tesi si pone come obiettivo la realizzazione di un algoritmo in grado di generalizzare la rete ferroviaria nelle scale appena menzionate.

Come si può immaginare, la rappresentazione di tutto il reticolo ferroviario, nei suoi minimi dettagli, è impossibile alle scale 1:25000 o 1:50000; è quindi obbligatorio uno sfoltoimento dello stesso, in modo da rendere possibile una corretta rappresentazione, coerente con la scala, senza screditare l'informazione generale circa l'andamento di massima del tracciato.

Capitolo 1

La Cartografia

La cartografia è l'insieme delle tecniche necessarie alla realizzazione di carte geografiche, topografiche, geologiche e simili (Il Grande Dizionario Garzanti, 1987). Una carta geografica è una rappresentazione grafica che rappresenta un'astrazione della realtà (di una parte o di tutta la superficie terrestre) secondo delle convenzioni, dei vincoli imposti e delle funzionalità per cui questa carta è stata progettata.

La scala è un fattore fondamentale della cartografia: è definita come il rapporto tra le distanze rappresentate sulla scala stessa e le distanze reali. Le carte a scala grande contengono elementi più dettagliati, e presentano una grande quantità di informazione. Al contrario, le carte a scale minori, rappresentano un'area più vasta, ma poiché la densità di informazione è più ampia, occorre semplificarne gli elementi che la compongono, eliminandone alcuni o trasformarli.

1.1 Cenni Storici

Quello che spinse le popolazioni antiche a redigere mappe geografiche fu l'incontenibile desiderio di spostarsi. E' da qui che iniziamo a trovare pietre o legno sulle quali sono incisi segni che potrebbero raffigurare fiumi, scene di caccia, linee di confine, monti.

Alcuni dei più antichi documenti cartografici ancora esistenti sono scritti su tavolette d'argilla, come quella rinvenuta nel 1930 sulle rovine della località di Gar-Sur in Babilonia, risalente, secondo gli studiosi, all'epoca della dinastia di Sargon di Akkad, tra il 2300 e il 2500 a.c.

Stranamente, una delle più grandi potenze della storia, l'Egitto, non ha lasciato quasi nessun documento cartografico. Una delle poche carte redatte dagli egizi è quella che si trova ora al Museo Egizio di Torino, nella quale viene raffigurata la mappa schematica delle miniere d'oro della Nubia, creata, probabilmente, durante il regno di Ramsete IV.

Un grande passo in avanti si ottenne quando famosi pensatori, come Pitagora ed Aristotele introdussero l'ipotesi che la terra poteva essere sferica.

Con l'arrivo del Rinascimento, grazie alle scoperte di famosi esploratori (Colombo, Magellano, ...) e grazie anche alle innovazioni tecnologiche, la produzione di mappe raggiunse uno sviluppo mai visto prima e le carte prodotte erano molto simili a quelle in uso oggi.

Nel 1872, in Italia, venne fondato uno dei primi istituti cartografici, chiamato Istituto Topografico Militare, che si impegnò alla costruzione della prima Carta Topografica d'Italia (completata nel 1921) alla scala 1:100 000.

Al termine del XX secolo siamo presenti ad un'altra rivoluzione cartografica, dettata dalle prime apparizioni del GIS (Geographical Information System), un potente sistema informatico in grado di analizzare, visualizzare ed elaborare tutti gli elementi geografici di una carta.

1.2 La Carta Geografica e i GIS

La carta geografica è una rappresentazione figurativa convenzionale, generalmente piana, di una parte o di tutta la superficie terrestre considerata nella varietà di aspetti e fenomeni, concreti o astratti, su di essa localizzabili. Tale rappresentazione è necessariamente ridotta secondo una data proporzione (scala); inoltre è simbolica nel senso che i dati degli elementi del terreno e tutti i fenomeni sono riprodotti nella carta mediante segni appropriati essenzialmente convenzionali, ordinati in una legenda da allegare alla carta.

La produzione di una carta geografica è un processo assai complicato: è essenziale tenere conto di una moltitudine di fattori vincolanti, come la scala, la natura della superficie (che essendo curva deve essere disegnata su una superficie piana) e l'eterogeneità degli elementi; a causa di tali difficoltà gli oggetti che compongono la realtà subiscono trasformazioni di vario tipo, rendendo il risultato di queste operazioni una "approssimazione" della realtà.

1.2.1 Le Carte Geografiche

Le carte geografiche possono essere classificate in base a due categorie.

Se cataloghiamo le carte in base alla loro scala otteniamo quattro differenti gruppi:

- Carte geografiche (in senso stretto): hanno una scala minore di 1:1.000.000 e servono per rappresentare continenti o tutta la superficie terrestre.

- Carte corografiche: hanno una scala compresa tra 1:1.000.000 e 1:100.000 e rappresentano zone estese.
- Carte topografiche: hanno una scala compresa tra 1:100.000 e 1:10.000 e raffigurano grandi aree e riescono ad evidenziare alcuni particolari, come case o ponti.
- Piante o mappe: hanno una scala maggiore di 1:10.000 e possono rappresentare planimetrie e zone rurali: queste carte offrono particolari molto dettagliati.

Se invece suddividiamo le carte in base al contenuto otteniamo:

- Carte topografiche: rappresentano fisicamente il terreno dando il maggior numero possibile di informazioni a riguardo. In queste carte troviamo laghi, mari, fiumi, edifici, strade, ponti, etc.
- Carte tematiche: servono per rappresentare temi specifici per dimostrare qualitativamente e quantitativamente fenomeni concreti o concezioni astratte.

Un concetto molto importante, e forse il concetto da tener sempre presente quando si parla di cartografia, è la scala.

La scala è definita come il rapporto tra l'unità di lunghezza misurata sulla carta e il valore della stessa lunghezza misurata nella realtà: in altre parole 1 cm di lunghezza in una mappa a scala 1:10.000 corrisponde a 100 metri nella realtà ($0,01 \text{ metri} / 100 \text{ metri} = 0,0001 = 1 / 10.000$).

La scala consente di sapere quali e quante informazioni rappresentare in una carta, perché al cambiare di essa, cambia anche il più piccolo dettaglio percepibile. E' da tenere presente che una linea di spessore 0,5 mm in una carta in scala 1:25.000 copre una fascia di 12 metri: proprio per questo ogni carta viene parametrizzata da due parametri che dipendono dalla scala:

- Grado di risoluzione: è la dimensione del più piccolo particolare rappresentabile ed è dato dal minimo spessore del tratto grafico con cui la carta viene disegnata; per convenzione è uguale a 0,2 mm.
- Errore massimo di posizionamento: rappresenta il diametro del cerchio all'interno della quale un punto è sicuramente contenuto; questo valore è solitamente di 0,5 mm.

1.2.2 I GIS

Fino a qualche anno fa le carte geografiche venivano rappresentate unicamente in modo cartaceo. Ora con l'avvento dello sviluppo tecnologico e quindi dei computer, assistiamo alla nascita dei GIS e quindi ad una vera e propria transizione da un approccio statico ad un approccio dinamico.

Un GIS (Geographical Information System) è un sistema informativo che raccoglie, analizza, elabora e visualizza le informazioni geografiche (spaziali). Grazie a questo sistema, gli oggetti che compongono la mappa diventano entità dinamiche, con propri attributi, proprie caratteristiche e una propria geometria.

I campi di applicazione dei GIS sono ormai molteplici: pianificazione territoriale, economia, difesa, gestione forestale, istruzione, geologia, telecomunicazioni, gestione di infrastrutture, trasporti, sono solo alcune delle tematiche nella quale si può collocare un GIS.

Le entità, gli elementi base del GIS, sono poste in un database spaziale chiamato GeoDB, che ha la funzione di memorizzare, interrogare e manipolare dati spaziali.

Un GIS è composto da risorse (umane, strutturali, dati) e procedure (acquisizione, archiviazione, elaborazione, trasmissione, ...).

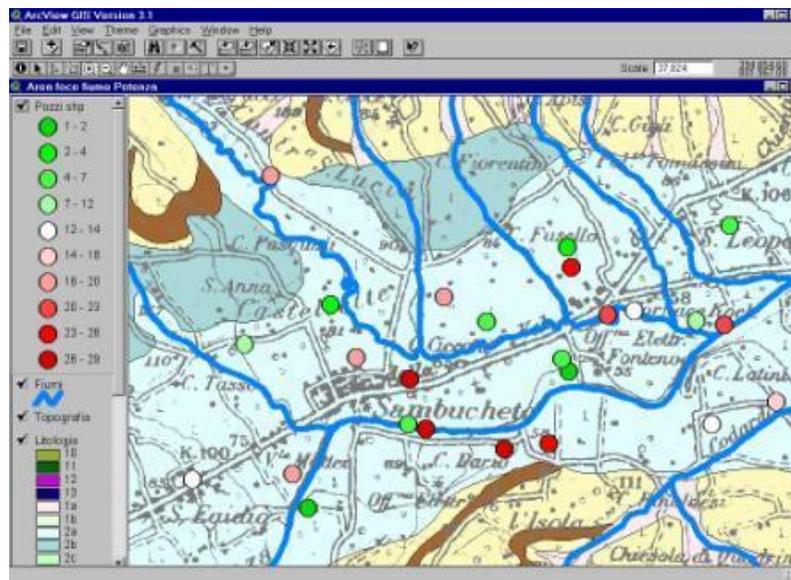


Figura 1: Interfaccia di un GIS

L'elemento più importante che costituisce un GIS sono i dati, che devono essere costantemente aggiornati.

Ogni dato è chiamato **feature** e contiene una componente spaziale (posizione, geometria, topologia) e non spaziale (descrizioni, attributi, metadati).

L'acquisizione dei dati prende le fonti da mappe cartacee o file contenenti le coordinate degli elementi: questa fase di raccolta dei dati avviene attraverso un processo di astrazione e generalizzazione.

Esistono due modelli di dati spaziali: **raster** e **vettoriale**.

I modelli raster sono immagini in cui ogni pixel contiene informazioni quali il colore, altitudine, ecc; in altre parole le immagini di questo tipo sono costituite da celle e la dimensione dei file che utilizzano questo modello son di grandi dimensioni.

I modelli vettoriali invece si comportano come dei contenitori spaziali di oggetti; questi oggetti sono le primitive geometriche che hanno i propri attributi e proprietà. Le primitive sono: punto (coordinate x,y), linea (sequenza di punti), poligono (sequenza di punti i quali il primo e l'ultimo sono lo stesso punto).

E' facile intuire che il modello vettoriale è nettamente superiore al modello raster in quanto la manipolazione, elaborazione e interrogazione delle primitive geometriche risulta molto più semplice della gestione dei singoli pixel. Inoltre l'utilizzo del modello vettoriale offre una facile integrazione in database relazionali.

Molti vantaggi sono legati all'uso dei GIS, tra cui la breve tempistica di produzione, elaborazione e archiviazione di mappe, una più facile gestione del processo cartografico, la presenza di una certa automazione, minori costi, un maggior risparmio di spazio e risorse e, non di meno, la possibilità di aggiornamento e condivisione dei dati.

1.3 Il processo cartografico

Quando si guarda una carta geografica, si sta guardando un prodotto la cui realizzazione ha impegnato molte persone e impiegato molto tempo. In effetti la produzione di una mappa vede una serie di fasi, che compongono il cosiddetto Processo Cartografico, che ospita tutte quelle operazioni necessarie alla creazione della mappa stessa. Le fasi sono le seguenti:

1. Definizione
2. Analisi
3. Raccolta dei dati
4. Costruzione della mappa
5. Collaudo

1. Nella definizione si decide la finalità dell'utilizzo della mappa, quindi cosa si dovrà rappresentare e come.
2. Nell'analisi si decidono le caratteristiche della carta: quale superficie rappresentare, tipo di rappresentazione, scala, tipologia di elementi, o decidere quali aspetti della realtà rappresentare (cartografia meristica od olistica)
3. La raccolta dei dati prevede due opzioni: ricavare i dati tramite campagne di acquisizione sul territorio (rilevazione) oppure ricavarli direttamente da altre carte esistenti (derivazione). Nel corso del tempo la rilevazione si è evoluta, partendo da strumenti come la tavoletta pretoriana e il teodolite (da terra), arrivando oggi al GPS e la fotogrammetria.
4. La costruzione della mappa produce la carta tenendo conto di alcuni vincoli, oltre a quelli specificati in fase di definizione: leggibilità e usabilità. In questa fase il cartografo svolge la generalizzazione, in cui viene scelta la forma, dimensione e posizione di ogni oggetto sulla carta finale, al fine di dargli una rappresentazione astratta e il più possibile efficace e informativa.
5. Il collaudo consiste di una serie di test sulla carta per validarne la consistenza e correttezza. Inoltre viene attivato un processo di affinamento

1.3.1 Differenze tra processo cartografico tradizionale e digitale

Il processo cartografico differisce di alcuni punti quando si parla di cartografia tradizionale e cartografia digitale. Nel processo tradizionale la costruzione della mappa è quasi sempre definita da un processo lineare e come output restituisce un prodotto finale ben definito (la carta). Al contrario, nel processo cartografico digitale la costruzione della mappa avviene attraverso un processo iterativo, con continue piccole modifiche ad un database già esistente, con una conseguente assenza di un prodotto finale ben preciso.

Questi due tipi di processi si chiamano rispettivamente *map-oriented* e *gis-oriented*.

Per esempio, le prime fasi del processo, quelle di definizione e analisi, definiscono una mappa che, nel processo *map-oriented* producono una tradizionale mappa cartacea, mentre, quando si parla di *gis-oriented*, il prodotto diventa un GeoDB.

Capitolo 2

La Generalizzazione

Molto spesso ci si trova a dover definire una carta geografica classificando secondo un ordine ben preciso alcuni elementi della realtà. In altre parole l'informazione che la mappa deve trasmettere all'utente finale deve essere chiara e precisa, riportando tutti gli elementi che sono rilevanti allo scopo per cui la carta è stata prodotta, tralasciando tutti quegli oggetti che non sono importanti o che, per qualche motivo, non possono essere riprodotti. E' qui che entra in gioco la generalizzazione cartografica.

2.1 Definizione di generalizzazione

La generalizzazione cartografica è un processo chiave nella fase di produzione di una carta derivata, perché avviene una ridefinizione, spostamento e semplificazione di tutti gli elementi coinvolti.

bisogna rappresentare elementi più importanti e non tutto, per avere una chiara leggibilità questa semplificazione deve rispettare i vincoli imposti di rappresentazione.

I fattori che devono essere tenuti in considerazione dal processo di generalizzazione sono molteplici; McMaster e Shea (1992) ne individuano sei:

1. Riduzione della complessità: il passaggio ad una scala inferiore porta naturalmente diversi oggetti ad entrare in conflitto per lo stesso spazio. Ridurre il numero di questi oggetti consente di far risaltare gli elementi più importanti e di mantenere la mappa leggibile.
2. Mantenimento dell'accuratezza spaziale: tanto più la scala è alta e tanto più l'accuratezza spaziale deve essere rispettata, limitando al minimo l'errore dovuto alla diversa posizione degli oggetti nella mappa rispetto alla realtà.

3. Mantenimento dell'accuratezza degli attributi: in particolare nelle mappe tematiche, l'obiettivo è quello di minimizzare le alterazioni non intenzionali degli attributi delle feature.
4. Mantenimento della qualità estetica: dai colori utilizzati alla simbologia, dal bilanciamento allo stile tipografico, vari fattori influenzano l'estetica complessiva di una mappa; l'arte del cartografo sta nel mixare opportunamente questi fattori in modo da ottenere una mappa esteticamente bella.
5. Mantenimento di una logica gerarchica: elementi di una stessa categoria, ma di importanza o dimensioni differenti, devono essere differenziati, in accordo con lo scopo della mappa; ad esempio, una città grande deve risultare molto più prominente di una città piccola, mostrando una maggiore densità del grafo stradale e dell'edificato.
6. Coerente applicazione delle regole di generalizzazione: allo scopo di ottenere una generalizzazione imparziale e coerente, il cartografo deve determinare esattamente quali algoritmi applicare e in quale ordine e i parametri di input necessari per ottenere il risultato voluto quando si opera ad una data scala.

Quando si crea una mappa derivata, cioè partendo da dati cartografici preesistenti, si parla di *map generalization* (generalizzazione cartografica), mentre se si disegna una mappa rilevata, il processo diventa *map compilation*.

Ciò nonostante, non esiste una reale distinzione tra questi due processi, in quanto le problematiche sono pressoché uguali: entrambi devono sottostare alle stesse specifiche della carta e agli stessi parametri.

La differenza sta invece nel fatto che, mentre quando si produce una carta rilevata si crea un modello della realtà partendo dalla realtà stessa, quando si produce una carta derivata si parte invece da un altro modello della realtà, ottenendo sempre un modello della realtà.

2.2 Il Processo di generalizzazione

Il processo di generalizzazione cartografica è solitamente molto lento e costoso ed incide molto su tempi e spese per la produzione della mappa ed è articolato in due fasi successive: la generalizzazione del modello e la generalizzazione cartografica.

La generalizzazione del modello ha l'obiettivo di creare un modello astratto che rappresenti la realtà di interesse. La scelta di quali aspetti riportare sulla mappa e quali eliminare è strettamente legata allo scopo per il quale la mappa è realizzata.

Quindi si procede ad una classificazione dei vari elementi per ottenere un primo livello di astrazione. Nell'ambito della cartografia digitale, la generalizzazione del modello corrisponde alla definizione di un GeoDatabase.

La generalizzazione cartografica consiste nella ridefinizione delle geometrie allo scopo di fornire una rappresentazione grafica del modello dei dati in precedenza creato. Per questa operazione non esiste uno schema rigido e nelle scelte operate in funzione della scala e del contesto il cartografo si deve affidare all'interpretazione e all'esperienza personale. In questa fase è introdotta un'ulteriore astrazione della realtà per l'aggiunta di approssimazioni al fine di migliorare la leggibilità della carta.

2.3 La generalizzazione automatica

Con l'arrivo dell'informatica si è arrivati ad un punto nella quale la maggior parte dei processi e attività, che erano svolte in passato manualmente, vengono svolte automaticamente dai computer.

Un passo di questo tipo molto importante trova luogo nell'ambito della generalizzazione cartografica.

In passato la generalizzazione manuale era un processo molto lungo e faticoso; con l'arrivo della tecnologia si è riusciti a ridurre considerevolmente i tempi e i costi.



Figura 2: Alcuni tipi di carte a scala diversa. Da *Atalante Geografico Illustrato* DeAgostini

Capitolo 3

Il progetto CARGEN

Il progetto di ricerca CARGEN (*CARtographic GENeralization*) ha come obiettivo lo sviluppo di un progetto di generalizzazione delle carte regionali tecniche (CTR) per produrre una base di dati con il modello 25DB dell'Istituto Geografico Militare (IGM).

Inizialmente, in Italia, la prima produzione dell' IGM fu la "Nuova Carta Topografica d'Italia" in scala

1:100 000.

Successivamente vennero prodotte diverse carte a diversa scala:

- Scala 1:25 000: sono state prodotte le serie 25V, 25 e 25DB;
- Scala 1:50 000: sono state prodotte le serie 50 e 50/L;
- Scala 1:100 000: sono state prodotte le serie 100V e 100L.

L'unica serie ad essere stata completata è la 25V (vecchio taglio), aggiornata al 1960. La produzione della serie 25 è stata interrotta per ospitare la nuova serie con supporto digitale, la 25DB.

Oltre all'IGM, che si occupa su cartografie con scala inferiore alla 1:25 000, vi sono altri enti che si occupano di carte a scale maggiori. La creazione di queste carte è affidata alle Regioni che si occupano in modo autonomo la definizione di tali carte, che offrono molti particolari e un aggiornamento frequente.

In Veneto esistono due serie di Carte Tecniche Regionali: una in scala 1:10 000 e una in scala 1:5 000.

Negli ultimi anni anche gli uffici regionali hanno adottato la digitalizzazione cartografica e questo evento apre un nuovo scenario nella cartografia italiana: adottare il metodo della derivazione per produrre la nuova serie 25DB a partire dalle carte regionali, in modo da limitare le rilevazioni dei dati sul territorio e ottenere quindi una accelerazione dei tempi di produzione. Questo ci conduce allo scopo del progetto CARGEN.

3.1 I modelli utilizzati nel progetto CARGEN

Nato da una convenzione tra la Regione Veneto e il Dipartimento di Ingegneria dell'Informazione dell'Università di Padova, il progetto CARGEN ha l'obiettivo di sperimentare delle metodologie informatiche per la derivazione del database in scala 1:25000 dell'IGM a partire dal database in scala 1:5000 della Regione Veneto.

Questo progetto è uno dei primi in Italia che tratta il tema della generalizzazione automatica in modo approfondito, inserendosi tra le principali iniziative già avviate in altri paesi (Francia, Spagna, Inghilterra).

Vengono descritti ora i principali tre modelli utilizzati nel progetto:

3.1.1 Carta tecnica regionale numerica

Chiamata anche CTRN, questa carta viene prodotta dalla Regione Veneto, è in formato vettoriale e la sua fonte di dati è il rilievo fotogrammetrico. Viene aggiornata di frequente.

Le scale di rappresentazione sono 1:5000 e 1:10000 (quest'ultima utilizzata per le zone scarsamente urbanizzate). L'organizzazione dei contenuti presenta Livelli e Codici: i Livelli sono una prima classe di suddivisione degli elementi, a loro volta divisi in Codici, che identificano particolari caratteristiche degli elementi.

In totale vi sono:

- 16 Livelli Principali
- 12 Livelli di servizio
- 6 Livelli Funzionali

Lo svantaggio risiede nel fatto che essendo realizzata da tecniche CAD, non contiene informazioni relative alla topologia e pertanto è obbligatorio procedere con fasi di controllo e pulizia dati.

3.1.2 DBT

Questo moderno modello dati è caratterizzato da una divisione dell'informazione territoriale in informazione base e strati tematici, nonché la gestione delle geometrie tridimensionali. Ogni oggetto spaziale viene rappresentato da una *feature* e gli elementi vengono suddivisi per livelli informativi a strati: vi è quindi una forte componente gerarchica.

3.1.3 DB25

Il DB25 è un *geodatabase* vettoriale di tipo relazionale (RDB) con geometria a struttura topologica.

Sono molte le fonti dalla quale proviene: stereorestituzione, derivazione da CTRN, digitalizzazione da cartografia preesistente. Il contenuto informativo è suddiviso in 10 strati tematici: trasporti, idrografia, limiti amministrativi e naturali, insediamenti, vegetazione, geomorfologia, industrie, servizi, altimetria e qualità dati. Gli oggetti sono presentati sottoforma di *feature*, suddivise in oggetti identificati dall'attributo LAB (*Label*). La codifica e la definizione segue lo standard DIGEST (*Digital Geographic Exchange Standard*). Le coordinate sono di tipo geografico, in gradi sessa decimale.

3.2 Le fasi del progetto CARGEN

Il piano di lavorazione di questo progetto è suddiviso in tre parti:

1. La creazione e popolamento del DBT
2. L'analisi delle differenze tra i modelli del DBT e DB25
3. La creazione del prototipo di DB25

Le prime due fasi possono essere eseguite separatamente o anche in parallelo, mentre la terza deve essere eseguita dopo che le prime due sono terminate.

3.2.1 Creazione e popolamento del DBT

La creazione del DBT è un passo importante per procedere poi al popolamento del DB25. Questa prima fase è divisa in due parti: la prima riguarda la costruzione delle tabelle del db, tenendo conto delle specifiche del DBT e creando le tabelle nel DBMS *Oracle Spatial 10g*. La seconda parte completa il tutto con il popolamento dei dati, che vengono opportunamente convertiti dal modello CTRN a quello del DBT. Tale conversione ha richiesto un non banale intervento, incentrato principalmente sulla trasformazione delle geometrie delle varie *feature* delle CTRN. Questo processo è stato svolto tramite algoritmi realizzati in Java.

Parallelamente a queste due fasi, è stato anche svolto un processo di controllo topologico per verificare la correttezza e la coerenza semantica dei dati immessi. Queste due fasi parallele

(creazione/popoloamento e controllo) hanno favorito la correzione di molti errori presenti, oltre alla garanzia di una migliore qualità dei dati generati.

3.2.2 Analisi dei modelli

L'analisi dei modelli è un'operazione necessaria al fine di premettere la definizione di un metodo di migrazione dei dati tra i vari differenti modelli (CTRΝ, DBT e DB25) che sono coinvolti nel progetto.

Quest'analisi analizza e prevede eventuali errori e problemi che possono verificarsi. Vengono analizzate una per una le *feature* del modello DB25 trovando i corrispondenti elementi nei modelli DBT e CTRΝ, al fine di produrre un documento riassuntivo di tutte le problematiche sorte e le indicazioni da seguire per procedere ad una corretta conversione. Queste analisi hanno proposto la necessità di porre delle modifiche a tutti e tre i modelli di dati, in particolare per il DB25.

3.2.3 Creazione del DB25

L'ultima fase del progetto CARGEN è la creazione del DB25, quindi la realizzazione delle tabelle delle *feature class* nel DBMS Oracle e il successivo popoloamento.

Una delle maggiori difficoltà nel popoloamento dei dati del DBT nel DB25 riguarda le geometrie delle *feature* che, nel DBT e nel CTRΝ sono in scala 1:5000, mentre il DB25 deve contenere dati in scala 1:25000. E' quindi necessaria una fase di derivazione dei dati, tradotta in un intervento di generalizzazione cartografica.

Questa è la parte più delicata e interessante del progetto, poiché viene anche affiancata dalla ricerca per automatizzare questa fase.

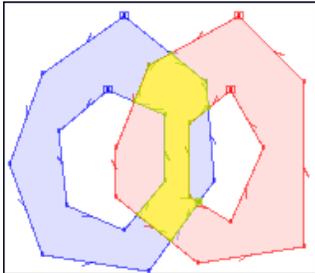
3.3 Strumenti e tecniche

All'inizio del progetto CARGEN era stato usato il modello client/server. La parte server era costituita da una macchina con il DBMS *Oracle Spatial 10g*, in cui venivano memorizzati e mantenuti i dati. Nel lato client erano installati i software *Dynamo/Dynagen* e *Geomedia Professional* (entrambi della Intergraph).

Quest'ultimo veniva utilizzato principalmente come strumento di accesso ai dati e come tramite tra il software *Dynamo/Dynagen* e Oracle, mentre gli strumenti *Dynamo/Dynagen* venivano usati per la generalizzazione cartografica.

Oggi questi strumenti sono stati in parte abbandonati, lasciando spazio a metodi e strumenti che offrono vantaggi prestazionali, temporali e che rendono più facile lo sviluppo di nuovi algoritmi.

3.3.1 JTS Topology Suite



La *JTS Topology Suite* è una API scritta interamente in Java e implementa un insieme di operazioni sui dati spaziali utilizzando un modello esplicito di precisione e algoritmi geometrici. Le tre geometrie fondamentali definite in questa libreria sono **Point**, **LineString** e **Polygon** (Punto, Linea, Poligono).

Le principali funzioni geometriche utilizzate nell'ambito di questa tesi sono:

- Buffer
- Albero KdTree
- Funzione di creazione LineString
- Involuppo delle geometrie
- Utilizzo dei LineSegment e calcolo angoli tra essi

3.3.2 OpenJump



OpenJump è un GIS open source scritto interamente in Java. È sviluppato e mantenuto da un gruppo di volontari provenienti da tutto il mondo. Questo software permette di caricare, gestire, modificare e interrogare dati spaziali e si basa sulla libreria JTS.

La sua caratteristica principale, e quella che ha offerto semplicità di sviluppo nel progetto, è la sua modularità, che permette di estendere le sue funzionalità mediante lo sviluppo di plug-in. In questo modo si può testare il proprio codice realizzando uno o più plug-in, integrandoli poi in OpenJump, al fine di visionare direttamente il risultato della propria applicazione.

Lo sviluppo di questi plug-in e il loro successivo test in OpenJump rende veloce ed efficace la creazione di algoritmi che manipolano geometrie e garantiscono una risposta grafica immediata per le fasi di testing e debug.

OpenJump lavora utilizzando i *layer*, livelli che si comportano come dei contenitori. Ogni contenitore ha al suo interno delle feature relative ad uno schema dati (*FeatureSchema*),

che definisce il numero e la tipologia di attributi che deve assumere le feature. In altre parole ogni livello si presenta come una tabella, alla quale si possono inviare delle query spaziali e non, al fine di recuperare tutte o alcune feature al suo interno.

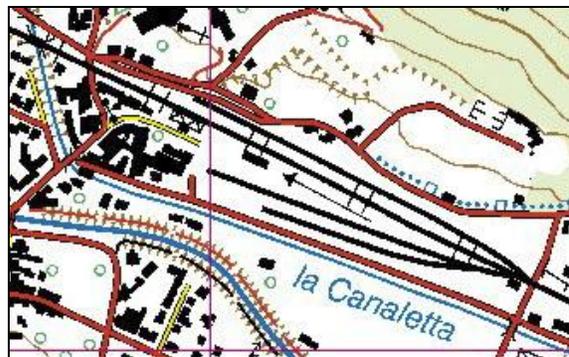
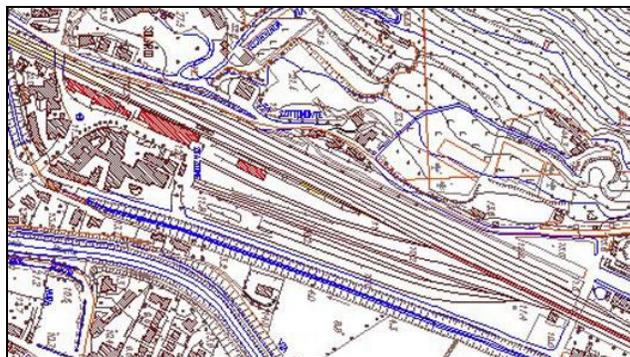
Capitolo 4

La Generalizzazione delle ferrovie

La generalizzazione cartografica delle linee ferroviarie è un aspetto della generalizzazione che apparentemente risulta di semplice risoluzione. In realtà, tralasciando le singole linee che collegano le diverse stazioni, è proprio in prossimità di queste ultime che nascono i problemi.

Lo scopo di questo elaborato è generalizzare una rete ferroviaria generica in modo automatico, ottenendo quindi una rappresentazione più semplice, in accordo con le specifiche imposte dall'IGM. Trasformando la rete ad una scala più piccola, ma senza generalizzare, si otterrebbero linee troppo vicine, sovrapposte e del tutto incomprensibili.

Grazie alla generalizzazione, invece, si ottengono dati più leggibili, linee più spaziate e quindi un tracciato pulito e semplice. Le due figure sottostanti mostrano un esempio di generalizzazione (*grafo ferroviario della stazione di Monselice - a scala 1:5000 e a scala 1:25000 rispettivamente*).



Il processo di generalizzazione è stato progettato dividendolo in quattro fasi principali:

1. Pulizia del grafo ferroviario;
2. Creazione e scelta delle coppie;
3. Creazione delle linee mediane;
4. Gestione delle connessioni.

La prima fase esegue una pulizia approfondita del tracciato, che consiste in un controllo e una successiva correzione delle connessioni e nell'eliminazione e collasso delle linee superflue.

Ovviamente l'algoritmo deve tener conto delle connessioni tra le linee, in modo da evitare di cancellare tratti che, nonostante abbiano una lunghezza minore, siano vitali per la continuità del tracciato.

La seconda e terza fase perseguono il vero e proprio obiettivo di questa tesi, utilizzando un approccio già studiato in un precedente lavoro di tesi nel progetto CARGEN. L'idea nasce dalla **scelta di coppie di linee (binari)**, che vengono coese in linee mediane, in modo da semplificarne via via il tracciato da queste composto.

L'ultima fase è una parte molto importante, non presente nei precedenti lavori, che ha il compito di garantire la connessione del grafo.

4.1 Dataset utilizzato per i test

Le cartografie utilizzate per testare l'algoritmo sono le reti ferroviarie di:

- Monselice
- Belluno
- Padova
- Mestre / Venezia / Marghera

Questi dataset presentano complessità differenti tra loro (vedi figure nella pagina seguente): le stazioni di Monselice e Belluno sono piccole e contengono pochi elementi; la stazione di Padova invece, ha molte più linee e quindi una complessità maggiore; infine vi è la stazione di Mestre che ospita un numero di linee quasi 10 volte il numero delle linee di Monselice. Come si può notare, quindi, il problema non è affatto banale, poiché ogni stazione ha una sua propria complessità e caratteristica.

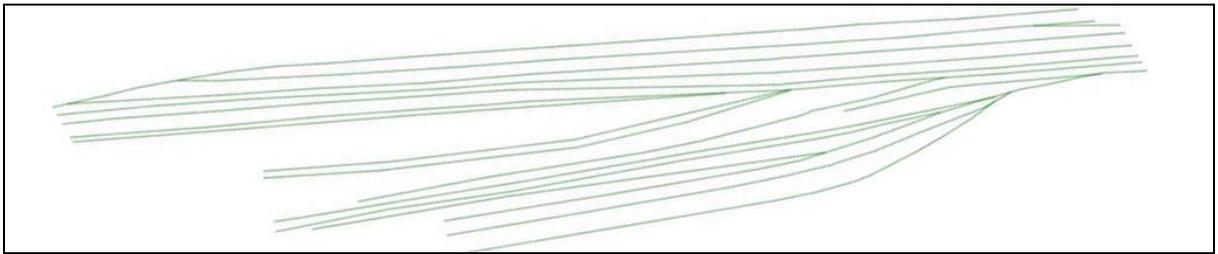


Figura 3: Dataset della stazione di Monselice

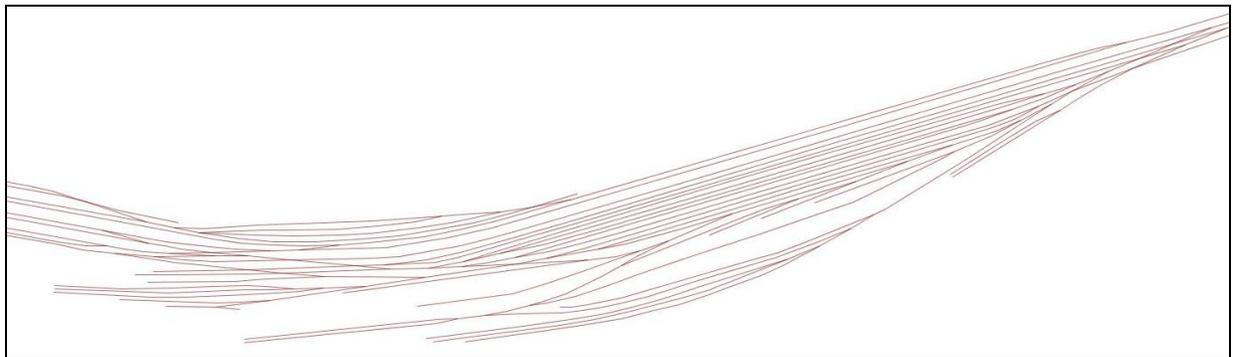


Figura 4: Dataset della stazione di Padova

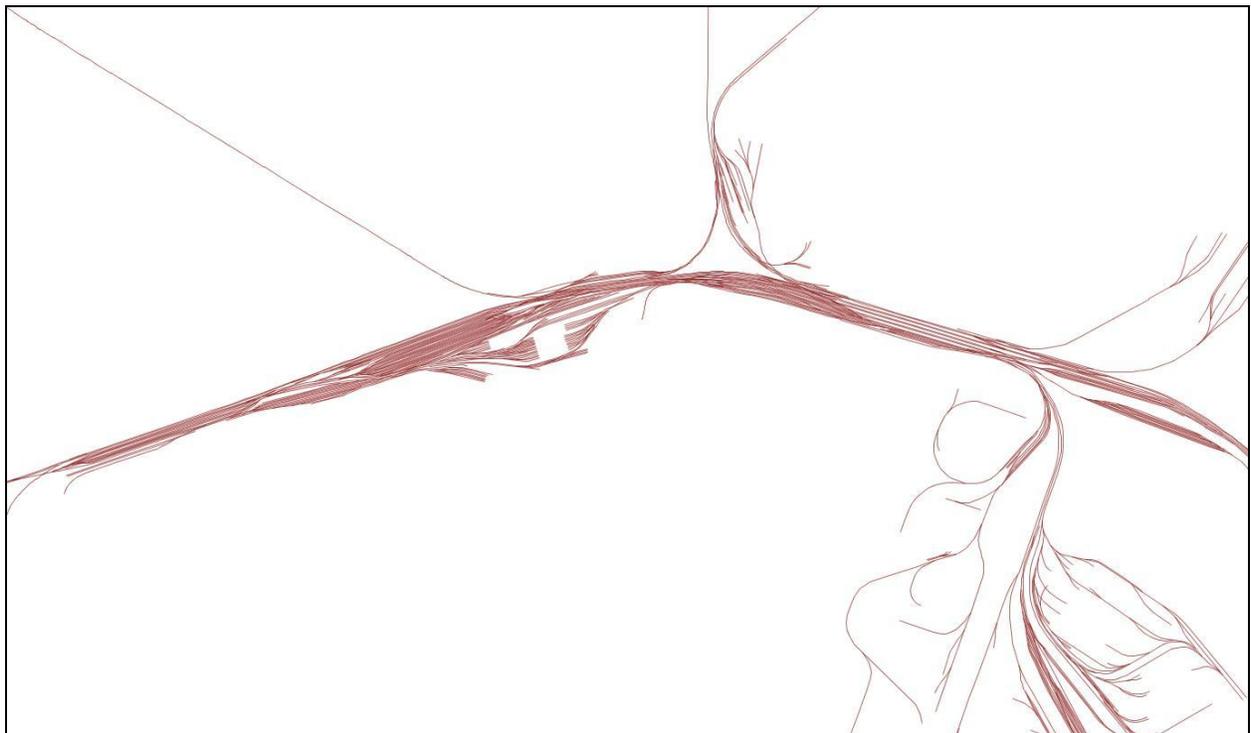


Figura 5: Dataset della stazione di Mestre

4.2 Pulizia del grafo

Prima di procedere all'esecuzione del vero e proprio algoritmo di generalizzazione, occorre inserire nel plug-in una procedura di pulizia del grafo, in quanto vi è il rischio di forti anomalie che, durante l'esecuzione della fase seguente, possono creare problemi non trascurabili.

Si tratta per lo più di problemi legati alle connessioni e alla continuità: linee non connesse perfettamente, nodi degli scambi non perfettamente puntuali, sono esempi che all'occhio possono essere trascurate, ma che, per un calcolatore, rappresentano grandi differenze.

Inoltre, per semplificare il grafo, vengono eliminate linee che non hanno alcuna rilevanza al fine del prodotto finale: sono linee isolate o molto corte, che possono solo appesantire il lavoro del processo di generalizzazione; il valore sopra il quale deve assumere la lunghezza di ogni linea per non essere eliminata verrà chiamata "*length*".

In alcune precedenti tesi si è già trattato il problema della pulizia; in questo progetto, è stato studiato un nuovo approccio strategico a questo problema, rendendo la fase di pulizia più efficiente e completa.

Di seguito verranno presentati quattro diverse fasi di pulizia e correzione; prima però verrà trattato "come" il processo riconosce quali dati devono essere corretti e che correzione svolgere.

4.2.1 KdTree

Per sviluppare questo plug-in è stata utilizzata la struttura dati ad albero KdTree¹, all'interno della quale vengono inserite una o più coordinate (eventualmente associate ad un dato generico). Quando viene eseguita una query, vengono restituiti dei KdNode. Ogni KdNode rappresenta una coordinata e il numero di volte che è stata inserita la stessa (eventualmente assieme al dato).

Usando questa struttura, l'algoritmo è stato progettato in questo modo: ogni punto terminale di ogni linea del grafo viene inserito all'interno del KdTree, assieme all'ID della linea stessa.

¹ Nella libreria JTS il KdTree è una struttura dati ad albero binario che permette una facile ricerca di elementi. Fornisce una partizione dello spazio di dimensione k e in ogni nodo dell'albero è presente un punto k-dimensionale. Il KdTree è utilizzato in molte applicazioni informatiche.

Successivamente viene restituita una lista di KdNode di tutti i punti inseriti, per poi essere analizzati uno alla volta.

Per ogni KdNode:

- Se il KdNode contiene 3 o più punti, significa che vi è uno scambio tra tre o più binari. Viene eseguito uno *"Snapping"*.
- Se il KdNode contiene 2 punti, allora si tratta di un nodo edge e pertanto si deve procedere con il *"Dissolve"*.
- Se il KdNode contiene 1 punto possiamo avere tre casi:
 - a. Il punto si trova in forte prossimità di una linea (ma non in prossimità degli estremi di questa): è uno scambio di tre binari e si deve procedere con un *"Noding"*.
 - b. Il punto terminale non si trova vicino nessuna linea:
 - i. la lunghezza della linea alla quale appartiene il punto è minore di "length": viene eliminata la linea.
 - ii. la lunghezza della linea è maggiore di "length": non si fa nulla.

Tutto il processo di controllo e correzione (comprendente tutte le fasi sopra descritte) viene ripetuto fino al punto in cui il dataset non presenta cambiamenti dal suo stato precedente, cioè quando un altro ciclo di pulizia sarebbe inutile.

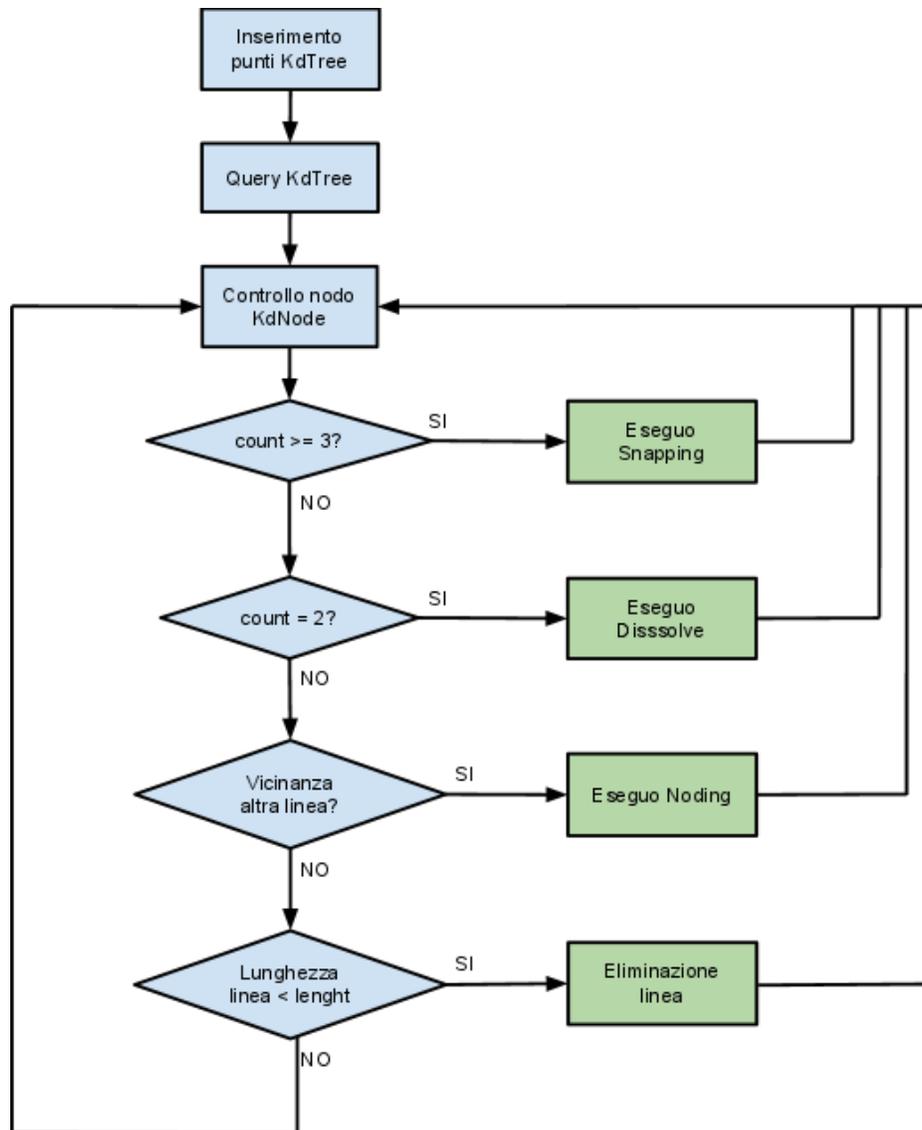
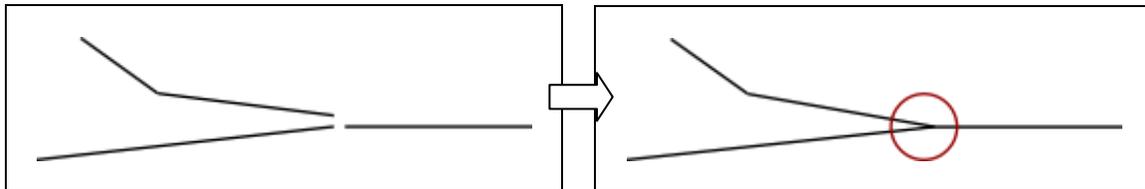


Figura 6: Percorso di decisioni dell' algoritmo di pulizia

4.2.2 Snapping

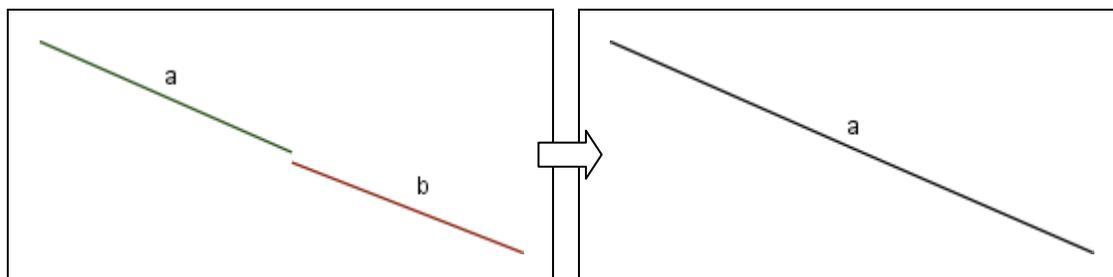
La prima fase di semplificazione del grafo è lo *Snapping*: si tratta di trovare tutti i nodi del grafo e controllare, per ogni nodo, se i punti terminali delle linee che, appunto, terminano in quel nodo, sono perfettamente connessi. Se non è così viene eseguita un'azione correttiva che posiziona i punti terminali in modo corretto, al fine di garantire la connessione del tracciato.



4.2.3 Dissolve

Questa fase cerca eventuali discrepanze nella connessione tra due linee, che, a causa di fattori che ne modificano le geometrie, per esempio i tagli dei fogli, non sono perfettamente collegate (i due punti terminali non sono sovrapposti in modo corretto).

Questa fase è simile allo *Snapping*, con la sola differenza che, oltre a sistemare i punti terminali, fonde le due linee connesse in quel punto in una sola linea, cancellando una delle due linee e modificando l'altra, in modo che vi sia una sola linea al posto di due.

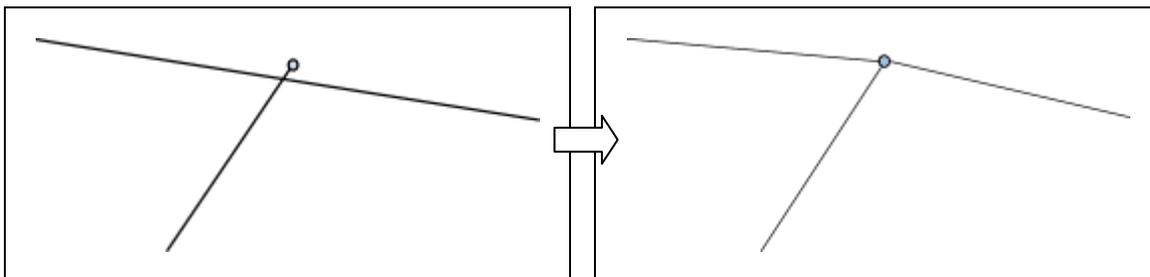


4.2.4 Noding

Alcune parti del tracciato presentano delle discontinuità negli scambi dei binari: vi possono essere linee che ne raggiungono un'altra tramite un nodo ma che di fatto non è connessa perfettamente a quel nodo, semplicemente termina nelle vicinanze di quella, sovrappondendosi di pochissimo o terminando un po prima di raggiungerla. Bisogna evidenziare che l'altra linea non presenta nessuna interruzione in prossimità di quel punto, che però – come detto prima - rappresenterebbe uno scambio. Queste *micro interruzioni* non si notano quando la scala è bassa, ma emergono solo se si osservano le mappe con scala molto alta, e per il calcolatore e per la corretta esecuzione del plug-in sono assai gravi.

L'obiettivo del *Noding* è quello di trovare queste anomalie e di correggerle.

Viene creato un buffer, di diametro uguale alla tolleranza scelta, che verrà a sovrapporsi ad una linea nel caso ve ne sia una nelle vicinanze di quel punto. Questa linea, resa nota dal suo ID, viene divisa in due parti, proprio in prossimità del punto in questione. Queste due linee, alla fine, vengono collegate proprio al nodo esaminato.

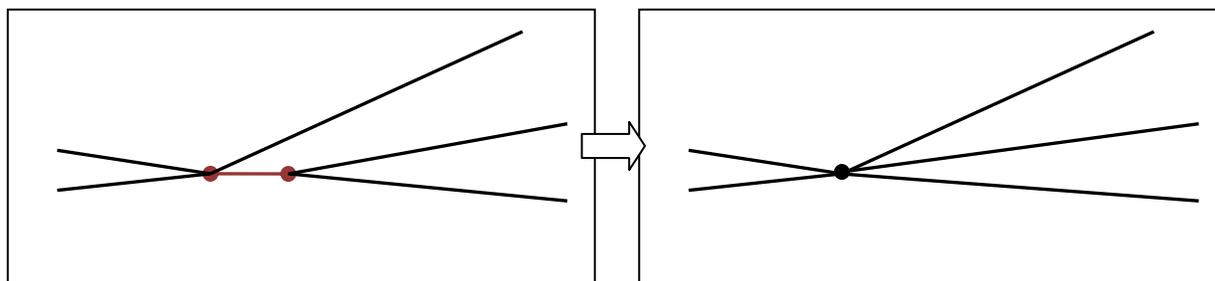


4.2.5 Eliminazione linea

Questo processo elimina semplicemente tutte le linee che hanno una lunghezza inferiore ad un certo valore. Le linee che vengono eliminate non devono essere vitali per la continuità del tracciato; in altre parole tutte le linee che non presentano caratteristiche tali da essere considerate nei casi precedenti (cioè quelle linee che hanno uno o tutti e due i terminali liberi) possono essere trascurate e quindi eliminate. Il controllo della lunghezza viene perciò svolto solo sui nodi del KdTree che hanno un solo punto terminale.

4.2.6 Collapse

Oltre alle fasi sopra descritte, occorre dar merito ad un'altra piccola funzione che semplifica ulteriormente il grafo: il suo scopo è collassare una linea molto corta, che collega due o più linee, in un punto. In questo modo viene ridotta la complessità in quanto due nodi vengono collassati in un unico connettore.



4.3 Creazione e scelta delle coppie

Per affrontare il problema della generalizzazione delle ferrovie è stata seguita una linea di studio già sviluppata in un lavoro di tesi precedente, che partiva dall'ipotesi che due linee solitamente parallele e molto vicine tra loro possano essere fuse in un'unica linea.

Ovviamente questa ipotesi non può essere applicata in tutto e per tutto, perché la struttura stessa del grafo ferroviario non presenta solo questo tipo di coppie.

Ciò nonostante questo approccio sembra dare una soluzione anche a quelle parti di tracciato che appaiono di difficile elaborazione.

Di seguito verrà descritta una strategia che, partendo da questa idea di base, è stata migliorata ulteriormente per quanto riguarda prestazioni, efficienza ed efficacia, ed estesa in modo significativo, al fine di garantire una certa qualità del risultato finale.

4.3.1 Fase 1: Creazione delle linee ortogonali

Per trovare le coppie di linee vicine è stata seguita una strategia che vede l'utilizzo di una serie di linee ortogonali alle linee del livello. Il processo accetta alcuni parametri quali la lunghezza delle linee ortogonali (che da ora in poi chiameremo "*lineort*") e la distanza tra di esse. Ogni linea del tracciato presenta quindi una serie di *lineort* che seguono il suo andamento. E' emerso che il ricorso a questa operazione, oltre che essere di semplice implementazione, offre uno strumento molto potente: le *lineort* ci danno una direzione e possono informarci su quale linea sta "sopra" o "sotto" la linea in esame e soprattutto "quante volte" una linea ne tocca un'altra. Questo approccio evidenzia informazioni molto importanti, poiché ci mette conoscenza delle linee che stanno più vicine ad una e quali hanno più requisiti di stare in coppia con quella.

Ogni *lineort* viene rappresentata come una Feature con la sua *geometry* e l'*id* della linea alla quale appartiene.

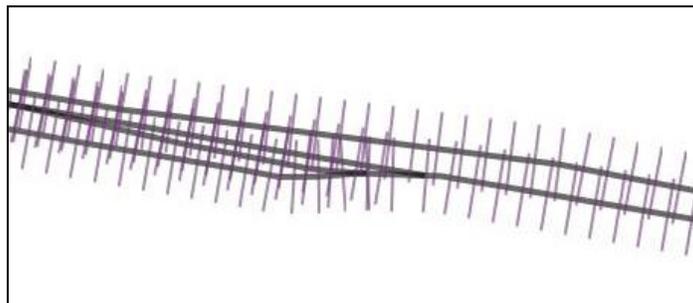


Figura 7: Linee ortogonali applicate al tracciato

4.3.2 Fase 2: Creazione delle coppie candidate

A questo punto si esegue, in linea di massima, uno *spatialJoin* tra tutte le *lineort* e le linee del tracciato. In questo modo si viene a creare una tabella con tutte le interazioni tra le due parti.

Ci troviamo di fronte ad una tabella formata da due colonne: Id1 e Id2.

Id1 rappresenta l'id di una linea, Id2 rappresenta l'id di una linea la cui *lineort* tocca id1.

A questo punto è necessaria l'implementazione di una struttura che rappresenta una "coppia":

Ogni coppia contiene le seguenti informazioni:

1. ID1 e ID2: l'id delle due linee candidate ad essere accoppiate;
2. NC: numero di *lineort* (totali) che toccano la propria contro-linea;
3. DM: distanza tra le due linee;
4. %COP_{ID1} e %COP_{ID2}: percentuale di copertura; ne parleremo in seguito.

Questa classe è molto utile perché permette una facile organizzazione e un facile reperimento delle informazioni relative ad ogni coppia di linee.

Una volta creata la classe si procede con la scansione di ogni linea ortogonale: viene caricata la lista di tutte le linee che tocca quella *lineort* e poi, per ognuna di quelle linee, viene creata una istanza Coppia contenente prima di tutto:

1. ID1 = id della linea alla quale appartiene la *lineort*;
2. ID2 = id della linea corrente della lista;
3. DM = distanza tra le linee ID1 e ID2.

Questa Coppia viene aggiunta ad una `ArrayList<Coppia>` e se è la prima coppia ad essere inserita, viene istanziato anche il numero di *lineort* in comune NC (posto ovviamente a 1, dato che questa è la prima *lineort* esaminata) e la percentuale di copertura.

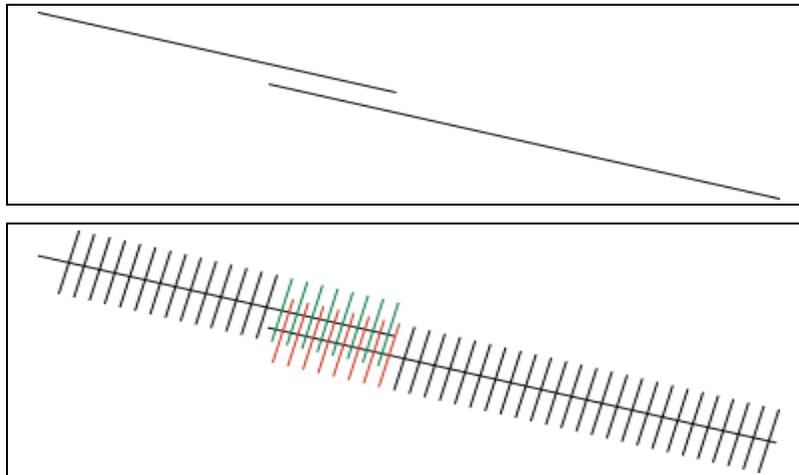
Se invece l'`ArrayList<Coppia>` non è vuota, si esamina se la Coppia(ID1, ID2) corrente è già presente: se non lo è allora la si aggiunge come prima, altrimenti viene aggiornata la coppia già presente, aumentando di 1 l'NC e aggiornando di conseguenza la %COP.

In questo modo otteniamo una lista di coppie di linee con le informazioni relative ad ognuna.

La percentuale di copertura

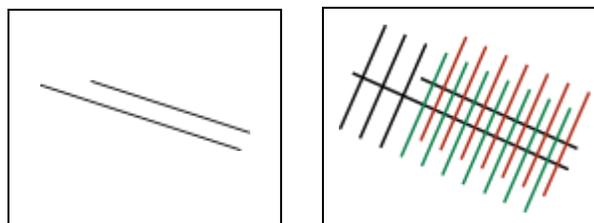
Questo valore è molto, se non l'unico, importante. Consente di valutare quanto due linee sono "accoppiate". Per spiegare il senso di questo accoppiamento, è bene farne un esempio.

Immaginiamo due linee come quelle in figura:



Se ci limitassimo a contare semplicemente il numero di linee ortogonali che toccano l'altra linea otteniamo **16 linee ortogonali in comune**.

Passiamo ora ad un'altra coppia di linee, nella stessa scala della prima:



Possiamo notare che in questo caso abbiamo **14 linee ortogonali in comune**.

Analizzando i due casi e osservando solo l'NC si può dire che la coppia più "accoppiata" è quella con NC maggiore, cioè quella con NC = 16.

Ma è facile vedere che non è così: la coppia del primo caso è formata da linee che sono parallele ma che si coprono molto parzialmente, al contrario della seconda coppia nella quale invece c'è una più correlazione tra le due linee.

E' per questo che viene introdotta la percentuale di copertura definita come:

$$\%COP_{ID1} = (distanzaLinee * nSbarreInComune) / lunghezzaLinea_{ID1}$$

$$\%COP_{ID2} = (distanzaLinee * nSbarreInComune) / lunghezzaLinea_{ID2}$$

Descritta a parole, si tratta del rapporto tra la lunghezza della parte di linea che copre l'altra (o che viene coperta dall'altra) e la lunghezza della linea stessa. In questo modo riusciamo a capire il grado di adiacenza tra due linee.

4.3.3 Fase 3: Scelta delle coppie

Il prossimo step sta nello scegliere la coppie più valide. A tal proposito viene aggiunto ad ogni coppia un altro parametro VAL, che ha lo scopo di dare un "voto" alla coppia, che dipende dalla distanza tra le due linee, dal numero di *lineort* in comune e soprattutto dalla percentuale di copertura. Ognuno di questi fattori viene pesato da un coefficiente appropriato. Il parametro VAL viene calcolato in questo modo:

$$VAL = \frac{NC * c + cop_{ID1} * cop_{ID2} * p}{(dm+1) * d}$$

dove:

- c = coefficiente per il numero di *lineort* in comune
- p = coefficiente per la copertura
- d = coefficiente per la distanza

Come si può notare il parametro VAL aumenta all'aumentare del numero di *lineort* e della copertura, ma diminuisce all'aumentare della distanza tra le linee.

Il processo inizia caricando una coppia, passandola attraverso un filtro che controlla la sua validità, cioè se soddisfa i parametri minimi impostati. Viene cercata la coppia con VAL più alto, aggiungendola ad un'altra lista, COPPIE_SCELTE. Dal secondo ciclo in poi l'algoritmo cerca la coppia successiva e poi controlla se nella lista COPPIE_SCELTE c'è almeno una coppia con almeno una linea uguale ad una linea della coppia che si sta esaminando. In caso affermativo la coppia non viene aggiunta, altrimenti sì. In questo modo si evita che una linea faccia parte di due o più coppie, perché, nella prossima fase, queste coppie devono formare una sola linea intermedia.

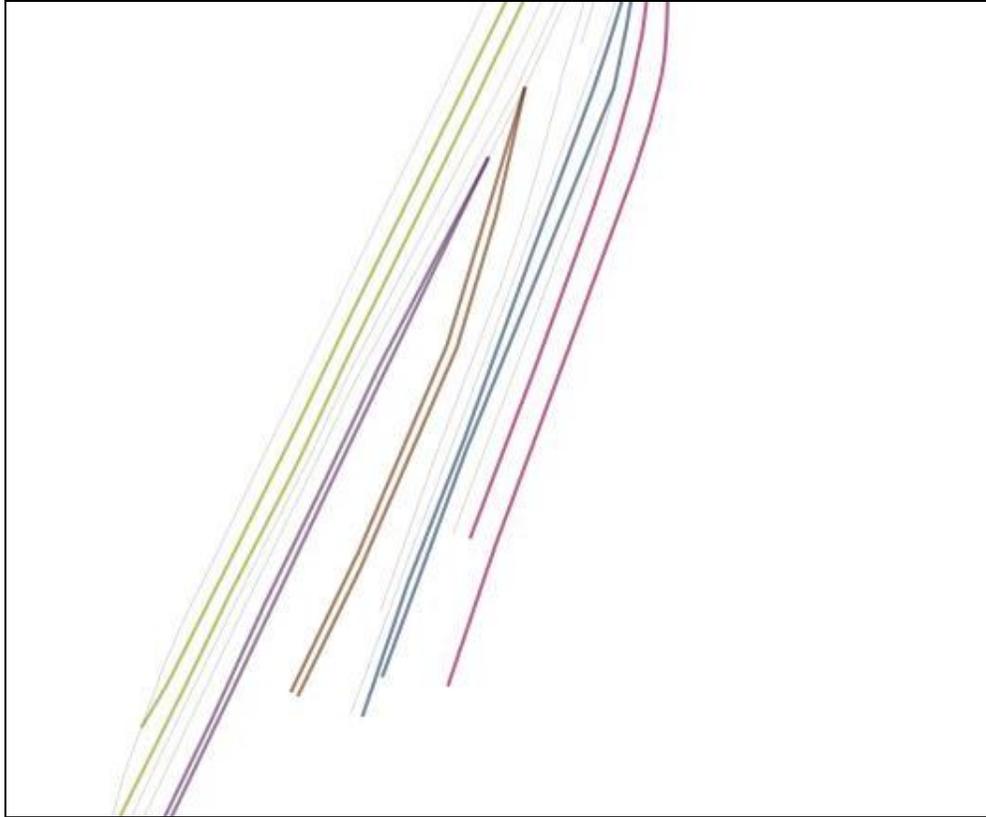


Figura 8: Le coppie scelte

4.4 Creazione delle *mid-line*

A questo punto abbiamo ottenuto una tabella con la lista di tutte le coppie valide del tracciato. Ora bisogna procedere con una trasformazione che consente di rendere le linee più distanziate tra loro, in modo da renderle adattabili alla scala scelta.

La soluzione scelta è calcolare una semplice linea intermedia tra le due linee di una coppia. Il metodo utilizzato per la formazione della *mid-line* è quello per punti mediani: si effettua un campionamento a entrambe le linee e, per ogni coppia di vertici, si crea il punto intermedio. Una volta completato il processo, si connettono tutti i punti intermedi fino a formare la *mid-line*.

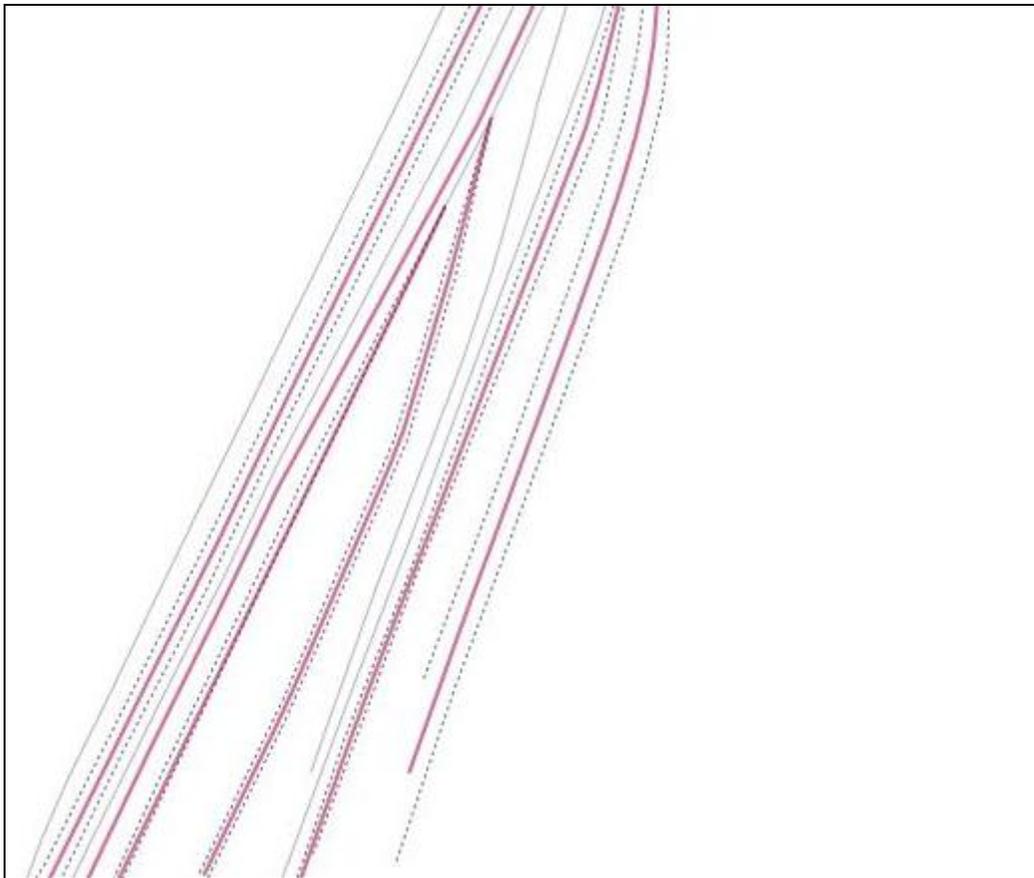


Figura 9: Le *mid-line* create dalle coppie

4.5 Gestione delle connessioni

L'ultima parte del plug-in, è di vitale importanza per la connettività del grafo.

Se ci si fosse limitati a creare le mid-line, si avrebbe ottenuto un grafo con gravi problemi di connettività: una volta eliminate le linee delle coppie e tenendo le mediane, queste ultime non sono connesse alle linee alle quali erano connesse quelle di prima. Si è deciso quindi di estendere l'algoritmo aggiungendo una importante funzione che gestisce questo tipo di problema.

Prima di tutto si sono caricati tutti i punti terminali delle linee in un KdTree; per ogni coppia di linee si sono create quattro liste contenenti le coordinate dei 4 punti terminali e i rispettivi id delle linee connesse a quei punti.

Vengono quindi esaminate le due coppie di liste (quelle dei punti iniziali e finali) controllando quali e quante linee sono in comune.

Si ottengono tre casi:

1. Le due liste non hanno linee in comune: è il caso peggiore per quanto riguarda la complessità del grafo; la sola cosa per rendere il grafo connesso è connettere la mediana ai due punti tramite due linee. Questa soluzione presenta uno svantaggio, poiché si vengono a creare nuove linee quando invece la natura stessa della generalizzazione dovrebbe eliminarne o al massimo semplificarle.
2. I nodi sono sovrapposti: in questo caso la mediana si connette automaticamente ai due nodi, che hanno le stesse coordinate. Non si fa nulla.
3. Le due liste hanno una linea in comune: in questo caso i due nodi vengono fusi in uno e le linee connesse a tali nodi (compresa la mediana) vengono connesse a questo nodo. La linea in comune viene naturalmente eliminata.

Le due figure nella prossima pagina mostrano due esempi dei tre casi sopra riportati.

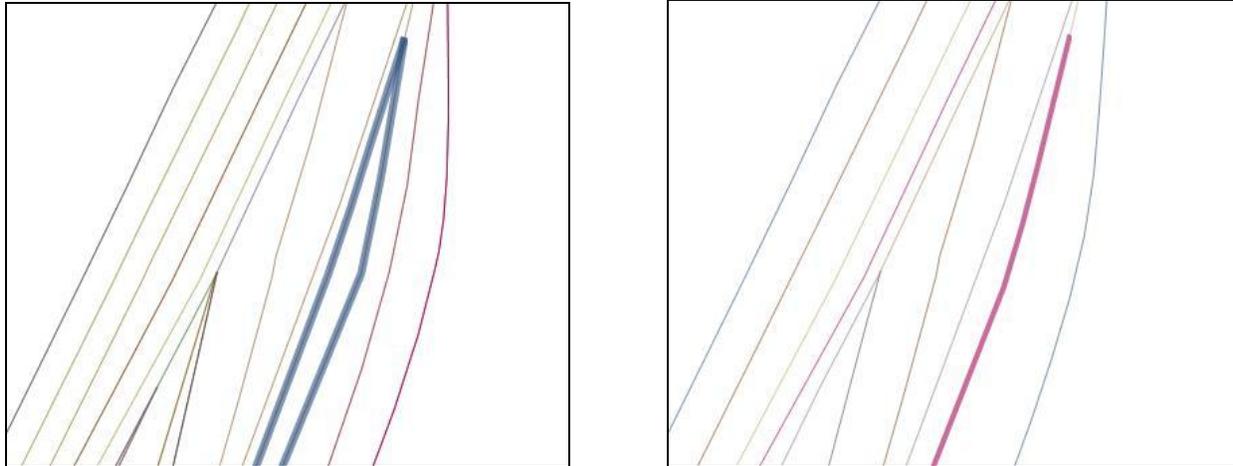


Figura 10: La figura mostra una coppia di linee che si riconduce al caso 2, caso cioè nella quale i due punti terminali si sovrappongono, formando un unico punto. L'algoritmo in questo caso si limita a creare la mid-line, poiché il punto terminale di quest'ultima si congiunge automaticamente al nodo.

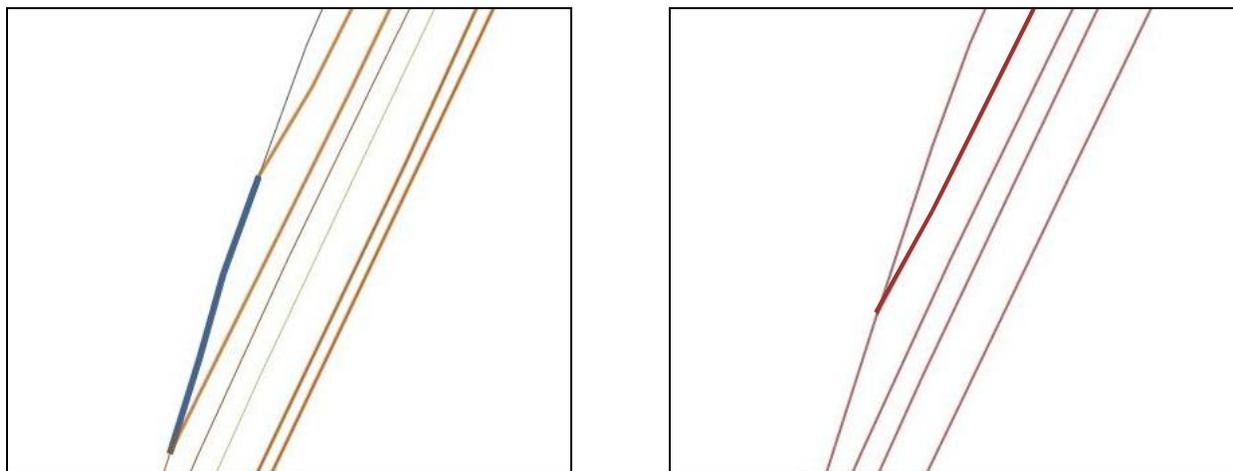


Figura 11: Viene mostrato il caso 3, in cui le due linee della coppia sono connesse da una terza linea comune. In questo caso, una volta creata la mid-line, la linea comune viene eliminata e i due nodi di connessione vengono sovrapposti nel punto terminale della mid-line stessa. Le altre linee connesse a tali nodi vengono automaticamente collegate a quest'ultimo.

Capitolo 5

Conclusioni

Lo sviluppo dell'algoritmo tramite approccio "a coppie" è risultato abbastanza vincente. La fase iniziale di pulizia del grafo offre una considerevole diminuzione della complessità, con una conseguente miglior efficienza della fase successiva.

La struttura dati che ha permesso il controllo dei punti terminali delle linee è il KdTree, che viene appunto utilizzato più volte per diverse analisi.

I risultati ottenuti alla fine del lavoro sono stati buoni, con alcune difficoltà nella fase finale per la gestione delle connessioni.

Per quanto riguarda gli sviluppi futuri, la cosa che salta maggiormente all'occhio sono alcune discrepanze e intersezioni di linee in prossimità degli scambi. Infatti in alcuni casi, in presenza di molti scambi vicini tra loro, l'algoritmo ha diverse difficoltà a generalizzare quelle porzioni, poiché il numero di linee e di connessioni è elevato.

Sarebbe importante quindi gestire questi casi particolari, al fine di semplificare ulteriormente il tracciato e renderlo più pulito.

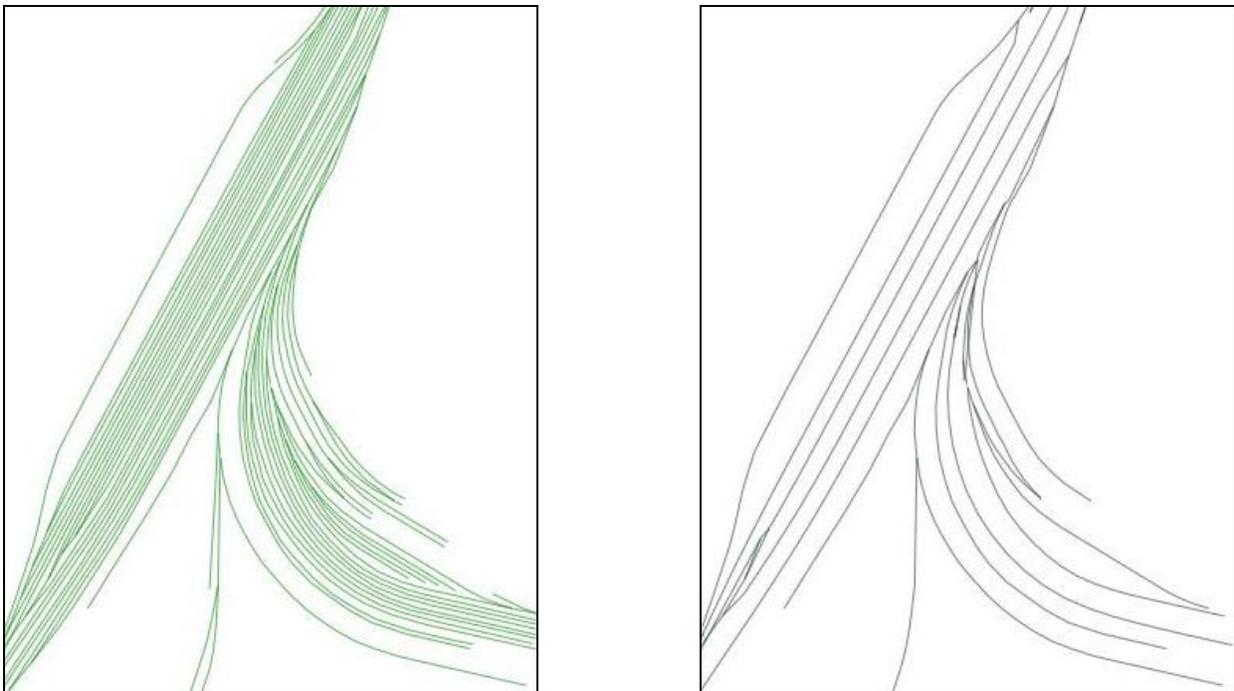


Figura 12: Il grafo della stazione ferroviaria di Padova prima (scala 1:5 000, a sinistra) e dopo la generalizzazione (scala 1:25 000, a destra).

Bibliografia

IGM (2004) *Norme e Segni convenzionali per la realizzazione dei fogli della Carta d'Italia alla scala 1:50 000* - Firenze - Istituto Geografico Militare.

Michael T. Goodrich, Roberto Tamassia (2007) *Strutture Dati e Algoritmi in Java* - Prima edizione italiana - Zanichelli.

McMaster R. B. e Shea K. S. (1992) *Generalization in Digital Cartography* - Association of American Geographers, Washington.

CARGEN (2009) *La generalizzazione del DB25 a partire dal GeoDBR 1:5000 - Schema della derivazione, query per il popolamento e operazioni sulle geometrie* - Dipartimento di Ingegneria dell'Informazione, Padova.

De Gennaro M., Rumor M., Savino S. (2009) *Le procedure per la derivazione del DB25 dal DBT della Regione del Veneto: risultati del progetto CARGEN* - Bollettino della Associazione Italiana di Cartografia, AIC, 135 aprile 2009.

Istituto Geografico De Agostini (1999) *Atlante Geografico Illustrato* – Istituti Geografico DeAgostini, Novara, 1999.

Sandro Savino (2011) *A solution to the problem of the generalization of the Italian geographical databases from large to medium scale: approach definition, process desing and operators implementation* - Tesi di dottorato, Università degli studi di Padova, 2011.

OpenJUMP (2004) *Jump Pilot Project online: <http://www.openjump.org/>*

Vivid Solutions (2003), *JTS Topology Developer'S Guide*, Version 1.4
URL: <http://www.vividsolutions.com/jts/jtshome.htm>.