

Alma Mater Studiorum - Università di Bologna

SCUOLA DI SCIENZE
Corso di Laurea in Scienze dell'Informazione

**PATTERN ANTICIPI E RITARDI
DEL TRASPORTO PUBBLICO:
UN'APPLICAZIONE DI DATA
VISUALIZATION SU DATASET
REALI**

Relatore: Chiar.mo
Prof. FABIO VITALI

Presentata da:
MAXIMILIANO MARSIGLI

Sessione III
A.A. 2013/2014

Indice

1. INTRODUZIONE.....	5
2. DATA VISUALIZATION E CARTOGRAFIA	11
2.1 Data Visualization	11
2.2 Visualizzazioni di valori scalari	13
2.3 Rappresentazione topologica della rete di trasporto.....	18
2.4 Cartografia digitale.....	20
2.5 GIS e WebGIS.....	24
2.6 Data Visualization nel Trasporto Pubblico	27
2.7 Visualizzazioni di anticipi e ritardi nel trasporto pubblico	29
2.8 Considerazioni sulle soluzioni in letteratura	30
3. GARTP: GEOVISUALIZZAZIONE DI ANTICIPI E RITARDI DEL TRASPORTO PUBBLICO	33
3.1 Che cosa fa GARTP	33
3.2 A cosa serve GARTP.....	34
3.3 Descrizione dell'interfaccia di GARTP	35
4. ARCHITETTURA DI GARTP	51
4.1 Struttura concettuale e concreta dell'applicazione.....	51
4.2 Principali tecnologie utilizzate	52
4.3 Utilizzo e personalizzazione del software	55
5. USABILITA' DI GARTP.....	59
5.1 Test di usabilità.....	61
5.2 Risultato della SUS.....	62
6. CONCLUSIONI	63
7. BIBLIOGRAFIA	65
Appendice A – Servizio giornaliero.....	67
1. Copertura del servizio giornaliero.....	67
2. Turno macchina.....	68

Indice figure

Figura 1 - Interfaccia di GARTP.....	8
Figura 2 - Carte figurative des pertes successives en hommes de l'Armée Française dans la campagne de Russie 1812-1813.....	13
Figura 3 - Mappa arcobaleno	16
Figura 4 - Altre scale	17
Figura 5 - Raster	23
Figura 6 - Schermata di avvio del GARTP.....	35
Figura 7 - Comandi di navigazione	36
Figura 8 - Parametri di filtro dati	36
Figura 9 - Catalog.....	37
Figura 10 - Zona mappa	38
Figura 11 - Barra di stato.....	38
Figura 12 - Visualizzazione di una linea.....	40
Figura 13 - Visualizzazione di due linee.....	41
Figura 14 - Visualizzazione archi senza rilevazione.....	42
Figura 15 - Attivazione Zoom In.....	43
Figura 16 - Risultato Zoom In.....	44
Figura 17 - Attivazione misurazione.....	45
Figura 18 - Fase di misurazione	46
Figura 19 - Attivazione identifica.	47
Figura 20 - Selezione identifica.	48
Figura 21 - Esito identifica.....	49
Figura 22 - Struttura di GARTP.....	51
Figura 23 - Percentile SUS.....	60
<i>Figura 24 - Associazione percentile SUS aggettivo.....</i>	<i>61</i>

Indice tabelle

Tabella 1 - Primitive della cartografia digitale	21
Tabella 2 - Risultato test di usabilità.....	62

1. INTRODUZIONE

Scopo di questo progetto è realizzare un sistema software, denominato GARTP (Geovisualizzatore Anticipi Ritardi del Trasporto Pubblico), che consenta di visualizzare e navigare pattern di anticipi e ritardi del trasporto pubblico. Il dataset visualizzato è costituito dalle rilevazioni dell'intera flotta telecontrollata di una reale azienda di trasporti nel periodo compreso tra gli anni 2011 e 2014. Un test di usabilità, secondo la *System Usability Scale (SUS)*, è stato condotto su alcuni addetti alla progettazione del servizio della stessa azienda per verificare il funzionamento di GARTP.

La progettazione di un servizio di trasporto prevede diversi passi: l'analisi della domanda di mobilità, la scelta del percorso delle linee, la stima dei viaggiatori da trasportare, la determinazione dei tempi di percorrenza della linea, l'identificazione del veicolo più adatto all'esercizio e il calcolo della frequenza del passaggio alle fermate. La progettazione è seguita dall'erogazione del servizio. E' in questa fase che si presentano i problemi che richiedono una riprogettazione del servizio. Uno dei motivi principali è dovuto alla differenza tra l'orario programmato e l'orario effettivo di transito. Le moderne aziende di trasporto pubblico hanno attrezzato le proprie flotte di veicoli con sistemi AVM, cioè di Automatic Vehicle Monitoring, che utilizzano la localizzazione come elemento principale del monitoraggio. In genere, la localizzazione è effettuata con il GPS, e il meccanismo di trasmissione è un satellite, una radio o una connessione via cellulare [MBPTT04]. Gli AVM possono valutare se il bus arriverà alla prossima fermata all'orario programmato o, in caso contrario, con quanto anticipo o con quanto ritardo. Questa stima è registrata e trasferita poi a un database dedicato all'archiviazione del monitoraggio del servizio. La mole dei dati raccolti dal monitoraggio, anche sul breve periodo, può essere enorme per flotte composte di numerosi veicoli. Anche limitandosi alla sola informazione dell'anticipo/ritardo nell'erogazione del servizio, la dimensione del dato da analizzare rimane molto vasta. Come consentire ai programmatori del servizio di capire e impiegare l'informazione di cui necessitano?

L'assenza di uno strumento specifico per la visualizzazione degli anticipi/ritardi porta gli operatori a utilizzare strumenti di analisi in grado di gestire solo una parte molto limitata dell'intero insieme di dati disponibili. I confronti tra i dati rilevati in periodi diversi dello

stesso anno o di anni diversi richiedono una serie così articolata di operazioni da disincentivare gli addetti dal procedere nell'elaborazione. L'effetto ultimo è che la base di conoscenze, su cui sono scelti gli interventi da operare per migliorare il servizio, è molto più piccola di quanto potrebbe essere e la riprogettazione potrebbe essere incompleta o insufficiente a risolvere i problemi sperimentati dal cliente del trasporto pubblico.

La Data Visualization è un'area relativamente nuova degli studi che sta acquisendo sempre più l'attenzione degli ambienti sia accademici sia industriali. Il suo obiettivo è definire un insieme di procedure e di tecniche per produrre delle rappresentazioni visive che rendano esplicite all'osservatore le strutture o le relazioni esistenti tra i dati oggetto della rappresentazione. La Data Visualization è un'area multidisciplinare della scienza e include elementi dell'interazione uomo-computer, informatica, grafica, progettazione visuale, psicologia, e processi economici. Attualmente la Data Visualization ha un ampio raggio di applicazioni che includono: ricerca scientifica, librerie digitali, data mining, analisi di dati finanziari, studi di mercato, controlli della produzione manifatturiera e mappatura dei crimini.

Una rete di trasporto pubblico può essere espressa attraverso l'uso di un grafo orientato in $G(N, A)$. I nodi dell'insieme N rappresentano i punti strategici sul territorio che sono raggiunti dalla rete di trasporti. Ogni nodo è caratterizzato da una sigla, da un nome e da una posizione geografica. Gli archi dell'insieme A rappresentano i tratti di strada che devono essere percorsi per andare da un nodo origine a un nodo destinazione. Entrambi i nodi appartengono all'insieme N . Ogni arco appartenente all'insieme A è caratterizzato da una lunghezza e dal tempo impiegato per percorrerlo detto tempo di percorrenza.

Un percorso appartenente alla rete G è descritto da una sequenza ordinata di archi.

Ogni corsa effettuata sulla rete G segue un percorso ed è descritta dalla successione di coppie (nodo, orario di transito). Il servizio di trasporto pubblico erogato su una rete G è l'insieme delle corse definite sulla rete G .

La rappresentazione cartografica si è evoluta nel tempo. Si è passati dalla rappresentazione su carta alla rappresentazione numerica consentendo al mondo dell'informatica di impiegare questo tipo d'informazione. Esistono due formati per rappresentare una cartografia: i vector e i raster. I vector utilizzano le primitive punto, linea, area e volume per descrivere gli

oggetti nella cartografia. I raster ricorrono a delle matrici di punti colorati per rappresentare la cartografia.

Gli strumenti nati per la gestione, la modifica e la consultazione della cartografia digitale ricadono nella categoria detti Geographic Information System (GIS).

Sono detti WebGIS i sistemi informativi geografici (GIS) che sono pubblicati su web. Un WebGIS è quindi la migrazione al web degli applicativi nati e sviluppati per gestire la cartografia numerica. Un progetto WebGIS si distingue da un progetto GIS per le specifiche finalità di comunicazione e di condivisione delle informazioni con altri utenti.

Con i WebGIS le applicazioni GIS sviluppate per utenze stand-alone o in ambienti LAN possono essere implementate su web server (anche detto *map-server*) consentendo l'interazione attraverso internet con la cartografia e con i dati a essa associati. Gli esempi più noti di WebGIS sono gli applicativi web per la localizzazione cartografica, lo stradario oppure l'atlante online. Le applicazioni WebGIS sono utilizzabili attraverso i browser internet, talvolta con l'impiego di specifici plug-in, oppure per mezzo di software distinti come Google Earth. Altri esempi tipici di applicazioni GIS pubblicate in versione WebGIS sono i sistemi informativi territoriali (SIT) delle Regioni e di diversi Comuni: questi rendono accessibili ai cittadini le informazioni di carattere ambientale, urbanistico, territoriale offrendo una navigazione su base cartografica delle stesse.

Lo strumento GARTP è un'applicazione web, funzionante su tutti i browser più recenti, che consente all'operatore di selezionare la o le linee di cui è interessato vedere l'anticipo o il ritardo e di indicare l'arco temporale su cui compiere l'analisi, le condizioni meteo, la presenza di blocchi alla viabilità e il tipo di alimentazione dei bus.

La risposta ottenuta è rappresentata da un insieme di segmenti colorati – come descritto nella leggenda visibile nel lato inferiore destro della schermata - che rappresentano la differenza tra il tempo di percorrenza programmato per percorrere un certo tratto di strada e il tempo medio impiegato a percorrere lo stesso tratto di strada dai bus, telerilevati, impiegati nel servizio di trasporto pubblico.

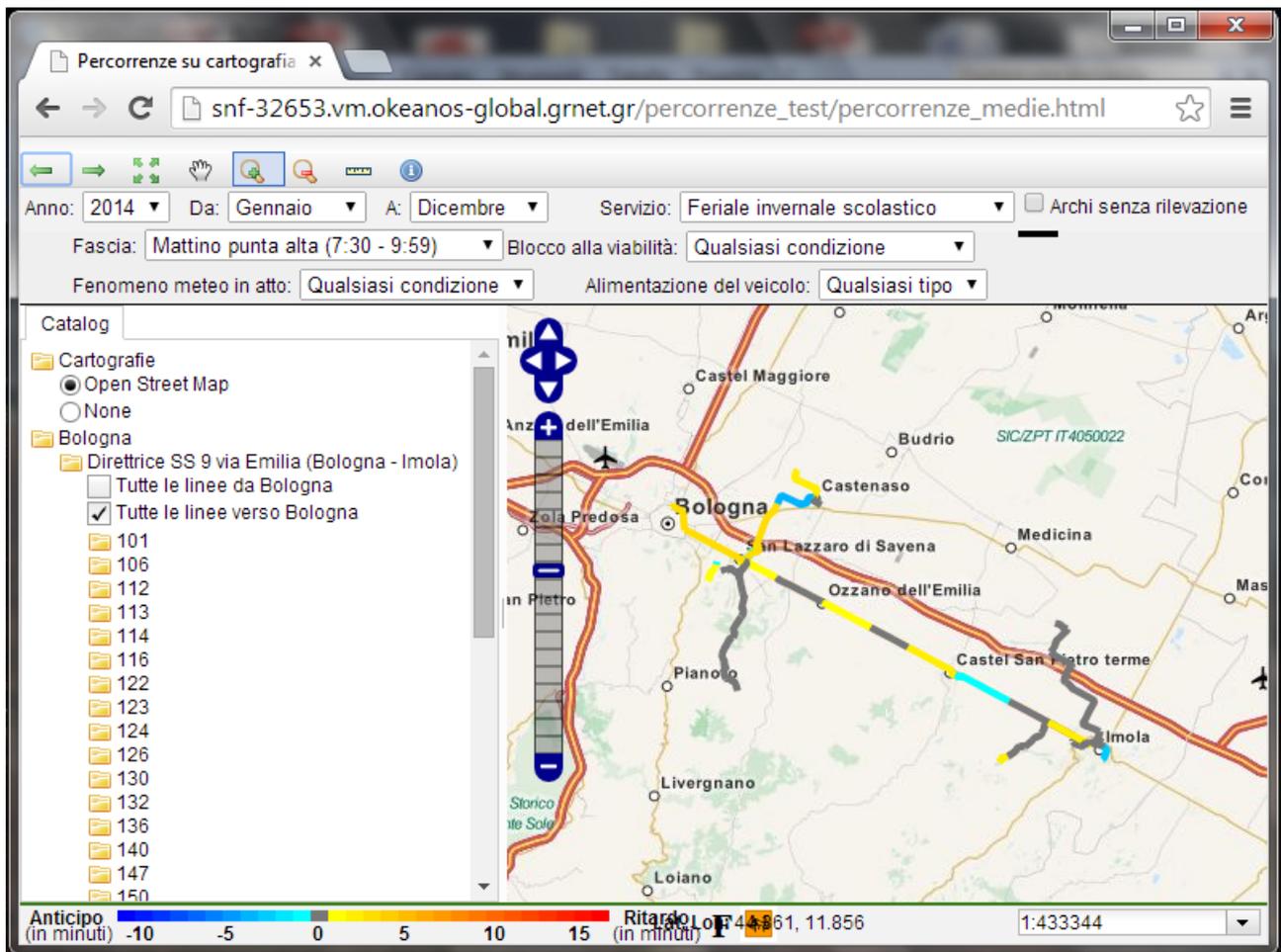


Figura 1 - Interfaccia di GARTP

Nella *Figura 1* è visibile la risposta di un'interazione tra l'operatore e l'applicazione Web.

L'applicazione web presentata appartiene alla classe delle applicazioni WebGIS Client, Oltre alla visualizzazione dei pattern di anticipo e ritardo, consente di zoomare su una zona del territorio, di spostarsi, di eseguire misure di distanze, di visualizzare altre informazioni di un segmento cliccandoci sopra.

La soluzione è stata realizzata per coadiuvare l'analisi del settore progettazione e pianificazione del servizio. E' una soluzione specialistica mirata a specifici operatori. Un sottoinsieme degli addetti a tali settori è stato sottoposto a un test di usabilità secondo la System Usability Scale (SUS). Il risultato è stato di 86,4 equivalenti all'aggettivo "Excellent". I commenti sono stati positivi e hanno riguardato sia l'immediatezza con cui l'informazione era colta nel suo insieme sia la comodità nel poterla navigare. GARTP nel

suo complesso si è rivelato stabile e piuttosto veloce nel rispondere alle richieste degli operatori. L'aspetto pratico più apprezzato è stato la scelta di impiegare tecnologie Web. Gli addetti si sono trovati a loro agio proprio perché la loro interazione con l'applicazione avviene con un comune browser web, uno strumento normalmente impiegato dagli intervistati.

GARTP è un esempio di una geovisualizzazione, cioè un'applicazione delle tecniche di Data Visualization a dati collocati spazialmente, nel campo del trasporto pubblico. Nel capitolo 2 saranno date la definizione di Data Visualization e di alcune sue tecniche di rappresentazione del dato. Quindi introdurremo la rappresentazione topologica della rete dei trasporti, i concetti fondamentali della cartografia digitale e gli strumenti nati per la sua gestione. Termineremo il capitolo presentando alcuni esempi reali di geovisualizzazione nel campo dei trasporti. Le caratteristiche fondamentali e le istruzioni d'uso di GARTP saranno presentate nel capitolo 3. L'architettura complessiva di GARTP e la scelta delle tecnologie impiegate per realizzarlo sono riportate nel capitolo 4. Una test sulla usabilità di GARTP, su un campione reale di addetti alla progettazione del trasporto pubblico, è illustrato nel capitolo 5. L'analisi conclusiva e possibili sviluppi di GARTP sono proposti nel capitolo 6. Il capitolo finale riassume la bibliografia. In appendice è proposta un'introduzione ai concetti di servizio giornaliero e di turno macchina.

2. DATA VISUALIZATION E CARTOGRAFIA

Introduciamo, in questo capitolo, l'insieme delle conoscenze e delle tecniche che sono alla base della progettazione di GARTP.

2.1 *Data Visualization*

Definiamo Data Visualization la raffigurazione e la presentazione di dati che sfruttano la nostra capacità di percezione visiva per amplificare la cognizione [W04].

Analizziamo gli elementi chiave della definizione:

- la raffigurazione del dato è il modo nel quale decidiamo di rappresentare il dato attraverso una forma fisica. Che sia una linea, una barra, un cerchio, o un'altra forma visuale, si considera il dato come materia grezza e si crea una rappresentazione che meglio ritragga i suoi attributi.
- La presentazione del dato va oltre la sua raffigurazione e si preoccupa di come integrare la raffigurazione nel complesso della sua comunicazione, includendo la scelta dei colori, le annotazioni e le funzionalità d'interazione.
- Sfruttare la nostra capacità di percezione visiva fa riferimento alla comprensione scientifica di come i nostri occhi e il nostro cervello elaborano l'informazione più efficacemente. Questo ha a che fare con l'utilizzo delle nostre capacità di ragionamento spaziale, riconoscimento di schemi e visione d'insieme.
- Amplificare la cognizione riguarda il massimizzare quanto efficientemente ed efficacemente siamo capaci di processare l'informazione in idee, intuizioni e conoscenza.

Essenzialmente l'obiettivo della Data Visualization dovrebbe essere far sentire un lettore o un utente più informato su una materia.

Una delle motivazioni più convincenti sul valore della Data Visualization è espressa dall'affermazione di John W Tukey (*Exploratory Data Analysis*) [W04]:

Il più grande valore di un'immagine è quando ci forza a notare ciò che non ci saremmo mai aspettati di vedere.

La Data Visualization è impiegata in una moltitudine di situazioni. Limitandosi al solo ambito della medicina, una delle tante applicazioni è l'uso della Data Visualization per

rendere comprensibile al radiologo l'insieme di dati acquisiti attraverso una risonanza magnetica [C01].

Poiché la natura del dato da visualizzare cambia al variare dell'area applicativa, la scelta di quali tecniche di visualizzazione selezionare tra le disponibili e come applicarle al dato dipende strettamente dal contesto del dato. Ad esempio, nel caso sapessimo che esistono relazioni di ereditarietà tra gli elementi di un insieme dei dati, potremmo ricorrere ai grafi per evidenziare e visualizzare questa informazione [HMM00]. Nella visualizzazione d'insiemi di dati vasti e complessi, emergono problemi e questioni diverse. Visualizzare l'intero insieme di dati in un solo grafico potrebbe essere inutile agli utenti e, anzi, potrebbe causare problemi di performance, aumentare la complessità e, di conseguenza, diminuire la leggibilità del risultato. Inoltre potrebbe essere più efficace visualizzare alcuni elementi di un colore o di una forma diversa per aumentare la leggibilità del grafico visualizzato.

Diversi algoritmi di raffigurazione sono stati sviluppati per risolvere i problemi di visualizzazione d'insiemi vasti e complessi di dati [K02]. Alcuni di questi algoritmi si focalizzano sui criteri estetici da soddisfare perché la rappresentazione sia soddisfacente [C01]. Spesso la soluzione ottimale è l'applicazione di un misto di tecniche.

Sono state proposte diverse tassonomie per le applicazioni di Data Visualization.

Spence [S07] classifica le suddivide in applicazioni *scivis*, che trattano dati fisici, *infovis*, che considerano dati astratti che non hanno una posizione fisica, e *geovis* (*geovisualizzazioni*), che trattano delle realtà che sono tra i due tipi precedenti, nelle quali dati potenzialmente astratti sono rappresentati in un contesto spaziale.

La Geovisualizzazione è un ambito abbastanza vago che raggruppa le esplorazioni visive, l'analisi, la sintesi e la presentazione di dati geospaziali integrando varie discipline abbinano la cartografia con quelle della visualizzazione scientifica, dell'analisi delle immagini, dell'analisi visuale e della scienza dell'informazione geografica. [K03]

La più celebre visualizzazione eseguita è la mappa di Minard del 1861 'Carte figurative des pertes successives en hommes de l'Armée Française dans la campagne de Russie 1812–1813', detta anche 'Napoleon's March on Moscow', in cui è raccontata la fallimentare campagna di Mosca di Napoleone sintetizzata in una sola auto esplicativa immagine. [K03]

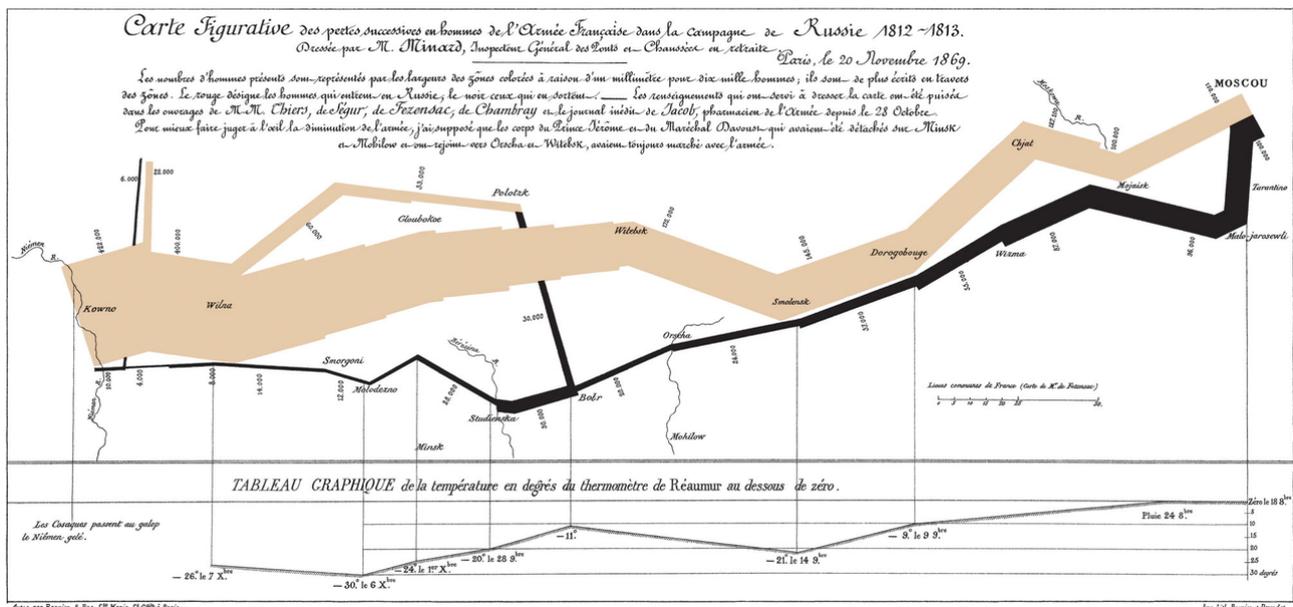


Figura 2 - Carte figurative des pertes successives en hommes de l'Armée Française dans la campagne de Russie 1812-1813

La tesi ricade nell'ambito della geovisualizzazione perché l'informazione numerica, che specifica l'anticipo o il ritardo, è associata a uno specifico tratto di strada che è un'informazione geografica. Vediamo quali tecniche della Data Visualization sono le più adeguate al nostro caso.

2.2 Visualizzazioni di valori scalari

Il ritardo e l'anticipo sono grandezze scalari. Sono dette grandezze scalari, quelle che, come per esempio la temperatura o il tempo, sono completamente descritte da un numero, che ne rappresenta il valore. Per definire univocamente una grandezza scalare è quindi sufficiente indicare un valore numerico accompagnato dalla relativa unità di misura (la durata di un intervallo di tempo è pari a 5 secondi, la temperatura di una stanza è di 20 °C ecc.). Il numero che definisce la misura di uno scalare è indicato con il termine di modulo, o più frequentemente intensità.

2.2.1 Color mapping

Il color mapping è il metodo di visualizzazione più diffuso per i valori scalari. In sostanza si tratta di associare un colore diverso per ogni valore scalare da rappresentare. Il color mapping applica una funzione $c: R \rightarrow Colors$ che assegna a quel punto un colore $c(s)$ □ $Colors$ che dipende dal valore scalare s in quel punto.

Esistono diversi modi per definire una tale funzione c .

Due tra queste funzioni sono le *color look-up tables* e le *transfer*.

Il primo tipo è il più semplice da implementare.

Una funzione *color look-up tables* C , detta *colormap*, è un campionamento uniforme della funzione c che mappa i colori:

$$C = \{c_i\} \quad i=1..N, \quad \text{dove } c_i = c\left(\frac{(N-i)f_{\min} + if_{\max}}{N}\right) \quad (2.1)$$

In pratica, l'equazione (2.1) è implementata come una tabella di N colori c_1, \dots, c_N che sono associati ai valori dell'insieme f di dati scalari che ha valori compresi nell'intervallo $[f_{\min}, f_{\max}]$. Conoscere l'intervallo scalare ci permette di costruire una mappa dei colori avente un chiaro e semplice significato: i colori c_i con bassi indici i rappresentano, nella mappa dei colori, valori vicini a f_{\min} , mentre colori nella mappa dei colori con indici i vicini a N rappresentano valori scalari grandi vicino a f_{\max} . Nella pratica, f_{\min} e f_{\max} possono essere entrambi determinati automaticamente esaminando campioni dell'insieme dei dati oppure possono essere indicati dall'utente. Nel secondo caso, i valori esterni all'intervallo $[f_{\min}, f_{\max}]$ sono ridotti a questo intervallo per ricondursi a valori validi della predefinita mappa dei colori. Abbinare i valori dell'insieme dei dati al loro intervallo è fattibile quando è possibile conoscere l'intervallo prima di realizzare la visualizzazione. S'immagini un'applicazione in cui si voglia visualizzare un valore scalare $f(t)$ dipendente dal tempo con $t \in [t_{\min}, t_{\max}]$. La tipica soluzione è di visualizzare dei valori seguendo la mappa dei colori dello scalare $f(t)$ per valori consecutivi di $t \in [t_{\min}, t_{\max}]$. Se non conosciamo $f(t)$ per ogni t prima di procedere alla visualizzazione, non potremo determinare l'intervallo assoluto dei valori $[f_{\min}, f_{\max}]$ per applicare l'equazione (1.1). In alcuni casi, anche quando questo è possibile, potrebbe non essere consigliabile. Se l'intervallo $[f_{\min}(t_i), f_{\max}(t_i)]$ di un certo istante $t_i \in [t_{\min}, t_{\max}]$ è molto più piccolo dell'intervallo assoluto $[f_{\min}, f_{\max}]$, normalizzare f all'intervallo assoluto porterebbe a visualizzare meno dettagli quando si osserva la singola inquadratura. In tali situazioni una migliore soluzione potrebbe essere normalizzare l'intervallo degli scalari per ogni singola inquadratura di $f(t)$. Naturalmente questo implica mostrare differenti leggende per ogni inquadratura.

Oltre alla funzione che mappa i valori in colori all'interno di una tabella predefinita, si potrebbe definire una funzione c analiticamente, se si preferisce. Poiché i colori sono di

solito rappresentati da una terna nel sistema di colori RGB o HSV, questo è fatto solitamente definendo tre funzioni scalari $c_R : R \rightarrow R$, $c_G : R \rightarrow R$, e $c_B : R \rightarrow R$ dove $c = (c_R, c_G, c_B)$. Le funzioni c_R, c_G, c_B sono anche dette funzioni di trasferimento.

Nella pratica, si utilizza funzioni *color look-up tables* quando non vi è la necessità di cambiare i colori durante la visualizzazione le funzioni *transfer* quando l'obiettivo della visualizzazione richiede una variazione dinamica delle funzioni che mappano i colori. [T14]

2.2.2 Scelta della mappa dei colori

La scelta della mappa dei colori dipende da qual è l'obiettivo dall'analisi. Alcuni obiettivi sono:

1. Valore assoluto, ad esempio rappresentare il valore assoluto di ognuno dei punti visualizzati.
2. Ordinamento dei valori, ad esempio dati due punti visualizzati qualsiasi, esprimere quale dei due ha il valore maggiore.
3. Differenza di valore, ad esempio dati due punti visualizzati qualsiasi, esprimere qual è la differenza tra i valori di questi due punti.
4. Valore selezionato, ad esempio dato un particolare valore $f_{interest}$, mostrare quali tra i punti visualizzati hanno valore $f_{interest}$. Una variante prevede di sostituire il singolo valore $f_{interest}$ con un intervallo compatto di valori.
5. Cambio di valore, ad esempio esprime la velocità della variazione del valore, o derivata prima, associato a un dato punto visualizzato.

Alcuni obiettivi sono più complicati da raggiungere degli altri.

L'obiettivo 2 richiede solo esprimere se un valore è maggiore (o minore) di un altro valore, ma non di quanto e quale sia il valore assoluto di entrambi. L'obiettivo 3 richiede di esprimere quanto due valori siano lontani tra loro, ma non quale sia il più grande e quale il più piccolo, ossia quale sia l'ordine tra i valori. L'obiettivo 4 richiede di esprimere dove nell'insieme dei dati appare un certo valore $f_{interest}$, ma non richiede di dire nulla o quasi nulla sui punti che hanno valori diversi da $f_{interest}$ (ad eccezione che non hanno quel valore). L'obiettivo 5 richiede di esprimere quanto velocemente cambino i valori in o attorno ad un dato punto ma nulla sul loro valore assoluto.

Progettare una mappa dei colori che raggiunga tutti gli obiettivi sopra indicati nella stessa misura, è un compito arduo. Per questo, la mappa dei colori è scelta focalizzandosi sull'ottimizzazione di un sottoinsieme degli obiettivi sopraindicati.

2.2.3 La legenda dei colori

Il primo obiettivo è il più complicato tra i cinque.

Per raggiungere quest'obiettivo dobbiamo essere in grado mentalmente di tradurre un colore in un valore. Per riuscirci, ricorriamo alla cosiddetta legenda dei colori, una striscia contenente i colori impiegati nella visualizzazione e delle etichette che indicano il valore associato a tutti o a una parte dei colori. Affinché ciò sia possibile, ogni valore deve essere associato a un solo colore. I colori devono essere distinguibili chiaramente dall'osservatore. Inoltre la risoluzione spaziale dell'insieme di dati visualizzati deve essere sufficientemente elevata come comparata alla velocità di variazione del dato scalare f che siamo capaci di distinguere visivamente regioni separate che hanno colori diversi.

Le legende dei colori sono necessarie in tutte le applicazioni di mappatura di colori in cui richiediamo di associare un colore a una quantità associata al dato.

La mappa dei colori è influenzata dal tipo di applicazione e dalle specifiche tradizioni o convenzioni.

Molto utilizzata è la mappa arcobaleno sebbene abbia alcuni importanti limiti che devono essere valutati e corretti rispetto alla visualizzazione che si vuole ottenere. (Figura 3)

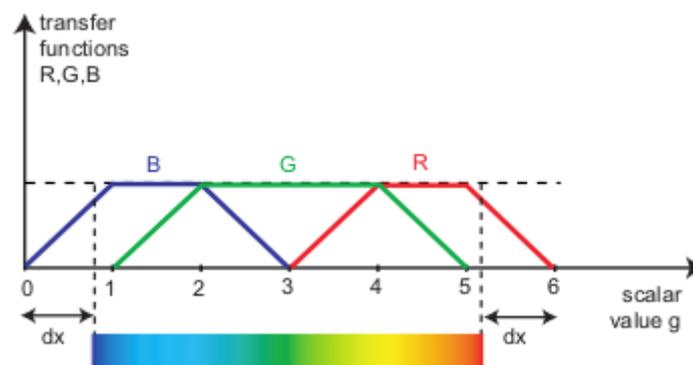


Figura 3 - Mappa arcobaleno

Altre scale possibili sono (Figura 4):

- greyscale;
- two-hue;
- heatmap;
- diverging;
- zebra.

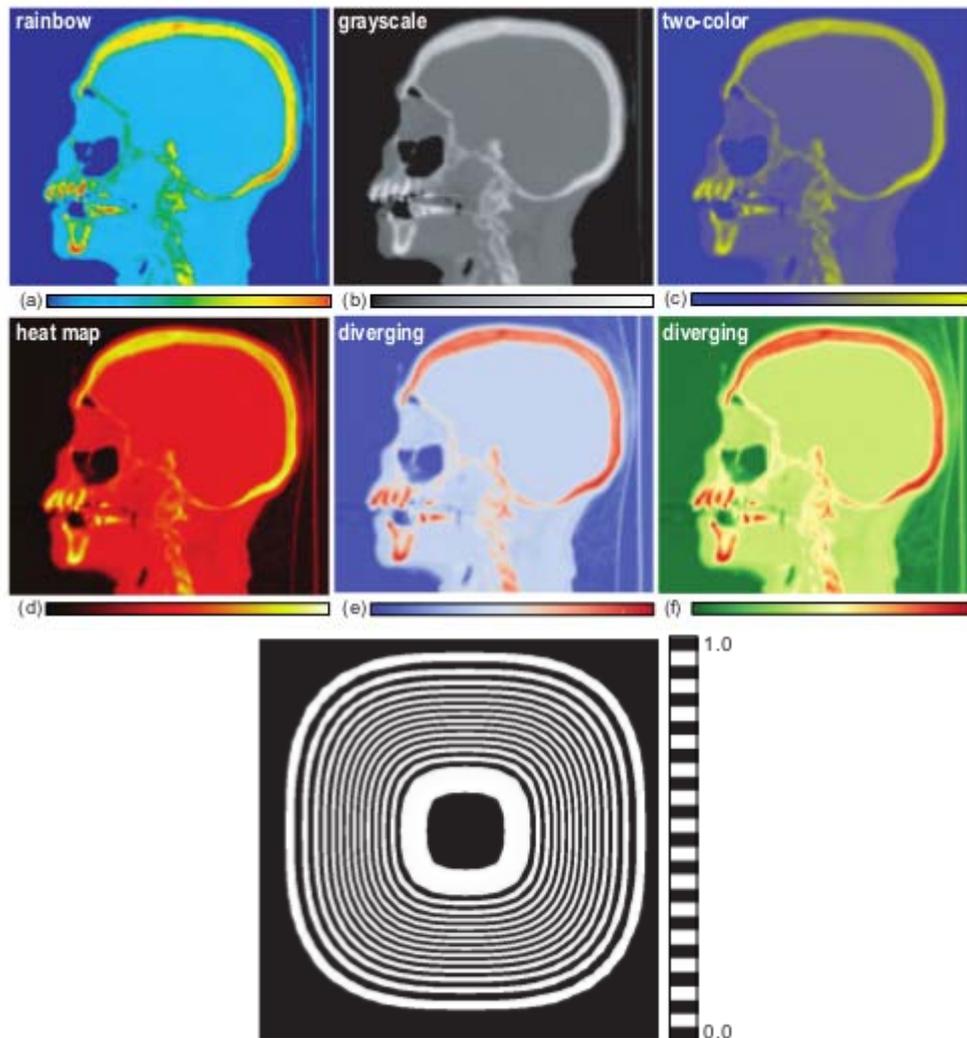


Figura 4 - Altre scale

Un importante aspetto da valutare nel definire la mappa dei colori è il numero massimo di colori N . Un valore piccolo di N porta all'effetto bandiera. Quando si visualizzano insiemi di dati continui, è un effetto indesiderato poiché crea degli artefatti visuali che non riflettono il reale andamento dei dati. Più velocemente il dato cambia spazialmente, più omogenea deve essere la mappa per evitare l'effetto bandiera. Il valore suggerito di N è tra 64 e 256, più grande è N meglio è.

Altri aspetti da considerare sono:

- la geometria, le aree uguali colorate diversamente possono essere percepite di dimensione distinta. Certe aree colorate densamente dello stesso colore possono sembrare
- il gruppo di utenti, persone diverse possono percepire diversamente la stessa mappa di colori (daltonismo)
- il medium, il tipo di medium su cui viene realizzata la visualizzazione può far percepire in modo diverso i colori
- contrasto, la mappatura dei colori descritta inizialmente genera colori attraverso funzioni colormap o transfer. Un altro insieme di tecniche consiste nel modificare i valori di grigio di una scala di grigi o dei colori di una immagine a colori al fine di enfatizzare i dettagli di interesse. [T14]

2.3 Rappresentazione topologica della rete di trasporto

Un servizio pubblico di trasporti garantisce la mobilità dei cittadini su di un territorio. Gli autobus circolano seguendo percorsi predefiniti, arrestandosi, all'orario indicato, alle fermate. Durante la fase di progettazione è impiegata una rappresentazione schematica del territorio e delle caratteristiche del servizio.

I grafi costituiscono un potente strumento di rappresentazione che può essere impiegato per descrivere realtà molto diverse. I nodi costituenti l'insieme N possono individuare i punti di un territorio, le diverse componenti di un sistema o le diverse attività di un viaggio. Un arco sta a indicare l'esistenza di una relazione di qualsiasi tipo fra la coppia di nodi che lo definisce. Ad esempio, se due nodi sono i punti di un territorio, un arco che li congiunge può rappresentare una strada che li unisce. Si osservi che un grafo costituisce una rappresentazione esclusivamente "topologica", essa consente unicamente di sapere se fra due qualunque elementi del sistema esiste la relazione che definisce gli archi, ma nessuna informazione quantitativa è associata a tale relazione.

Nell'ambito della programmazione del servizio, la rete del trasporto pubblico è descritta attraverso:

- un insieme S di assi stradali;
- un insieme F di punti di salita e/o discesa, detti fermate;

- un insieme N di punti significativi, detti nodi;
- un insieme A di assi stradali di S giustapposti che collegano un nodo B con un nodo C , detti archi.

L'asse stradale definisce il tracciato generalizzato di una strada. L'unione di tutti gli assi stradali fornisce la descrizione della viabilità disponibile. Ogni asse strada è caratterizzato da una lunghezza e da una rappresentazione grafica.

Ogni fermata è caratterizzata da un codice identificativo, da un nome e da una posizione geografica.

A una coppia di fermate (f_i, f_j) è associato un sottoinsieme di S contenente gli assi strada da percorrere per andare dalla fermata f_i alla fermata f_j . Indicheremo con $D(f_i, f_j)$ la somma delle lunghezze degli assi strada presenti in S .

Ogni nodo è caratterizzato da una sigla, da un nome e da una posizione geografica. I nodi si distinguono di transito e di sosta. Il nodo di transito individua un luogo in cui il bus può transitare o sostare brevemente come avviene nelle fermate di capolinea. Il nodo di sosta individua un luogo in cui il bus può essere lasciato in sosta per lunghi periodi e con l'assenza del conducente. Esempi tipici parcheggi o depositi.

Tra fermate e nodi esiste una relazione suriettiva: ogni fermata può essere nodo e ogni nodo deve essere fermata. Le coordinate del nodo sono le stesse della fermata associata al nodo. Può capitare che lo stesso nodo sia associato a fermate distinte.

Ogni arco BC dal nodo B al nodo C è caratterizzato da una lunghezza l e da uno o più tempi di percorrenza t . Un arco è detto di linea se è associato a un insieme F di fermate altrimenti parliamo di arco a vuoto. L'arco a vuoto rappresenta il percorso compiuto da autobus in uscita o in rientro da depositi o parcheggi.

Ogni fermata in F è caratterizzata dalla distanza sia assoluta sia percentuale dal nodo iniziale B . Se la cardinalità di F è k , avremo che la fermata f_0 , associata al nodo B , è a distanza 0, che la fermata f_k , associata al nodo C , è a distanza L , e la distanza di f_i è minore della distanza di f_{i+1} con $i=0,1,2,\dots,k-1$. La lunghezza l di un arco (B, C) è data dalla somma degli assi strada che collegano le coppie di fermate dell'arco seguendo la loro distanza

crescente dalla fermata f ossia $L_{(A,B)} = \sum_{i=0}^k D(f_i, f_{i+1})$.

Il tempo di percorrenza - o più semplicemente percorrenza - di un arco (B, C) è il tempo impiegato per percorrere gli assi strada che conducono da B a C . La percorrenza è espressa

in minuti e può variare durante la giornata o dell'anno per adattarsi alle condizioni della viabilità. Ad esempio la congestione stradale nel periodo invernale è maggiore rispetto al periodo estivo quindi percorrere lo stesso tratto di strada in inverno richiede maggior tempo rispetto all'estate.

Un percorso è una sequenza di archi giustapposti (B_i, C_i) di A con $i=0,1,2,3,..r$ in cui $C_i = B_{i+1}$ per ogni $i < r$.

Una linea della rete di trasporto è definita da un insieme dei percorsi. Il modo più semplice di rappresentare una rete di trasporti, evidenziando la posizione delle fermate o le strade percorse dalle linee, è quello di disegnarla su una mappa. Nel mondo digitale si ricorre alle mappe digitali.

2.4 Cartografia digitale

Dalla cartografia tradizionale realizzata e pensata su supporto cartaceo, negli ultimi decenni, si è passati, attraverso diverse fasi evolutive, alla cartografia digitale, cioè basata su supporti informatizzati.

L'utilizzo dell'informatica e la parallela diffusione dei Sistemi informativi geografici hanno determinato una progressiva evoluzione della cartografia digitale da semplice "disegno" del territorio a "dato numerico" utile per il disegno automatico e a "informazione geografica" (database geografico) utile per attività di analisi spaziale.

2.4.1 Evoluzione della cartografia digitale

L'evoluzione della cartografia su supporto informatizzato ha attraversato diverse tappe (identificate con denominazioni non sempre univoche):

- cartografia automatizzata, basata sull'utilizzo di strumenti automatici (elaboratori + software CAD + plotter) per la produzione di cartografia di tipo tradizionale;
- cartografia numerica, cioè cartografia gestita attraverso strumenti informatici e basata sulla rappresentazione (strutturata in livelli) degli oggetti del mondo reale corredata dagli attributi che li definiscono;
- database geografico (topografico), cioè un "contenitore" di dati geografici, organizzati secondo uno schema predefinito, adatto a essere consultato, analizzato, elaborato per ogni possibile applicazione.

Con altra terminologia, oggi forse un po' meno utilizzata rispetto a qualche anno fa, si può distinguere tra cartografia Map oriented e cartografia GIS oriented.

L'utilizzo di applicazioni informatiche, oltre a semplificare il lavoro di produzione della cartografia e ad aumentarne notevolmente il potenziale conoscitivo, ha imposto nuovi modelli d'interpretazione (e descrizione) del mondo reale:

- nella cartografia tradizionale, è l'utente che interpretando i simboli grafici deduce l'informazione geografica (contenuto associativo) che la cartografia veicola, attraverso un processo mentale detto sintesi associativa;
- nella cartografia numerica, dove l'utilizzo di un elaboratore rende inapplicabile la sintesi associativa, il contenuto informativo della cartografia è suddiviso in due parti distinte:
 1. contenuto cartografico, cioè la rappresentazione degli elementi del mondo reale;
 2. contenuto strutturale, cioè l'informazione geografica veicolata dalla cartografia.

2.4.2 Primitive della cartografia digitale

Nella cartografia digitale, gli oggetti del mondo reale possono essere descritti attraverso le primitive geometriche fondamentali, come punti, linee, poligoni e volumi in relazione alle caratteristiche spaziali degli oggetti reali, ma anche in funzione delle finalità (e della scala) del modello dati.

Tabella 1 - Primitive della cartografia digitale

Punti - 0D	<ul style="list-style-type: none"> • privi di dimensioni (a una data scala) • oggetti troppo piccoli perché siano rappresentati come aree (o linee) 	pozzi, impianti industriali, città, scuole
Linee/Archi – 1D	<ul style="list-style-type: none"> • solo lunghezza • oggetti troppo “stretti” per essere rappresentati come aree 	strade, fiumi, linee elettriche
Aree/Poligoni – 2D	<ul style="list-style-type: none"> • comprendono un'area che è delimitata dai suoi confini • oggetti dotati di area 	laghi, entità amministrative, edifici
Volumi/superfici tridimensionali – 3D	<ul style="list-style-type: none"> • descrivono una superficie tridimensionale • Oggetti dotati di volume 	superficie topografica, edifici

2.4.3 Forme di rappresentazione

La rappresentazione degli oggetti del mondo reale può essere basata su diverse strutture dati, che si adattano alle caratteristiche dei dati e corrispondono alle modalità con le quali i dati sono archiviati e organizzati.

Le due principali forme di rappresentazione sono:

- vector, basata sulle primitive geometriche (punto, linea, poligono)
- raster, articolata come griglia di punti.

2.4.4 Vector

Nella struttura vettoriale (vector) gli oggetti geografici (secondo la natura dei dati spaziali che rappresentano) sono rappresentati da primitive geometriche, punti, linee o poligoni, le quali contengono in sé un riferimento spaziale in base ad un sistema di coordinate; alle geometrie sono in genere associati degli attributi, archiviati in un database relazionale.

Gli oggetti vettoriali elementari sono:

- punti, individuati da singole coppie di coordinate x,y;
- linee, individuate da sequenze ordinate di coppie coordinate x,y;
- poligoni, individuati da sequenze chiuse di coppie di coordinate x,y;

Dalle strutture vettoriali elementari derivano strutture complesse impiegate per la rappresentazione di caratteristiche specifiche: vettoriali con misure, vettoriali multi patch (solidi in 3D), TIN.

Nella struttura vettoriale agli oggetti geografici sono generalmente associate nelle informazioni di tipo non geometrico che ne specificano meglio le caratteristiche (attributi).

Possono essere di vari tipi:

- Testo (es. toponimi, nomi e cognomi dei proprietari di particelle catastali)
- Numerici (es. superficie, popolazione, portata di un fiume)
- Boolean (vero/falso)

2.4.5 Raster

Nella struttura raster (*Figura 5*), propria delle immagini digitali, l'informazione è rappresentata come una griglia (matrice) rettangolare, ciascun elemento della quale è detto pixel (Picture Element).

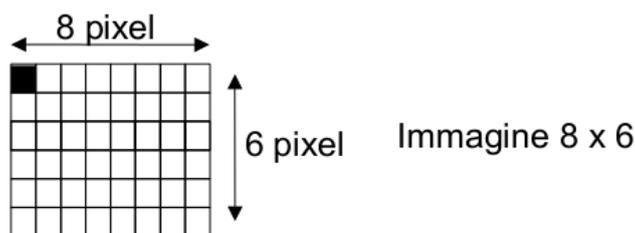


Figura 5 - Raster

A ciascun pixel dell'immagine è associato un valore d'intensità, che può rappresentare “soltanto” in colore nel caso d'immagini di cartografie, la risposta degli oggetti alla radiazione solare (fotografia, immagine da satellite) oppure la variazione di una caratteristica fisica (es. altimetria) o di un fenomeno (es. temperatura).

Dimensione del pixel e numero di livelli di grigio determinano:

- il più piccolo elemento rappresentabile, cioè risoluzione spaziale (in dpi o in unità di misura delle lunghezze);
- la più piccola variazione d'intensità rappresentabile (bitmap, toni di grigio, true color) o risoluzione radiometrica.

La posizione spaziale di una cella, come in ogni immagine digitale, è data dal numero di riga e di colonna (in relazione ad una determinata origine), la cosiddetta “coordinata immagine”. La georeferenziazione con il territorio si ottiene attribuendo alle celle il valore di una coppia di coordinate geografiche o piane (di solito a un particolare pixel dell'immagine, come quello in alto a sinistra).

Esistono tre tipologie di file raster, in base al significato del valore associato alle celle:

- immagini digitali, come ortofotografie, dati da telerilevamento, in cui l'informazione associata al pixel è un valore radiometrico;
- immagini cartografiche, immagini di carte, in cui l'informazione è veicolata sulla base dei colori associati ai pixel;
- raster tematici (o matrix), immagini in cui il valore associato al pixel è una grandezza che esprime un fenomeno fisico (precipitazioni, temperatura, etc) oppure un codice frutto di un'operazione di classificazione (es. uso del suolo).

2.4.6 Cartografia digitale nell'ambito del trasporto

Le cartografie digitali sono impiegate per rappresentare diversi tipi d'informazioni. Sono create e mantenute da molti soggetti pubblici o privati. Fra i più completi rilevamenti del

territorio italiano rientrano le carte redatte della Marina Militare e il sito del Portale Cartografico Nazionale.

Creare una cartografia stradale GPS è un compito assai lungo e difficile, da proseguire nel tempo, per questo, storicamente, ogni casa che produce software di navigazione GPS si affidava a dei fornitori di cartografie, che si occupavano di creare le cartografie digitali da concedere poi in licenza per l'uso con i software GPS, per i siti di mappe online, per usi aziendali, Questo scenario, però, è cambiato negli ultimi anni, man mano che la cartografia digitale acquistava importanza presso la clientela (anche aziendale) e assumeva un ruolo non più marginale anche nei dispositivi mobili di ultima generazione come smartphone e tablet.

I principali provider di mappe digitali sono due: Teleatlas e Navteq. Nel giugno del 2008 Teleatlas è stata acquisita da TomTom. Nel 2012 Navteq è confluita nel nuovo marchio di Nokia chiamato HERE.

Di là da piccole realtà “locali” alternative in zone poco sviluppate (a livello stradale), il mondo GPS si divideva tra questi due grandi fornitori, cui ogni produttore attingeva per i dati cartografici, troppo complicati da produrre in autonomia.

Le loro mappe sono fonte di dati cartografici per moltissime realtà quali navigatori GPS di varie marche, sistemi aziendali, da Google Maps a Mappe di Apple, passando per IGo e Microsoft, i marchi forniti sono veramente tanti.

La più valida alternativa gratuita a NavTeq e TeleAtlas è OpenStreetMap (OSM). E' un progetto collaborativo finalizzato a creare mappe a contenuto libero del mondo. Il progetto punta a una raccolta mondiale di dati geografici, con scopo principale della creazione di mappe e cartografie.

2.5 GIS e WebGIS

La migrazione dalla cartografia tradizionale alla cartografia digitale ha portato alla realizzazione degli strumenti necessari alla consultazione e alla manutenzione della cartografia digitale.

2.5.1 Geographic Information System

Un Geographic Information System (GIS) è un sistema progettato per ricevere, immagazzinare, elaborare, analizzare, gestire e rappresentare dati di tipo geografico. L'acronimo GIS è spesso usato per significare la scienza o gli studi sulle informazioni geografiche (dette anche geospaziali); gli studi sulle informazioni geospaziali si riferiscono a discipline accademiche o professioni che usano i GIS.

In termini semplici, col GIS si possono unire cartografie, eseguire analisi statistiche e gestire i dati attraverso tecnologie database.

L'innovazione più rilevante contenuta nei GIS è la loro capacità di gestire dati spazialmente suddividendoli in livelli e di sovrapporre questi livelli per eseguire analisi spaziali.

Dagli anni 70 i GIS iniziano a essere impiegati nell'analisi geografica ed economica nel settore sia pubblico sia privato.

Dai primi anni 90 i trasporti sono stati uno dei settori in cui cresce di più l'impiego dei GIS (T-GIS o GIS-T). In quel periodo sono aggiunte le specifiche funzioni per la gestione delle informazioni sul trasporto. Le agenzie di trasporto possono georeferenziare le fermate, i percorsi dei bus e altre informazioni associandole alla rappresentazione astratta delle strade. Negli stessi anni nascono aziende specializzate nel mondo GIS quali ESRI e MapInfo. La crescente popolarità dei GIS spinge i venditori di software specialistico per il trasporto a sviluppare per i propri prodotti delle interfacce aventi funzionalità GIS. Uno dei risultati è il problema d'incompatibilità che si presenta nel tentativo di utilizzare prodotti provenienti da fornitori distinti.

Agli inizi del 2000 i produttori di GIS si spostano verso standard più aperti e in molti aderiscono all'Open GIS Consortium (OGC) che promuove standard Open-GIS quali web feature standard (WFS) e map feature standard (MFS). Lo sviluppo dei web-services combinato a WFS e MFS ha portato allo sviluppo del WebGIS. [S05]

2.5.2 WebGIS

Le applicazioni webGIS permettono la distribuzione di dati geo-spaziali, in reti internet e intranet, sfruttando le analisi derivanti dai software GIS e per mezzo di classiche funzionalità di applicazioni web-based pubblicano informazioni geografiche nel World Wide Web. Un sistema WebGIS si basa su normali funzionalità Client - server, come una

classica architettura Web. Il client è un qualsiasi browser, come ad esempio Mozilla Firefox, il lato server consiste in Web - server (ad esempio Apache) e un software WebGIS (ad esempio UMN Mapserver) che si occupa di fornire le funzionalità di visualizzazione/interrogazione per l'interpretazione di dati georeferenziati.

Possiamo schematizzare il processo di funzionamento di un WebGIS nel seguente modo:

- l'utente, tramite un'apposita interfaccia web, invia dal proprio client mediante un generico browser una richiesta nella quale è definita l'area d'interesse (estensione geografica della zona da visualizzare) e i contenuti a cui si vuole accedere (elenco dei layer d'interesse);
- in base ai dati ricevuti il motore del WebGIS accede ricerca nel proprio archivio le informazioni richieste (file, come shape, TIFF, ...; compiendo una connessione a database, oppure mediante l'accesso ad altri server cartografici, ad esempio OGC Web Services) estraendo la porzione di territorio specificata;
- sono quindi generate una o più immagini che saranno inviate al client.

In modo simile avviene l'interrogazione dei dati: in questo caso il client invia al server una coppia di coordinate (nel caso d'interrogazione puntuale) o un insieme di coordinate per definire una regione e il server interroga gli attributi degli oggetti cartografici presenti nella zona d'interesse, restituendo un report con i valori estratti.

2.5.3 Open Geospatial Consortium

L'Open Geospatial Consortium (OGC) si occupa della definizione degli standard per i GeoServizi (OpenGIS Web Service - OWS).

Lo scambio dei dati si basa sul protocollo HTTP, utilizzando un web service si permette l'interazione tra servizi web e applicazioni mediante l'invio di messaggi in formato XML (eXtensible Markup Language). L'intero sistema di comunicazione è indipendente dalla piattaforma hardware, dal sistema operativo e dal formato originario dei dati: qualsiasi software client e server possono comunicare tra di loro purché implementino in modo corretto gli standard.

Possiamo descrivere il funzionamento come segue:

- il client contatta il server e lo interroga sulle sue funzionalità;

- il server fornisce al client un documento XML contenente la descrizione del servizio offerto, dal punto di vista delle funzionalità e dei dati disponibili;
- utilizzando come riferimento il documento di descrizione inviato dal server, il client genera la propria richiesta;
- il server fornisce i dati in un formato standard (GML o immagini) seguendo le richieste inoltrate dal client.

2.6 Data Visualization nel Trasporto Pubblico

La Data Visualization è impiegata ampiamente nell'ambito del trasporto. Negli Stati Uniti esiste un comitato stabile specifico, *Committee on Visualization in Transportation (ABJ95)*, inquadrato nella *Data and Information Technology*, all'interno del *Transportation Research Board (TRB)*. Lo scopo del comitato è di favorire e disseminare scambi e ricerche che migliorino e promuovano le conoscenze disponibili nonché l'applicazione delle metodiche e delle tecnologie di visualizzazione per il loro potenziale di rivolgersi alle attuali criticità del trasporto o promuovendo approcci innovativi per la soddisfazione delle necessità della futura società dei trasporti.

Il problema dell'analisi e della visualizzazione del buon funzionamento di una rete di trasporti è sentito tanto che Xerox Research Centre Europe ha organizzato un corso specifico, della durata di 6 mesi, intitolato "Public Transport Data Visualization" con inizio febbraio 2013.

In parecchi studi, i rilevamenti effettuati dai sistemi AVL sono impiegati per elaborare e distribuire informazioni, in tempo reale, agli utenti del trasporto pubblico, su tutti i media possibili.

Esistono soluzioni che interessano singole città come Calgary, Canada [H08] e altre che mostrano una collezione di realtà del trasporto.

Il team di sviluppo della geOps, in collaborazione con l'università di Friburgo, ha sviluppato il progetto TRAVIC. TRAVIC sta per Transit Visualization Client, e, il suo obiettivo è di visualizzare i dati di trasporto archiviati secondo le specifiche "General Transit Feed Specification – GTFS", il formato richiesto da Google alle agenzie di trasporto pubblico per apparire sul servizio Google Transit.

La copertura di Google Transit è diventata quasi globale ma ancora poche sono le agenzie di trasporto pubblico che rilasciano i dati per altri scopi.

Il team di sviluppo di TRAVIC ha deciso di utilizzare quello presente in questa pagina wiki - <https://code.google.com/p/googletransitdatafeed/wiki/PublicFeeds>, ottenendo quindi 202 sorgenti dati.

L'applicazione mostra, secondo le tabelle, la posizione prevista da un mezzo di trasporto pubblico nello stesso orario con cui si sta accedendo al servizio.

Un progetto interessante è Visual TPL [PMAR13], un'applicazione web creata per coadiuvare gli esperti nella pianificazione e nella valutazione del trasporto pubblico con l'obiettivo di individuare i punti critici che potrebbero essere studiati al fine di ottenere una loro ottimizzazione e ridurre l'onere sulla spesa pubblica. E' impiegato nella Regione Piemonte. Sfrutta la tecnologia Google Fusion Table (2012a) e Google Chart e la cartografia di Google Maps. Il risultato è un'interfaccia utente divisa in tre zone:

- la superiore consente di formulare la query da sottoporre al sistema;
- la centrale rappresenta su una mappa navigabile l'esito dell'elaborazione delle query;
- la inferiore contiene la tabella dei risultati ed estende le informazioni mostrate sulla mappa.

I dati possono essere filtrati sulla base della seguente lista di indicatori:

- contratto di servizio;
- linea;
- singolo percorso (una linea può averne di diversi);
- operatore del servizio;
- tipologia del percorso (ad esempio percorsi classificati in ordinario, servizio scolastico, servizio per lavoratori, mercatali, turistiche, turistiche a lunga percorrenza e atipico);
- carico medio;
- carico medio dal lunedì al venerdì durante il periodo invernale (da settembre a giugno);
- bus-km/anno;
- numero di corse nei giorni feriali invernali (da settembre a giugno) ossia del periodo di maggior servizio.

Inoltre è possibile effettuare operazioni sui dati filtrati tipo:

- ricercare di località o indirizzo;
- tematizzare la mappa rappresentando una certa informazione;
- tematizzare la tabella rappresentando una certa informazione;
- conoscere il numero di elementi selezionati, compresi quelli non visibili sulla mappa;
- conoscere il numero di elementi visualizzati;
- conoscere il numero totale di elementi.

2.7 Visualizzazioni di anticipi e ritardi nel trasporto pubblico

Uno studio è rivolto al supporto alle agenzie di trasporto per le loro decisioni [K07]. L'analisi impiega dei dati sono forniti da TriMet, l'azienda di trasporti regionali di Portland, OR, raccolti attraverso i sistemi AVL e APS a bordo dei propri mezzi e sono inviati all'agenzia pubblica deputata al controllo sul trasporto. È proposto un esame della letteratura dedicata alla rappresentazione delle informazioni e un esempio dell'applicazione delle tecniche di Data Visualization impiegate su strumenti GIS. Essendo uno studio datato non descrive soluzioni che utilizzino il mondo del web.

Un secondo studio considera le rilevazioni degli AVL dei mezzi impiegati nel sistema di trasporto di Melbourne e analizza il ritardo accumulato nell'effettuazione delle corse dalle linee che collegano il centro della città con la periferia nello stesso periodo dell'anno per 10 anni (2001 - 2011). [MCLN12]

L'analisi è espressa in una serie di visualizzazioni, ma non è prevista interazione per l'operatore.

Vediamo un paio di progetti che consentono un'interazione tra operatore e visualizzazione.

2.7.1 MetroViz

Il primo progetto è MetroViz [DBPEMS15] realizzato da CATT e University of Maryland – College Park. Partendo dal rilevamento triennale (2010 – 2013) della posizione dei veicoli e dal conteggio dei passeggeri trasportati sulle linee di Blacksburg, Virginia, hanno realizzato un sistema di visualizzazione che presenta contemporaneamente informazioni quali il numero passeggeri saliti a una certa fermata e il tipo di tariffa da loro pagata, l'elenco delle fermate per ogni linea e una vista del calendario che fornisce il dettaglio dell'andamento quotidiano a livello di fermata o di percorso. Gli utenti possono esplorare i dati che si

riferiscono ai carichi e all'aderenza all'orario a livello sia di fermata sia di percorso. Sono presenti delle sottoviste per fermata e per corsa per esplorare il dettaglio orario, come estensione della vista calendario. E' un'applicazione web che sfrutta i servizi di GoogleMaps per rappresentare la posizione delle fermate e il percorso dei bus su una mappa navigabile.

2.7.2 Web Portal for Reliability Diagnosis of Bus Regularity

Il secondo progetto è Web Portal for Reliability Diagnosis of Bus Regularity [BCDLMPT13]. Si concentra sui servizi ad alta frequenza per i quali la regolarità, ossia il rispetto della frequenza dei bus alle fermate, può essere usata come indicatore della qualità del servizio, in termini di affidabilità, sia dai viaggiatori sia dalle agenzie del trasporto. Il risultato dell'analisi è visualizzato da un'applicazione web che mostra diverse informazioni, tra le quali:

- la tabella della regolarità, una tabella che mostra il livello di servizio.

L'operatore può scegliere la direzione del percorso e l'intervallo di tempo. La tabella mostra per ogni fermata il livello di servizio della regolarità espresso in colori diversi.

- Il grafico a linea della regolarità – un grafico multi linea una linea per ogni intervallo di tempo.

Ogni linea ha un colore distinto e può essere immediatamente comparata con le altre.

- Il mapper – un'interfaccia basata sulla tecnologia di Google Maps.

La mappa mostra il cammino del percorso nella direzione selezionata. Le fermate associate al percorso sono visualizzate con colori diversi secondo il livello di servizio della regolarità. L'operatore può interagire con la mappa modificando la fascia oraria all'interno del periodo di tempo.

2.8 Considerazioni sulle soluzioni in letteratura

Gli esempi recuperati in letteratura sono per lo più concentrati sulla visualizzazione d'informazioni in tempo reale. Sono pochi gli esempi di visualizzazione di anticipi e ritardi nel trasporto pubblico.

Le due soluzioni raccontate nel paragrafo precedente consentono di visualizzare una sola linea alla volta. Inoltre l'analisi del dato non prende in considerazione né le condizioni meteorologiche né eventuali problemi alla viabilità. E' noto che, in caso di pioggia, il

traffico aumenta la sua congestione e che la presenza d'incidenti o manifestazioni può portare al rallentamento della velocità con cui i bus si muovono. L'inclusione, nell'analisi degli anticipi/ritardi medi, di tali situazioni anomale potrebbe alterare l'attendibilità del risultato.

3. GARTP: GEOVISUALIZZAZIONE DI ANTICIPI E RITARDI DEL TRASPORTO PUBBLICO

3.1 *Che cosa fa GARTP*

GARTP analizza la differenza tra percorrenze programmate e percorrenze reali degli archi di una rete di trasporti e la presenta su una base cartografica. E' uno strumento specialistico, dedicato agli esperti della progettazione del servizio pubblico.

Il sistema impiega gli orari dei transiti sui nodi della rete, sia programmati sia rilevati, degli autobus impiegati nel trasporto pubblico di linea nella provincia di Bologna. I transiti rilevati e i transiti programmati sono utilizzati per calcolare la differenza tra il tempo di percorrenza reale e il tempo di percorrenza programmato dei singoli archi della rete. Descriviamo brevemente il processo.

3.1.1 **Dati in GARTP**

Ogni corsa rilevata c_r è associata biunivocamente a una corsa programmata c_p che definisce i transiti teorici previsti nella progettazione del servizio su una certa rete di trasporto. Perciò è possibile per ogni corsa rilevata c_r selezionare sia i suoi transiti reali sia i suoi transiti programmati, leggendoli dalla corsa c_p . Poiché i nodi della rete hanno nomi univoci, possiamo associare ciascun nodo di transito e orario rilevato con un solo nodo e orario di transito teorico. Se riordiniamo la successione dei nodi e degli orari di transiti rispetto agli orari programmati e consideriamo due nodi successivi A e B , avremo che (AB) è un arco della rete e la differenza tra gli orari di transito in B e in A sarà la percorrenza t di quell'arco. Se la differenza è effettuata tra gli orari rilevati, avremo una percorrenza rilevata. Se la differenza è effettuata tra gli orari programmati, avremo una percorrenza programmata. Ripetendo lo stesso processo per ogni corsa rilevata avremo l'insieme di percorrenze che utilizza GARTP.

La rete di trasporti pubblici analizzata è costituita da 1184 nodi collegati tra loro da 3307 archi. L'estensione della rete, cioè la somma delle lunghezze di tutti gli archi, è di oltre 10500 Km. Le fermate sul territorio sono oltre 7000 e le linee oltre 230.

Le informazioni interessano il periodo compreso tra l'1 gennaio 2011 e il 31 maggio del 2014, durante il quale sono state accumulate oltre 38 milioni di registrazioni impiegabili per la visualizzazione.

GARTP utilizza altri dati esterni alla descrizione e alla rilevazione della rete di trasporti nel periodo compreso tra l'1 gennaio 2011 e il 31 maggio del 2014. Raccoglie le informazioni meteorologiche (temperatura, umidità, fenomeni meteo in corso, intensità e direzione del vento) della città di Bologna. Incorpora le informazioni sulle fiere tenutesi nel Fiera District, sui concerti tenutisi allo stadio o nei palazzi dello sport, sui maggiori eventi calcistici tenutisi allo stadio, sugli incidenti stradali (viabilità ordinaria e autostradale) riportati sulla stampa cittadina di Bologna. Contiene l'elenco e le caratteristiche dei mezzi impiegati per l'effettuazione delle corse telerilevate.

Le informazioni contenute in GARTP sono incrociate tra loro per ottenere una visualizzazione dell'andamento storico delle percorrenze.

3.2 A cosa serve GARTP

GARTP permette a un operatore di analizzare l'adeguatezza delle percorrenze di una o più linee del servizio di trasporto pubblico nella provincia di Bologna. L'operatore può selezionare contemporaneamente i percorsi di una o più linee. Come risposta vedrà visualizzati gli anticipi o i ritardi medi accumulati dagli autobus nell'esercire corse sui percorsi selezionati rappresentati attraverso la colorazione adeguata dei tratti di strada che compongono i percorsi disegnati sulla cartografia offerta da OpenStreetMap. I dati possono essere filtrati specificando uno o più dei seguenti parametri:

1. anno;
2. mese o intervallo di mesi;
3. fascia oraria;
4. presenza di eventi o d'incidenti di rilievo giornalistico;
5. condizioni meteorologiche;
6. tipo di alimentazione dell'autobus impiegato.

L'operatore può ottenere informazioni supplementari ricorrendo alla funzione identifica. Una volta attivata tale funzione, cliccando su un qualunque oggetto sulla mappa sarà restituita la collezione d'informazioni collegate a quell'oggetto, se ve ne sono.

3.3 Descrizione dell'interfaccia di GARTP

L'accesso a GARTP è effettuato mediante un qualunque browser che supporti la tecnologia javascript collegandosi al server che eroga il servizio. Per risposta è presentata la seguente pagina web. Dopo qualche secondo, sarà presentata la schermata in *Figura 7*

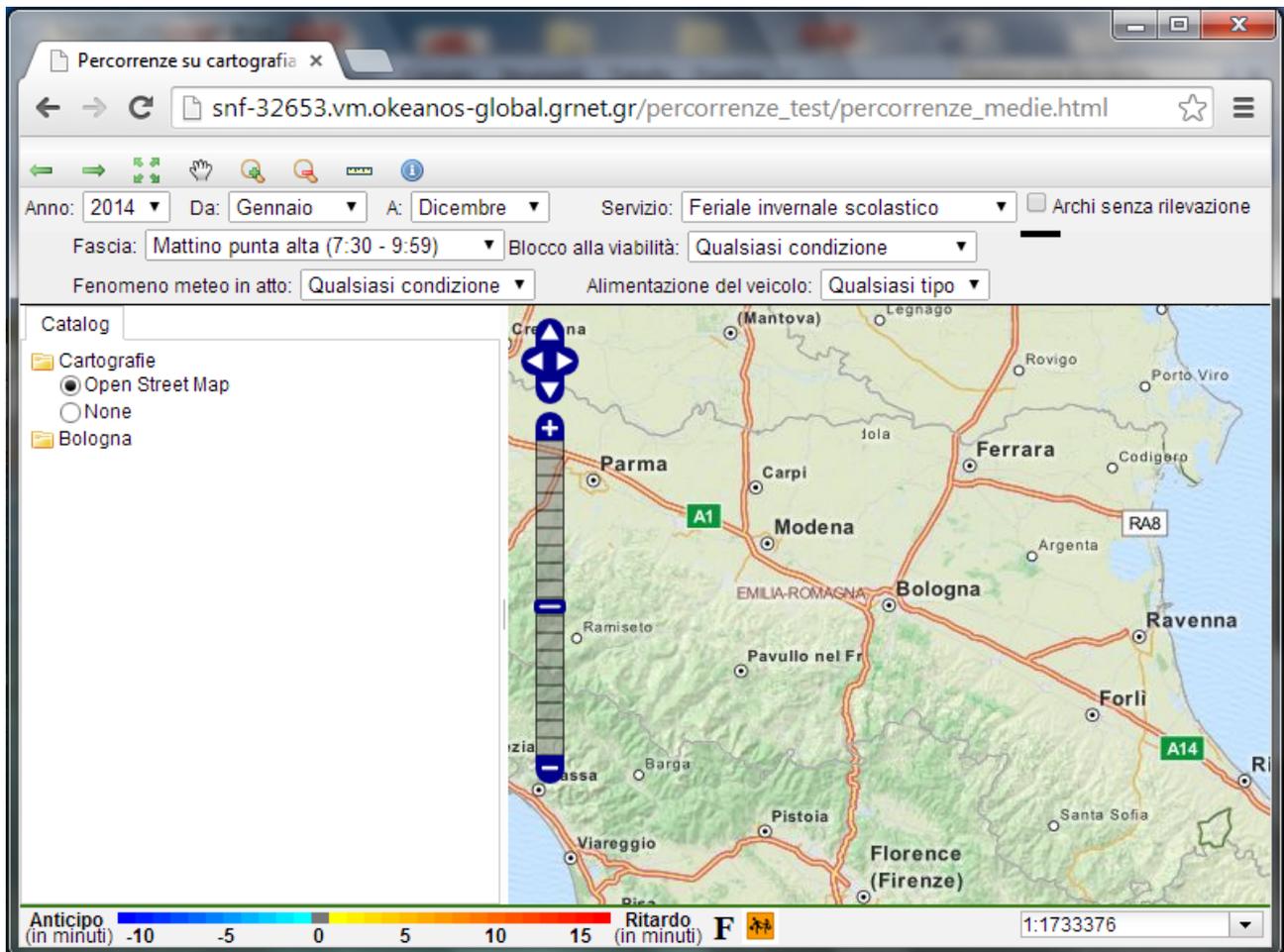


Figura 6 - Schermata di avvio del GARTP.

All'interno della schermata possiamo distinguere cinque zone:

- i comandi di navigazione;
- i parametri di filtro dei dati;
- l'elenco delle linee di cui sono disponibili i rilevamenti;
- la mappa;
- la barra di stato.

I comandi di navigazione attivano la cache della navigazione, lo zoom alla posizione iniziale, la modalità di spostamento del quadro, l'incremento e la diminuzione di zoom, le funzioni di misura e d'identificazione.



Figura 7 - Comandi di navigazione

Sotto i comandi di navigazione troviamo i parametri di filtro dati. L'operatore può indicare l'anno, l'intervallo di mesi, il tipo di servizio, la fascia oraria, il tipo dell'eventuale blocco alla viabilità, la particolare condizioni meteo e il tipo di alimentazione dei mezzi impiegati. L'intervallo dei mesi è inteso l'insieme dei giorni compreso tra il primo del mese indicato nella casella "Da" e l'ultimo del mese indicato nella casella "A". L'intervallo è relativo all'anno indicato nella casella "Anno". Inoltre può esprimere la volontà di vedere rappresentati, sulla mappa, tutti gli archi percorsi da una linea. Se il checkbox "Archi senza rilevazione" non è spuntato, saranno disegnati solo gli archi che appaiono nell'insieme dei dati filtrati.

Anno: 2014 ▼	Da: Gennaio ▼	A: Dicembre ▼	Servizio: Feriale invernale scolastico ▼	<input type="checkbox"/> Archi senza rilevazione
Fascia: Mattino punta alta (7:30 - 9:59) ▼		Blocco alla viabilità: Sciopero bus ▼		
Fenomeno meteo in atto: Pioggia ▼		Alimentazione del veicolo: Diesel ▼		

Figura 8 - Parametri di filtro dati

Sotto i parametri di filtro dati, la parte sinistra può contenere più tab. Inizialmente ne è presente solo uno che prende il nome di Catalog e contiene gli elenchi, strutturati ad albero, delle collezioni di layer interrogabili e visualizzabili sulla mappa. I cataloghi presenti sono Cartografie e Bologna. Cartografie elenca le cartografie visualizzabili sulla mappa. Al momento è disponibile solo la cartografia offerta da OpenStreetMap, visualizzata in automatico all'avvio. E' possibile disattivarla cliccando sul radio button None. Il catalogo Bologna, strutturato in altre sottocartelle, contiene le linee del trasporto pubblico esercito nella provincia di Bologna per le quali è disponibile almeno una rilevazione nel periodo 2011 - 2014. Le linee extraurbane sono raggruppate per direttrice o zona servita. Le linee

urbane sono divise per comune: Bologna, Casalecchio, Imola. La cartella linee speciali contiene le linee adibite a particolari servizi eserciti in alcuni giorni dell'anno. La cartella riservati contiene linee dedicate a particolari aziende.

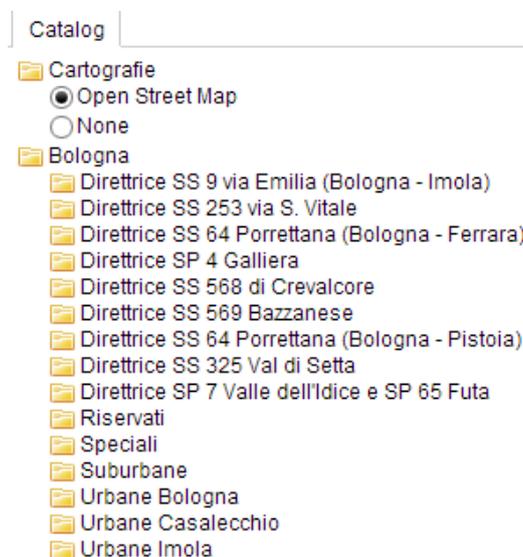


Figura 9 - Catalog

Sotto i parametri di filtro dati, la parte destra è dedicata a contenere la mappa sulla quale saranno rappresentate le visualizzazioni dei dati che soddisfano i requisiti indicati nella sezione filtro dati. I comandi navigazione operano sulla mappa. Una volta attivato un comando navigazione, il cursore potrebbe cambiare forma quando si trova nella zona mappa.



Figura 10 - Zona mappa

In fondo alla schermata sono presenti la legenda che consente l'interpretazione del risultato dell'interrogazione, la legenda che ripete graficamente alcuni il valore corrente dei parametri filtro - servizio, blocco alla viabilità, fenomeno meteo in corso e alimentazione del veicolo -, la posizione del cursore sulla mappa secondo il sistema WGS84 e la scala della mappa. (Figura 11)



Figura 11 - Barra di stato

All'avvio i parametri di filtro hanno il seguente valore predefinito: “Anno” è 2014, “Da” è Gennaio, e “A” è Dicembre, “Servizio” è Feriale invernale scolastico, “Fascia” è Mattino punta alta (7:30 – 09:59), “Blocco alla viabilità” è Qualsiasi condizione, “Fenomeno meteo in corso” è Qualsiasi condizione, “Alimentazione del veicolo” è Qualsiasi tipo.

I valori di default dei parametri di filtro sono scelti sulla base dell'esperienza. Le analisi sono condotte più frequentemente sui dati dell'ultimo anno disponibile e sul servizio più "forte", cioè con maggior numero di corse, erogato durante l'anno: il feriale invernale scolastico.

3.3.1 Visualizzazione dell'analisi di una linea

Quando il browser ha terminato il caricamento, l'operatore può iniziare a interagire con GARTP. Può caratterizzare l'insieme dei dati da analizzare modificando opportunamente i parametri filtro. Le richieste di elaborazione e visualizzazione dell'analisi sono inviate a GARTP selezionando almeno una delle linee presenti nel catalogo Bologna spuntando la checkbox accanto alla descrizione della direzione.

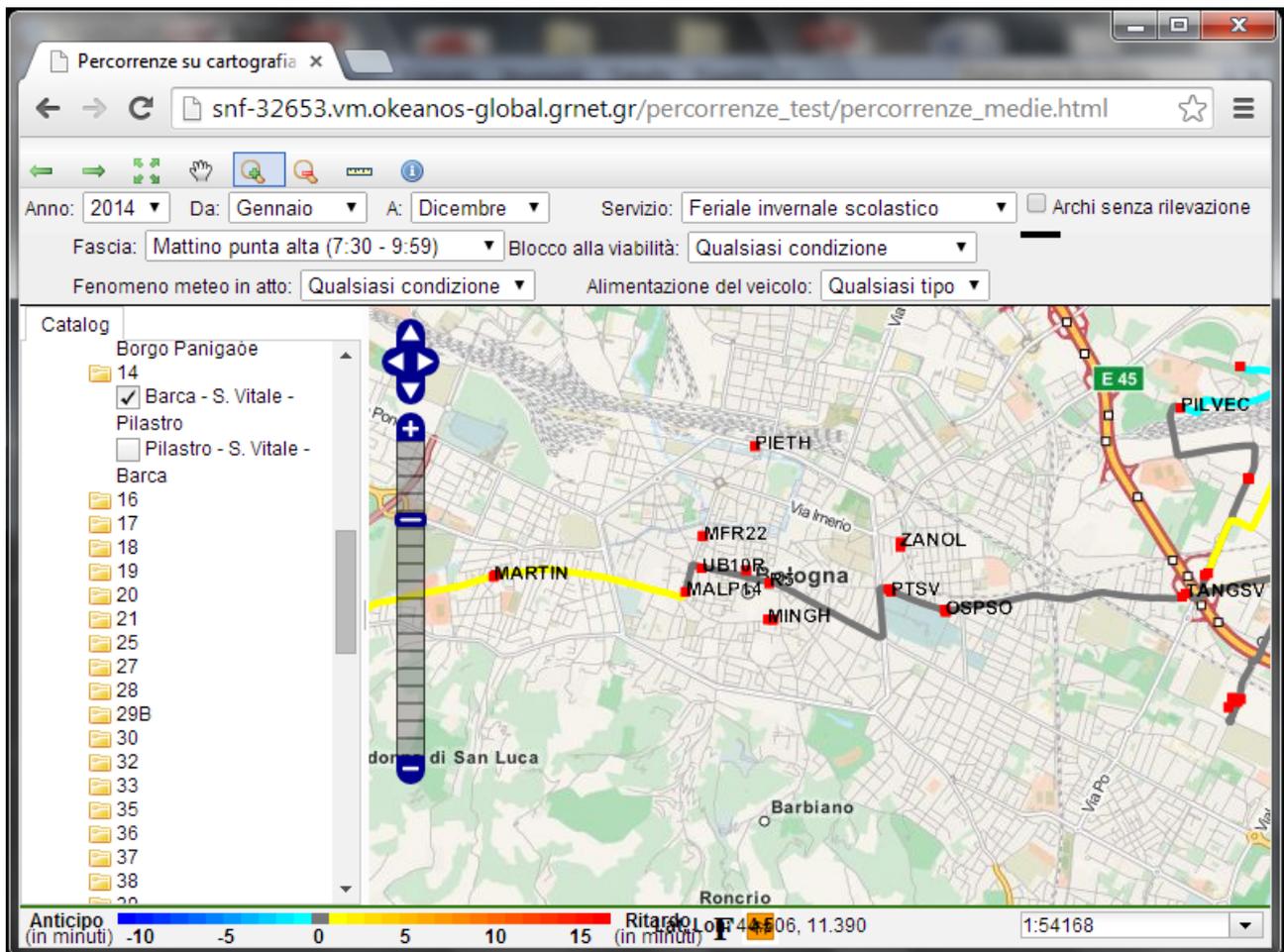


Figura 12 - Visualizzazione di una linea

In Figura 12 si può vedere l'esito della ricerca, con il valore predefinite dei parametri di filtro, per la linea 14 in direzione Barca – S. Vitale - Pilastro.

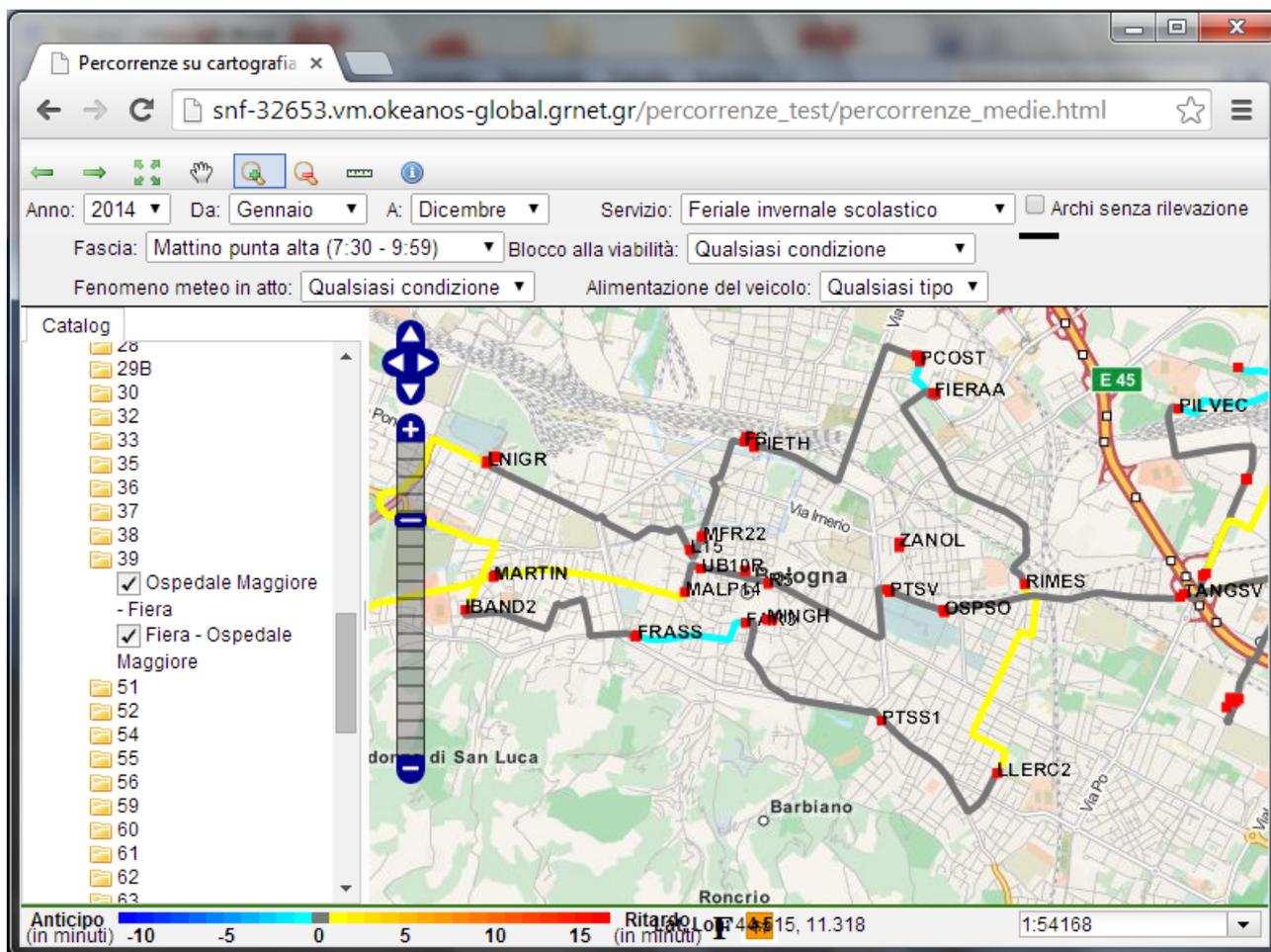


Figura 13 - Visualizzazione di due linee

In Figura 13 è stata aggiunta la linea urbana 39. Lo stato di anticipo/ritardo di ogni tratta della rete è rappresentato da segmento il cui colore cambia secondo il valore di ritardo seguendo la leggenda riportata in basso. I nodi utilizzati dalle due linee sono rappresentati da un punto rosso che ne indica la posizione sul territorio e da un'etichetta che ne riporta la sigla. Se un nodo non è raggiunto da alcun arco, questo significa che tra i dati filtrati non ve n'è alcuno che caratterizzi archi che inizino o terminino in quel nodo.

3.3.2 Visualizzazione della rete di una linea

E' possibile vedere la rete utilizzata dalla linea selezionata spuntando la checkbox Archi senza rilevazione. Dopo qualche istante saranno disegnati sulla mappa tutti gli archi percorsi dalla linea. Gli archi per i quali è disponibile l'informazione nell'insieme dei dati filtrati saranno colorati, i rimanenti avranno il colore nero, un colore che non rientra nella leggenda ed è facilmente distinguibile.

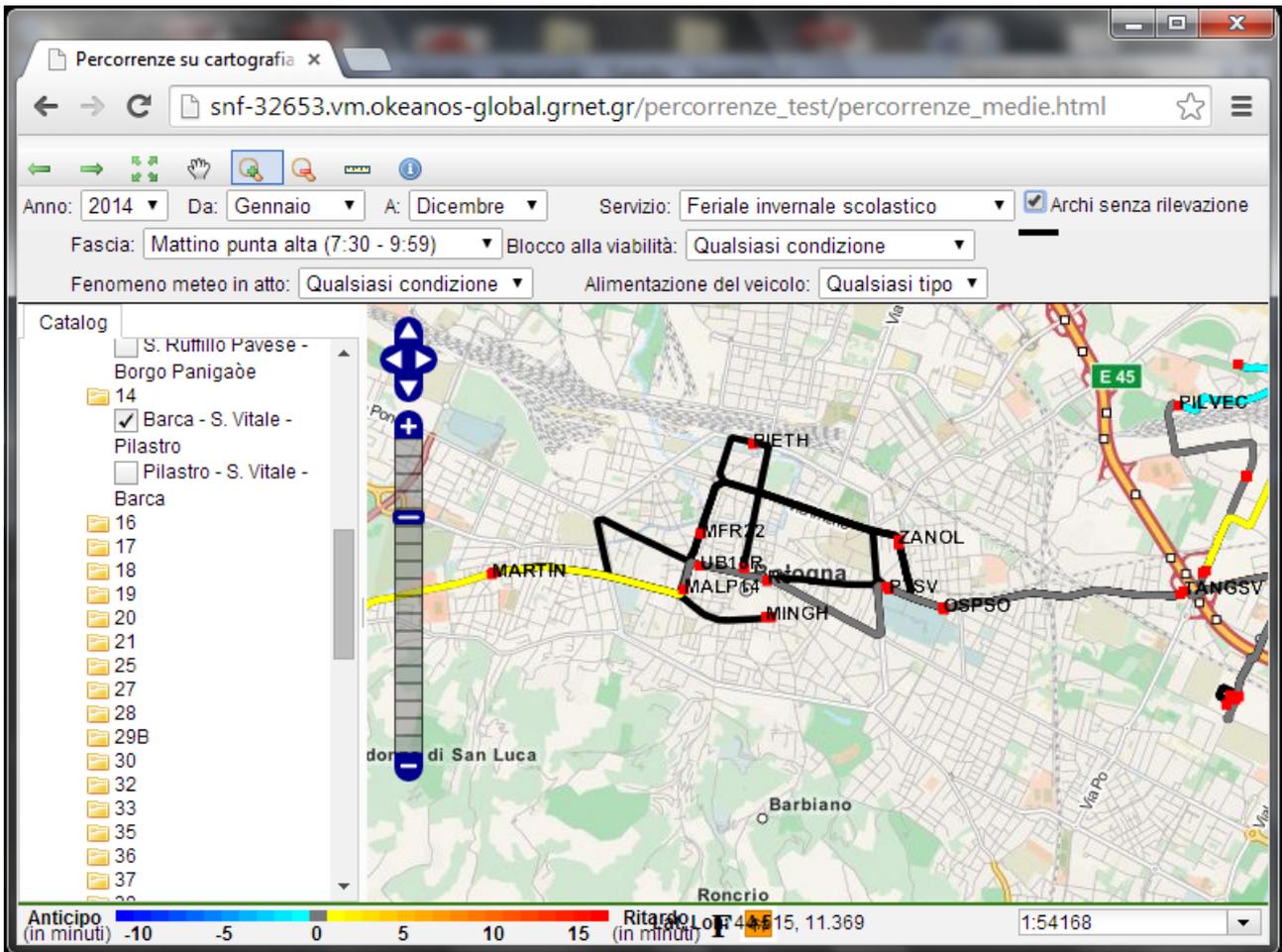


Figura 14 - Visualizzazione archi senza rilevazione

In figura 14 è rappresentata la linea 14 in direzione Bologna – S. Vitale – Pilastro già vista in precedenza. Il segmento nero rappresenta una tratta inutilizzata nell'anno 2014 ma in uso nel 2012. La parte del percorso disegnata in nero è la parte della rete abilitata alla circolazione dei veicoli per il trasporto pubblico e potrebbe essere reimpiegata in caso di ristrutturazione della rete. Se un operatore attivasse tutte le linee della rete e fosse spuntato il checkbox Archi senza rilevazione, la mappa visualizzata conterrebbe gli archi che possono essere utilizzati per il trasporto pubblico nella provincia di Bologna.

3.3.3 Comandi di navigazione

I comandi di navigazione consentono di interagire con la mappa, ad esempio aumentando lo zoom o muovendo visualizzando le informazioni di un oggetto presente nella mappa.

Vediamo solo alcuni comandi.

3.3.3.1 Iniziamo dallo zoom

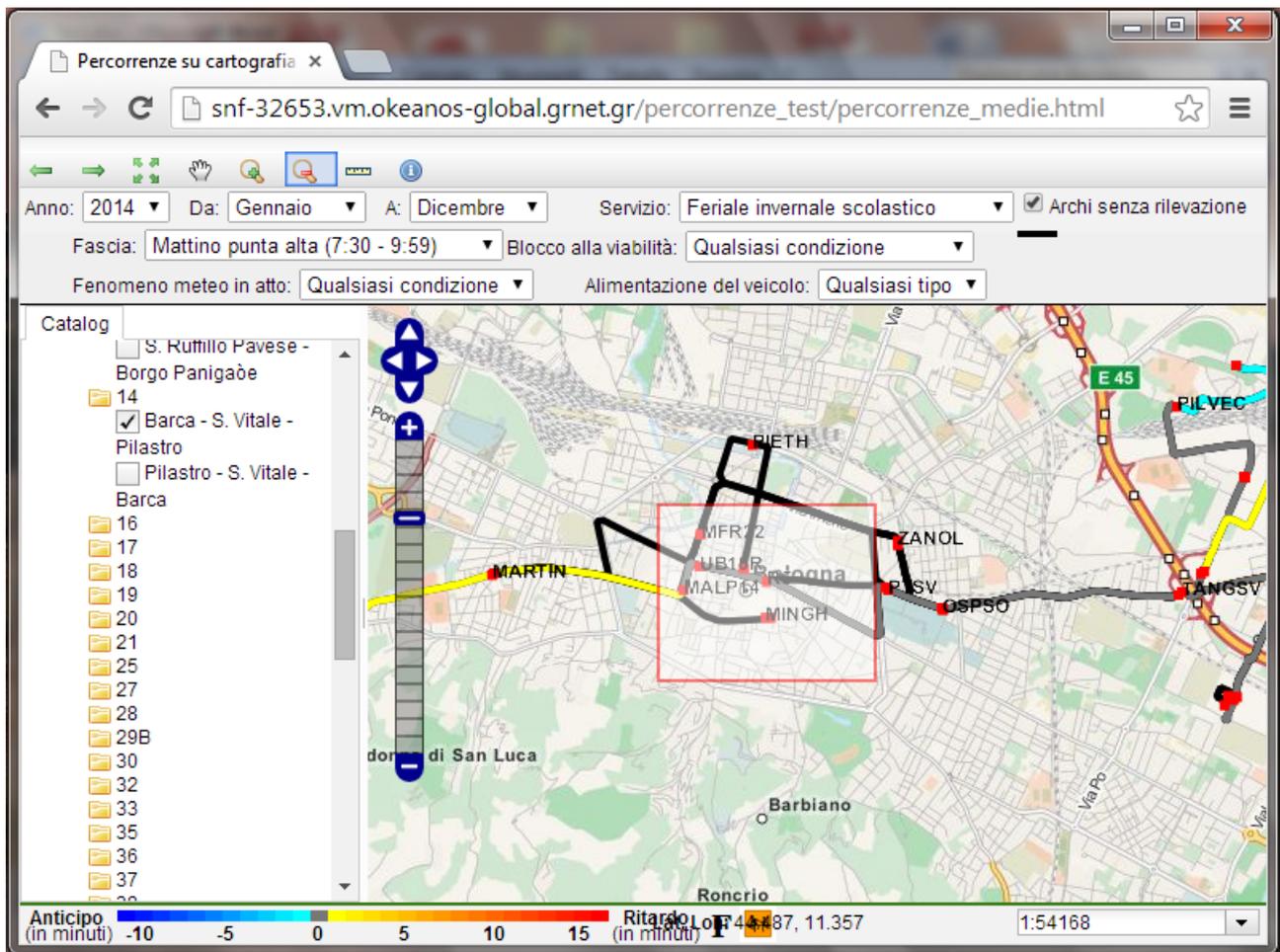


Figura 15 - Attivazione Zoom In

Premuto il tasto di zoom, ci si pone su un punto della mappa. Tenendo premuto il tasto di selezione del mouse e trascinandolo si otterrà il tracciamento di un rettangolo che rappresenta l'area di cui si vuole lo zoom. (Figura 15)

Il risultato sarà il seguente.

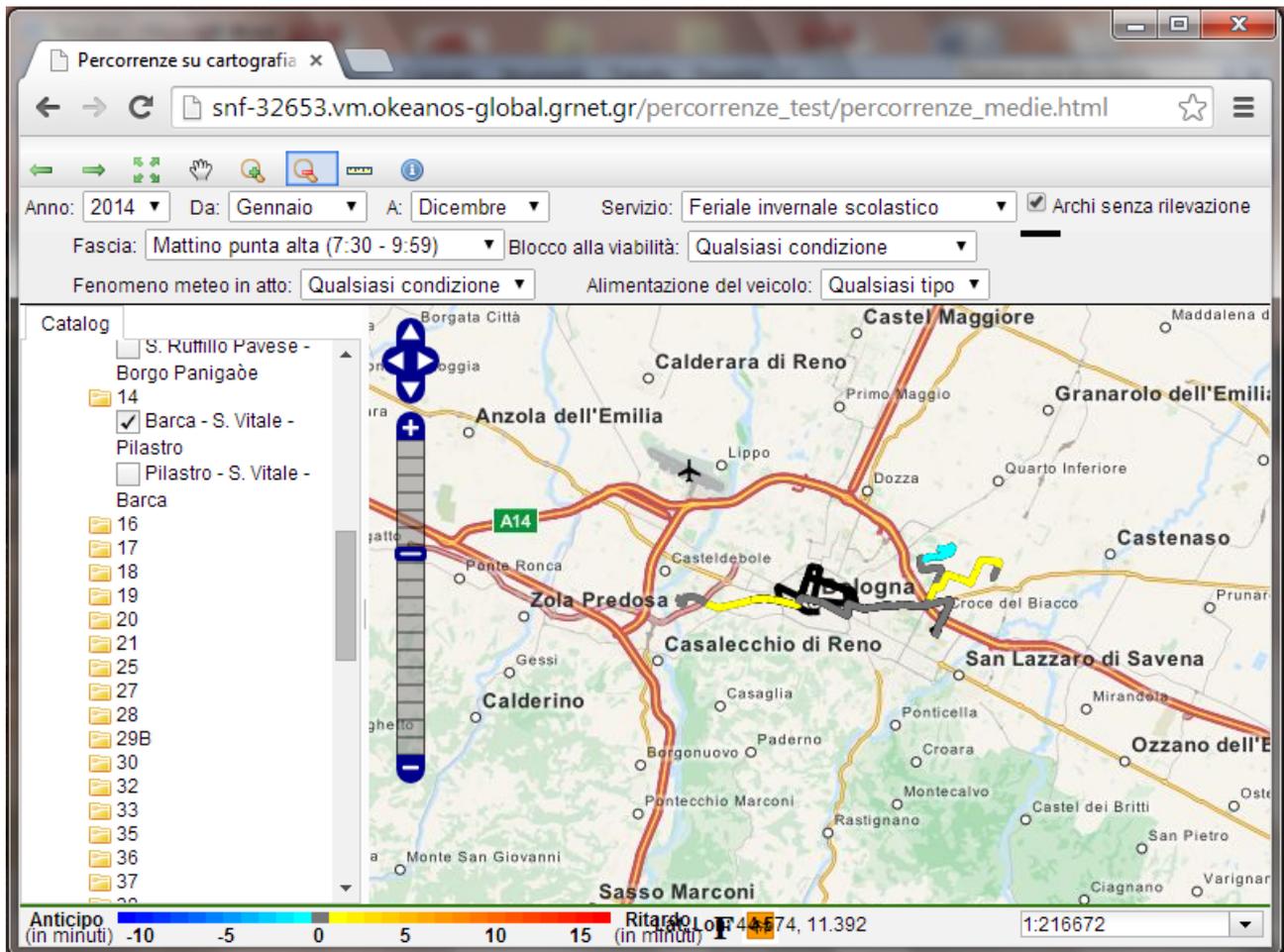


Figura 16 - Risultato Zoom In

Poiché la cartografia ha solo un numero finito di fattori di zoom, la visualizzazione ottenuta avrà un fattore di scala minore o al più uguale a quello desiderato. (Figura 16)

3.3.3.2 Misuratore di distanze

Lo strumento “Measure” consente di compiere misure di lunghezza. Si attiva cliccando sul pulsante. Il tab Measure appare accanto a Catalog. Il tab Measure conterrà i dati acquisiti durante la misurazione. Quando il mouse passa sulla mappa, il cursore cambia forma. Il primo punto della mappa su cui si clicca diventa il punto *A* da cui inizia la misurazione. (Figura 17)

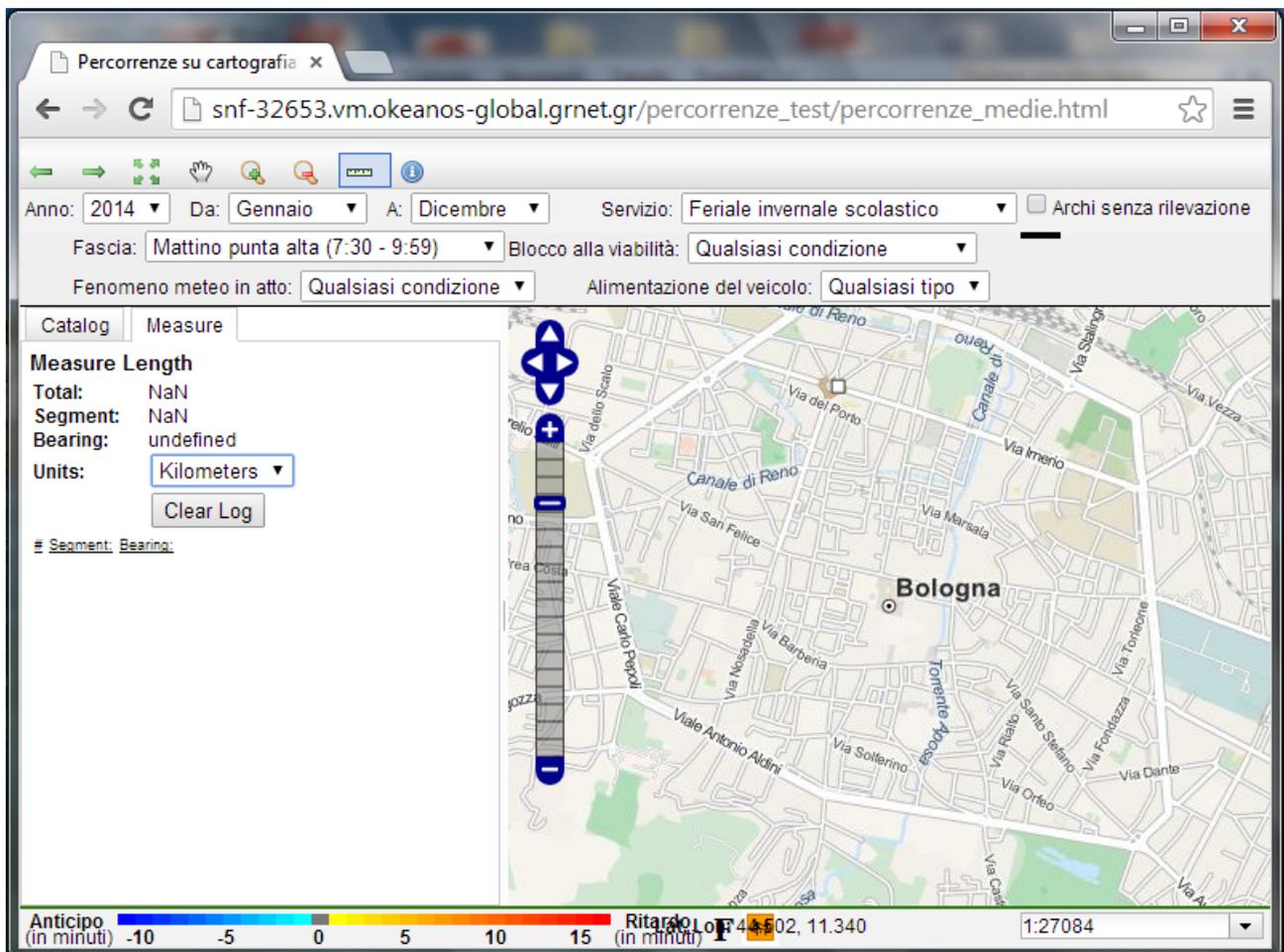


Figura 17 - Attivazione misurazione

Spostando il mouse in un altro punto della mappa, l'operatore vedrà apparire un segmento rosso che ha origine nel punto *A*. Dentro al tab *Measure* appariranno la direzione del segmento e la sua misura. Cliccando nuovamente nel punto *B*, la misura del segmento *AB* è registrata nel tab *Measure*. Il punto *B* diventa la nuova origine della misura verso un punto *C* e si ripete quanto avvenuto per il segmento *AB*. Il totale, costituito dalla somma della

misura di tutti i segmenti, è riportato nel tab Measure. Il pulsante “Clear Log” azzerava i dati della misurazione corrente. È disponibile più di unità di misura. In figura è visibile una misurazione in corso. (Figura 18)

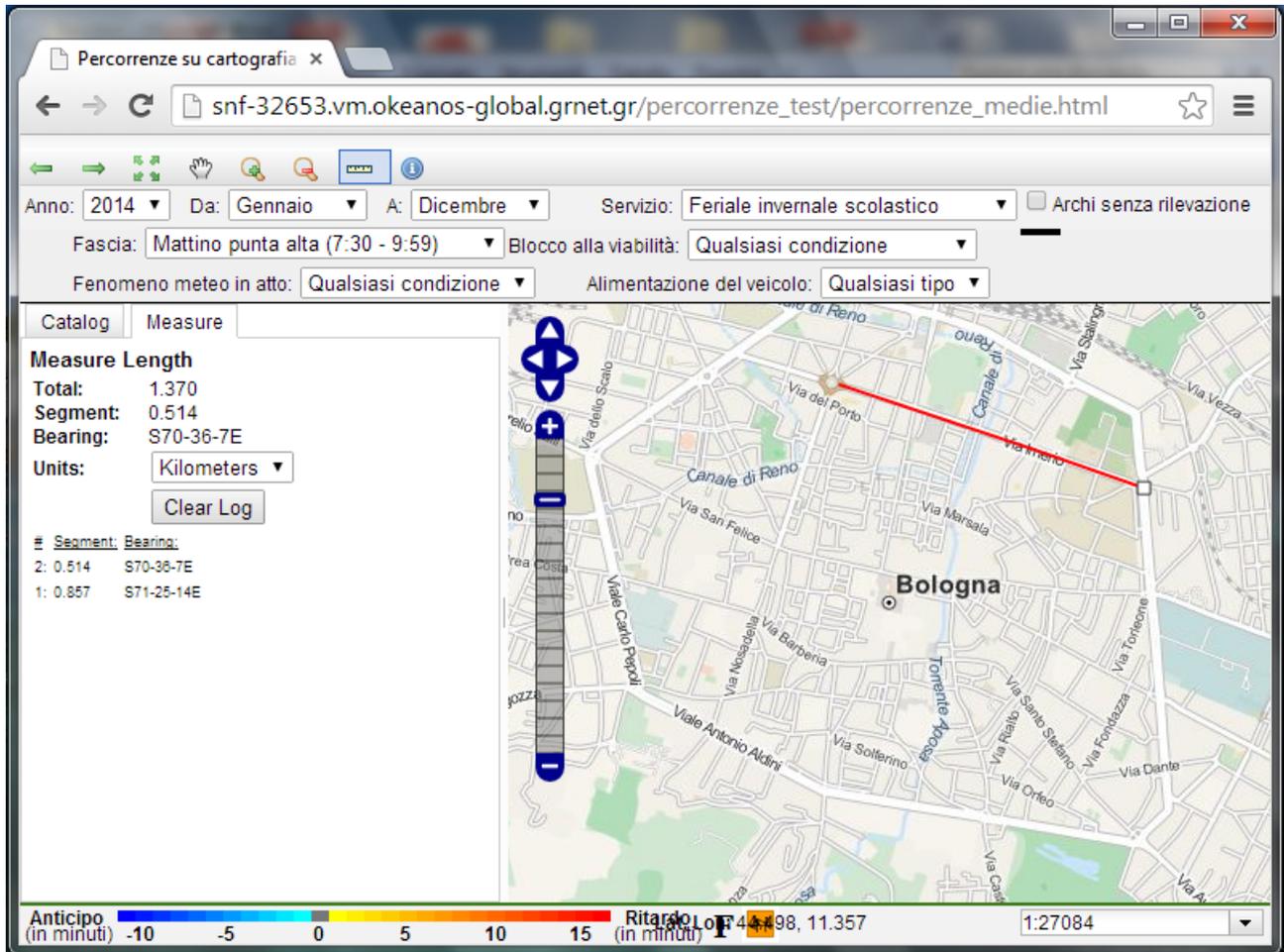


Figura 18 - Fase di misurazione

3.3.3.3 Identifica

Successivo strumento a disposizione dell'operatore è "Identifica" che consente di avere informazione sugli oggetti selezionati sulla mappa. Premendo il pulsante si vedrà questo:

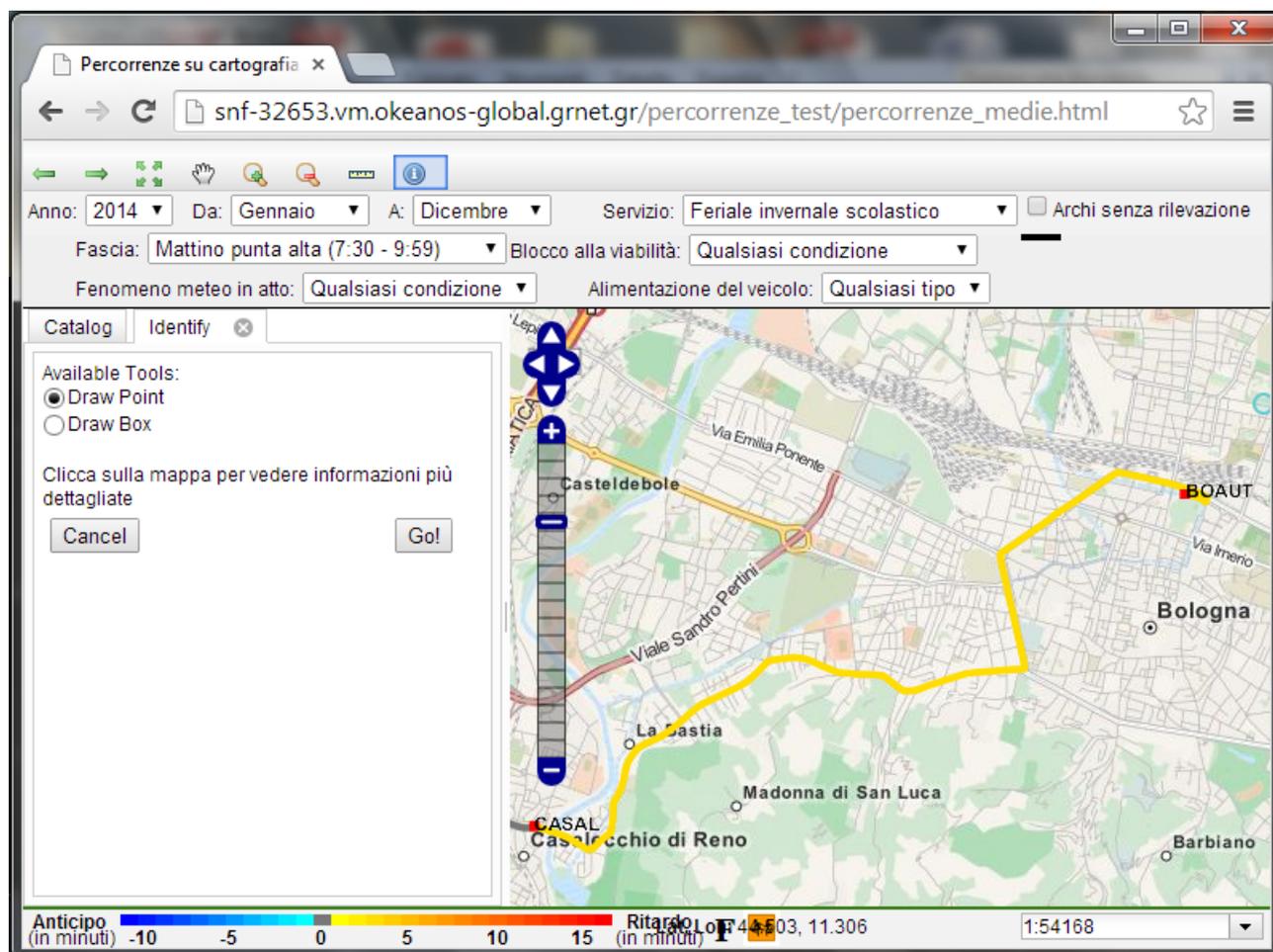


Figura 19 - Attivazione identifica.

La modalità "Draw Point" consente di ottenere le informazioni sul punto della mappa su cui sarà eseguito il successivo click con il mouse. (Figura 19)

La modalità “Draw Box” selezionerà il punto centrale del rettangolo tracciato dall’operatore.
(Figura 20)

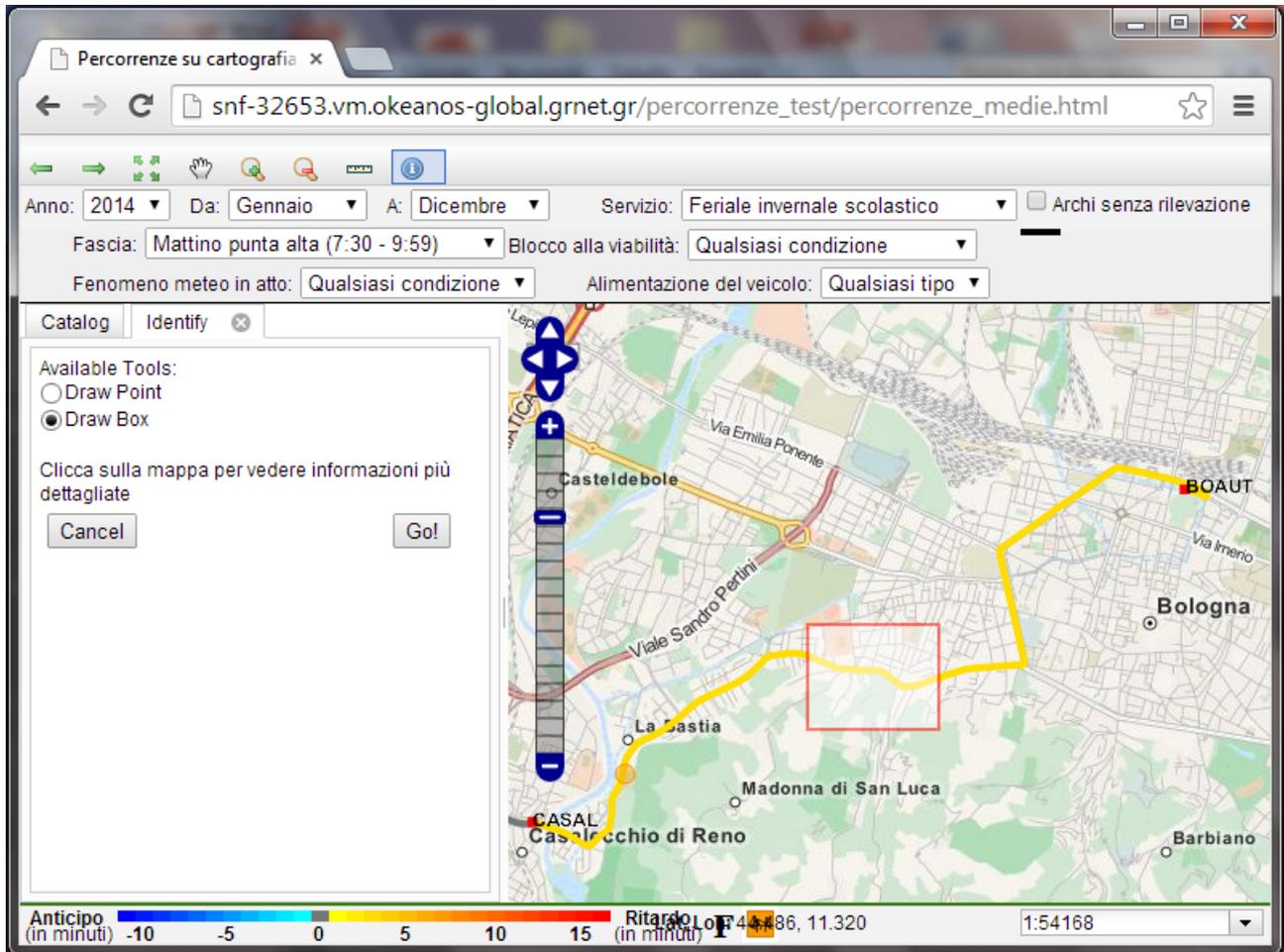


Figura 20 - Selezione identfica.

Il risultato sarà il seguente. (Figura 21)

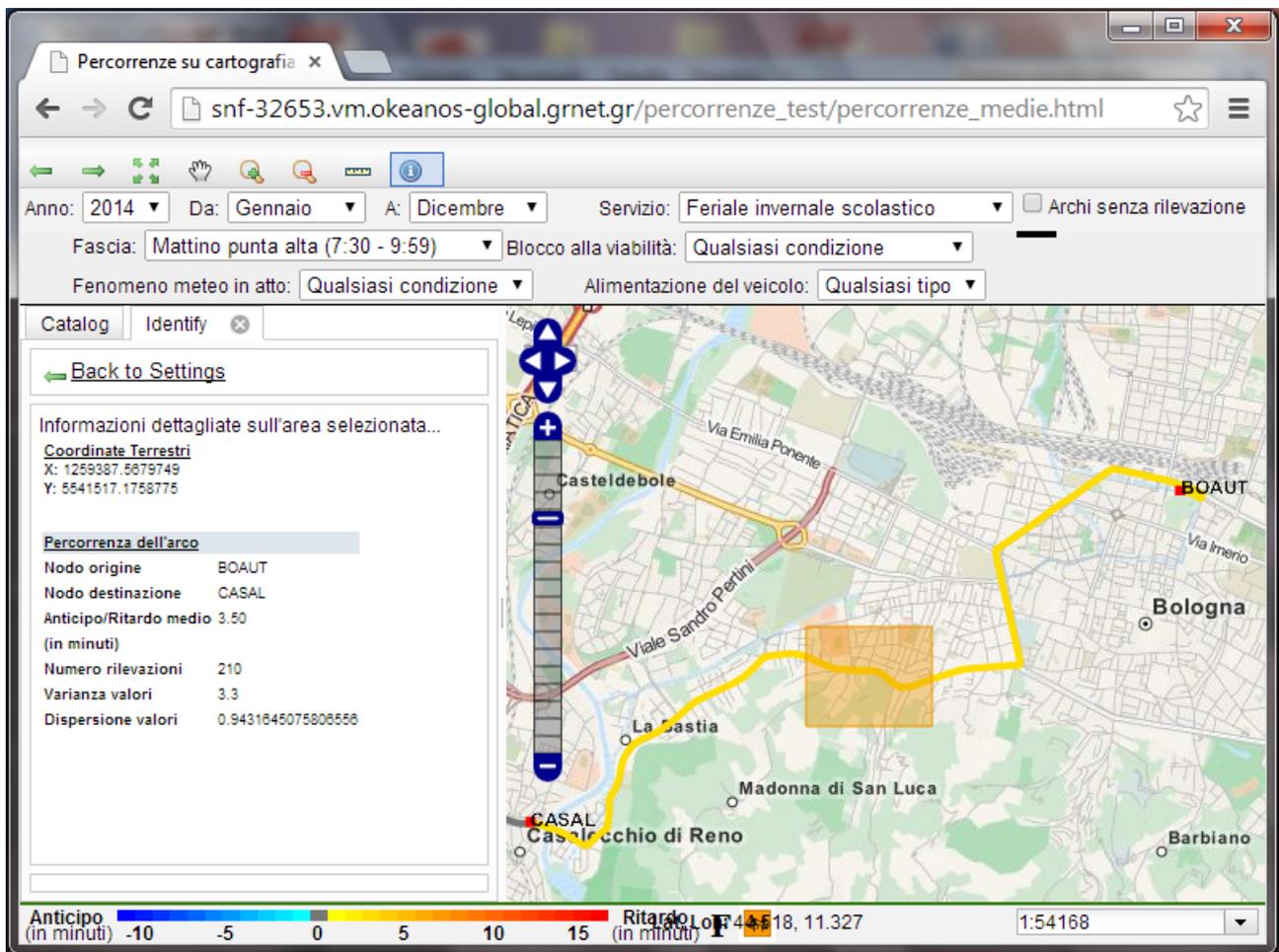


Figura 21 - Esito identifica.

Sulla mappa permane evidenziata l'area selezionata dall'operatore (). Sulla sinistra sono riportate le informazioni che si riferiscono alle tratte selezionate:

- nodo origine;
- nodo destinazione;
- anticipo/ritardo medio, se il valore è inferiore a 0 si tratta di anticipo, se il valore è maggiore di 0 si tratta di ritardo;
- numero di rilevazioni, ossia quante rilevazioni sono state trovate e analizzate;
- varianza dei valori, la varianza standard calcolata sull'insieme delle rilevazioni trovate;
- dispersione dei valori, ulteriore analisi sulla variabilità sull'insieme delle rilevazioni trovate.

4. ARCHITETTURA DI GARTP

GARTP è strutturato su più livelli. Ognuno ha una specifica funzione e comunica con il livello adiacente utilizzando protocolli standard.

4.1 Struttura concettuale e concreta dell'applicazione

GARTP è costituito da quattro livelli distinti:

- la banca dati, contenente tutte le informazioni sui transiti, sulle rilevazioni meteorologiche, sui blocchi alla circolazione, sulle caratteristiche degli autobus;
- un gestore della cartografia;
- un server http, che distribuisce il WebGIS Client ed elabora alcune risposte;
- il WebGIS Client, che fornisce l'interfaccia utente e che gestisce la comunicazione con le fonti dati.

GARTP è rappresentato nel suo complesso nel seguente schema.

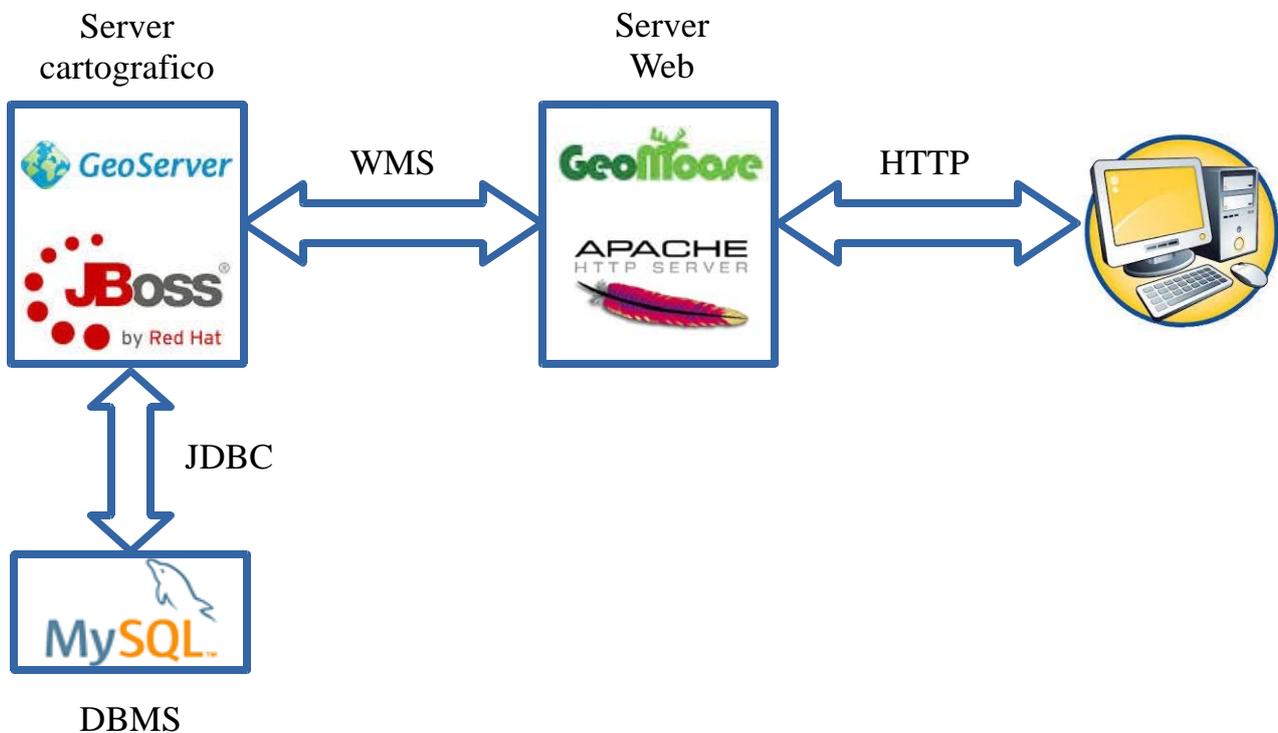


Figura 22 - Struttura di GARTP

L'intera architettura di GARTP può essere installata su uno stesso computer o le sue parti possono essere distribuite o replicate su più server. Nella scelta dei prodotti software da impiegare come elementi di GARTP, laddove esistessero e fossero performanti, sono preferiti gli OpenSource. Il primo motivo è contenere i costi totali di GARTP. Il secondo motivo è la possibilità di personalizzare o integrare il codice della parte. Terzo motivo è la disponibilità delle soluzioni su più piattaforme. L'Open Source Geospatial Foundation (OSGeo) è stata creata per sostenere lo sviluppo collaborativo del software geospaziale open source e per promuovere la diffusione del suo uso. Nel suo portale sono elencate diverse soluzioni software e alcune di esse sono impiegate in GARTP.

4.2 Principali tecnologie utilizzate

Analizziamo le tecnologie impiegate nell'architettura e quali prodotti software sono stati scelti per implementarle.

4.2.1 Database Management System

GARTP utilizza le banche dati delle percorrenze degli archi, rilevate e programmate, della descrizione della rete di trasporti, delle rilevazioni meteorologiche, degli incidenti stradali e autostradali, delle fiere, dei concerti, delle caratteristiche degli autobus in dotazione.

La dimensione delle banche dati è vasta e crescerà con il passare del tempo. Inoltre, altri tipi d'informazioni potrebbero essere aggiunte per ampliare le possibilità di analisi di GARTP.

Il DBMS candidato alla gestione delle banche dati deve:

- essere un RDBMS;
- avere la capacità di rispondere velocemente alle interrogazioni al crescere della dimensione dei dati;
- avere un protocollo di comunicazione comune alle altre parti dell'architettura che devono accedere al dato;
- implementare gli standard OGC necessari a rappresentare la cartografia digitale.

Sono diversi i DBMS che soddisfano i quattro requisiti. I più noti sono Oracle, SQL Server, Postgres SQL con estensione PostGIS, IBM DB2, IBM Informix, MySQL, ENSRI ARCSDE.

Limitando l'elenco ai soli prodotti OpenSource, la lista si riduce a due nomi: Postgres SQL con estensione PostGIS e MySQL.

La scelta è ricaduta su MySQL. La versione installata è la 5.6.21.

4.2.2 Gestore della cartografia

Il gestore della cartografia si occupa di leggere le informazioni contenute nel database cartografico, di elaborarle e di renderle disponibili implementando gli standard definiti dell'Open Geospatial Consortium (OGC).

Le soluzioni disponibili sono diverse. La più conosciuta sono ArcGIS, MapInfo, MapServer, GeoServer, Deegree. Tra queste sono OpenSource MapServer, GeoServer, Deegree.

La scelta è ricaduta su Geoserver, un server GeoSpaziale scritto in Java, seguendo le comuni pratiche Java Enterprise, per la gestione, la disseminazione e l'analisi di dati geospaziali. GeoServer permette di distribuire, manipolare e analizzare i dati usando i più diffusi standard OGC (WMS, WFS, WCS, WPS), senza però dimenticare estensioni specifiche per un'interazione trasparente con client come Google Earth e software commerciale in genere, né gli ormai diffusi approcci basati su protocolli REST e GeoJSON per la distribuzione semplificata di semplici dati vettoriali. Obiettivo della presentazione è fornire allo spettatore un'esaustiva panoramica delle funzionalità offerte da GeoServer per la creazione di Spatial Data Infrastructure interoperabili, con particolare attenzione alle nuove funzionalità introdotte in GeoServer 2.1 e all'analisi di dati spaziali con il protocollo WPS 1.0.

GeoServer è distribuito come applicazione sia stand alone sia per application server. Questa seconda possibilità è stata la scelta. Il progetto è sostenuto da OSGeo. La versione installata è la 2.4.4.

4.2.3 Application Server

L'application server è composto di moduli realizzati secondo standard ben definiti e accettati dalla comunità mondiale dei programmatori.

I moduli normalmente presenti in un application server sono:

- modulo web server che espone al client browser la logica di presentazione statica delle applicazioni e in diretta interazione con la sottostante logica di business;
- contenitore di componenti server-side detti anche logica di business;
- gestore degli accessi degli utenti e della sicurezza;
- gestione accesso a database o in generale a sorgenti di dati esterne;

- gestore transazioni;
- interfaccia per l'accesso a un sistema legacy;
- altri componenti per massimizzare le prestazioni, come connection pool, load balancer, caching, ecc.

Allo stato attuale delle cose, le tecnologie su cui possono basarsi gli application server sono diverse: Microsoft .NET, Java di Sun, Zope. Essendo una specifica, quella Java annovera numerose implementazioni sia commerciali sia open source: Apache Tomcat, JBoss, WebLogic, WebSphere, Geronimo, WildFly.

Gli standard di Java non sono unicamente frutto della Sun Microsystem, ma sono il risultato di un rapporto sinergico tra l'azienda americana e la partecipazione libera di sviluppatori in tutto il mondo. Gli standard su cui si fonda la tecnologia .NET sono invece stabiliti direttamente da Microsoft.

Nel mondo Java, gli application server compatibili con le specifiche della Sun sono detti application server J2EE.

La scelta è ricaduta su WildFly, un application server open source che implementa l'intera suite di servizi Java EE. Essendo basato su Java, WildFly è un application server multiplatforma, utilizzabile su qualsiasi sistema operativo per il quale esista un'implementazione conforme agli standard della Java virtual machine. La versione installata è la 8.2.0.

4.2.4 WebGIS Client

Il WebGIS Client si occupa di interagire con l'operatore e di presentare i risultati delle sue richieste al gestore della cartografia,

I prodotti disponibili sono diversi. Solo nell'ambito di OSGeo troviamo: Geomajas, GeoMOOSE, Mapbender, MapGuide.

La scelta è ricaduta su GeoMOOSE.

GeoMOOSE è un'infrastruttura in JavaScript che realizza un Web Client per la visualizzazione di dati cartografici. GeoMOOSE ha diversi punti di forza che includono la modularità, la configurabilità e definisce un insieme di funzionalità fondamentali nei suoi pacchetti. GeoMOOSE è anche leggerissimo per i server rendendo possibile la gestione di una grande quantità di utenti, con un ampio numero di layers e di servizio senza stressare

oltremodo un server. Il core di GeoMOOSE è scritto usando JavaScript e HTML. E' possibile utilizzare GeoMOOSE ricorrendo a nulla di più che un semplice web server (Nginx, Apache, IIS). Oltre al client di base, GeoMOOSE presenta anche un certo numero di servizi aggiuntivi scritti in PHP. Questi servizi aggiungono la capacità di eseguire operazioni d'identificazione degli oggetti visualizzati, di selezione e di ricerca d'insiemi di dati. GeoMOOSE consente di interrogare diversi tipi di fonti utilizzando gli standard definiti da OGC. La versione installata è la 2.6.1.

4.2.5 Server Web

Un server web (o web server) è un'applicazione software che, in esecuzione su un server fisico, è in grado di gestire le richieste di trasferimento di pagine web di un client, tipicamente un web browser.

Un web server può avere delle estensioni che gli consentono di produrre dinamicamente le pagine web.

La disponibilità d'implementazioni professionali di server web è ampia: Apache HTTP Server, Apache Tomcat, Cassini Server Web, ColdFusion, HTTP File Server, Internet Information Services (*IIS*), KLite, lighttpd, Mongoose, NCSA HTTPd, nginx, PowerFolder, Sun ONE, Zeus Web Server.

Il prodotto scelto è Apache HTTP Server, il server più utilizzato al mondo su oltre il 50% dei siti attivi. [N15]

Il progetto Apache HTTP Server Project ha la missione di sviluppare e mantenere un server HTTP OpenSource HTTP per i moderni sistemi operativi inclusi UNIX e Windows NT. L'obiettivo di questo progetto è di fornire un server sicuro, efficiente ed estensibile che fornisca servizi HTTP corrispondenti agli attuali standard HTTP. La versione installata è la 2.2.22.

4.3 Utilizzo e personalizzazione del software

Il software selezionato è stato installato su un server virtuale accessibile dalla rete e sono stati modificati alcuni settaggi e personalizzate alcune parti. Il server virtuale è caratterizzato da un sistema quadri processore, da 4GB di RAM, da un disco di 40GB. Il sistema operativo scelto è Windows Server 2012 R2.

4.3.1 MySQL

L'installazione del DBMS è stata compiuta come servizio.

Sono stati istanziati due distinti database: cartografia e telecontrollo. Il primo contiene le informazioni cartografiche: la descrizione e la posizione dei nodi, la forma degli archi e la loro dimensione. Il secondo contiene le informazioni sull'aderenza delle percorrenze reali alle programmate, sulle condizioni meteorologiche e su dove e quando sono avvenuti gli eventi che possono aver influito sulla viabilità, sulla definizione delle fasce orarie.

Entrambi i database utilizzano store engine MyISAM perché negli utilizzi di sola lettura, com'è nel caso di GARTP, ha ottime performance con un contenuto consumo di risorse.

E' stato creato un utente ad hoc, con i soli diritti di lettura, per accedere ai due database.

4.3.2 GeoServer

Rispetto alla release, disponibile sul sito in versione war, sono stati aggiunti i driver jdbc per accedere a MySQL, recuperabili dallo stesso sito di Geoserver. Sono state aggiunte, nella cartella workspaces/tper/vers, la descrizione della sorgente dati e le descrizioni dei layer che sono interrogati dal WebGIS Client. Sono state aggiunte, nella sezione styles, le icone per rappresentare i punti di blocco della viabilità e il file svg contenente le regole per colorare gli archi.

I layer creati sono:

- NodiLinea, per la visualizzazione dei nodi utilizzati nella descrizione della/e linea/e richieste;
- PercorsoLinea, per la visualizzazione di tutti gli archi utilizzati dalla/e linea/e richieste;
- PosizioneBlocchiViabilita, per la visualizzazione della posizione dell'evento che ha provocato un possibile blocco della viabilità;
- ViewConBloccoViabilita, per la visualizzazione dell'anticipo/ritardo della/e linea/e richieste in caso di blocco della viabilità;
- ViewConBloccoViabilitaConFenomenoMeteo, per la visualizzazione dell'anticipo/ritardo della/e linea/e richieste in caso di blocco della viabilità e con una certa condizione meteorologica;
- ViewConBloccoViabilitaConFenomenoMeteoConTipoAlimentazione, per la visualizzazione dell'anticipo/ritardo della/e linea/e richieste in caso di blocco della

viabilità, con una certa condizione meteorologica e per i veicoli aventi un certo tipo di alimentazione;

- ViewConBloccoViabilitaConTipoAlimentazione, per la visualizzazione dell'anticipo/ritardo della/e linea/e richieste in caso di blocco della viabilità e per i veicoli aventi un certo tipo di alimentazione;
- ViewConFenomenoMeteo, per la visualizzazione dell'anticipo/ritardo della/e linea/e richieste con una certa condizione meteorologica;
- ViewConFenomenoMeteoConTipoAlimentazione, per la visualizzazione dell'anticipo/ritardo della/e linea/e richieste con una certa condizione meteorologica e per i veicoli aventi un certo tipo di alimentazione;
- ViewConTipoAlimentazione, per la visualizzazione dell'anticipo/ritardo della/e linea/e richieste per i veicoli aventi un certo tipo di alimentazione;
- ViewPercorrenze, per la visualizzazione dell'anticipo/ritardo della/e linea/e richieste.

A ogni layer sono associati una specifica query parametrica sul database i cui parametri sono impostati rispetto ai valori inviati dall'interfaccia utente. L'esito della query è visualizzato concordemente allo stile indicato nella descrizione del layer.

Il deploy è stato compiuto sull'application server.

4.3.3 WildFly

L'application server è stato installato come servizio.

Sono stati aggiunti i driver jdbc per l'accesso a MySQL, collocandoli nell'apposizione sezione ed è stato integrando il file di configurazione del server per renderli disponibili. È stato attivato il broadcast per consentire l'accesso ai client remoti. Sono state aggiunte la descrizione delle fonti dati ponendo il loro file di configurazione nella sezione deploy dell'application server.

4.3.4 GeoMoose

La release base del programma è stata collocata nell'apposita cartella del server web.

L'homepage è stata modificata per aggiungere: la legenda colorata in basso a sinistra, il caricamento delle api di OpenStreetMap, la sezione contenente i parametri di filtro dati. È

stato aggiornato il file di configurazione mapbook.xml affinché contenesse le informazioni necessarie. In dettaglio:

- modificato il sistema di riferimento cartografico EPSG:4326 a EPSG:3857;
- impostati estensione iniziale e massima estensione della mappa da generare;
- definite le map-source e i relativi parametri per ognuno dei layer che possono essere visualizzati dall'utente;
- modificato il servizio identify aggiungendo i parametri specifici (fascia, giorno_inizio, giorno_fine, servizio, blocco_viabilita, fenomeno_meteo, tipo_alimentazione) presenti nell'interfaccia utente;
- definito il catalogo dei layer consultabili organizzandolo in Cartografie e Bologna: il primo contiene le possibili cartografie visualizzabili, il secondo contiene l'elenco dei layer associati a gruppi o a singole linee.

Sono state apportate alcune modifiche al file OpenLayers.js in modo tale che le richieste al servizio cartografico contenessero i corretti parametri. A tal fine sono state introdotte la verifica sulla correttezza dei valori dei parametri specificati dall'utente e la segnalazione del motivo per il quale tali valori sono errati, qualora ve ne fossero.

4.3.5 Http Server

Il server HTTP è installato dal pacchetto ms4w versione 3.0.6. Il server è stato installato come servizio. E' stato inserito il file di configurazione httpd_percorrenze_test.conf nella directory conf per attivare la sezione percorrenza_test. Nella directory apps è stato installato il webGIS Client GeoMOOSE.

5. USABILITA' DI GARTP

La valutazione dell'usabilità di GARTP è stata realizzata secondo la *System Usability Scale (SUS)*. Al test, tradotto in italiano e qui sotto riportato, sono stati sottoposti 9 addetti della programmazione e pianificazione del servizio dell'azienda che ha fornito i dati impiegati per popolare GARTP.

System Usability Scale

© Digital Equipment Corporation, 1986.

	Molto in disaccordo				Molto d'accordo
1. Penso che mi piacerà usare questo sistema frequentemente	1	2	3	4	5
2. Trovo il sistema inutilmente complesso	1	2	3	4	5
3. Ho pensato che il sistema fosse facile da usare	1	2	3	4	5
4. Penso che avrò bisogno dell'aiuto di un tecnico per essere in grado di usare questo sistema	1	2	3	4	5
5. Ho trovato che le varie funzioni di questo sistema fossero bene integrate	1	2	3	4	5
6. Ho pensato che c'era troppa inconsistenza in questo sistema	1	2	3	4	5
7. Immagino che la maggior parte delle persone potrebbe imparare molto velocemente ad usare questo sistema	1	2	3	4	5
8. Ho trovato il sistema molto complicato da usare	1	2	3	4	5
9. Mi sento in grado di usare il sistema	1	2	3	4	5
10. Ho bisogno di imparare un sacco di cose prima che io possa entrare in sintonia con questo sistema	1	2	3	4	5

La scala SUS è utilizzata generalmente dopo che i rispondenti hanno avuto l'opportunità di usare il sistema da valutare, prima che ogni debriefing o discussione avesse luogo. Ai rispondenti si dovrebbe chiedere di rispondere di getto a ogni voce presente nel test, piuttosto che meditare la risposta per lungo tempo.

Tutte le voci del test devono avere una risposta. Se il rispondente sente di non poter rispondere a una voce, dovrebbe barrare la casella centrale della scala.

SUS produce un numero singolo che rappresenta una misura composita dell'usabilità complessiva del sistema da valutare. Si noti che il punteggio della singola voce non ha significato in sé.

Per calcolare il punteggio SUS, per prima cosa si sommano i punteggi di ciascuna voce. Ogni voce contribuirà con un valore compreso tra 0 e 4. Per le voci 1,3,5,7 e 9 il contributo al punteggio è pari alla posizione nella scala meno 1. Per le voci 2,4,6,8 e 10 il contributo al punteggio è pari a 5 meno la posizione nella scala. Si moltiplica la somma per 2.5 per ottenere il valore complessivo.

Il punteggio di SUS rientra nell'intervallo tra 0 e 100.

Il percentile ottenuto dal calcolo del punteggio SUS può essere espresso da una lettera.

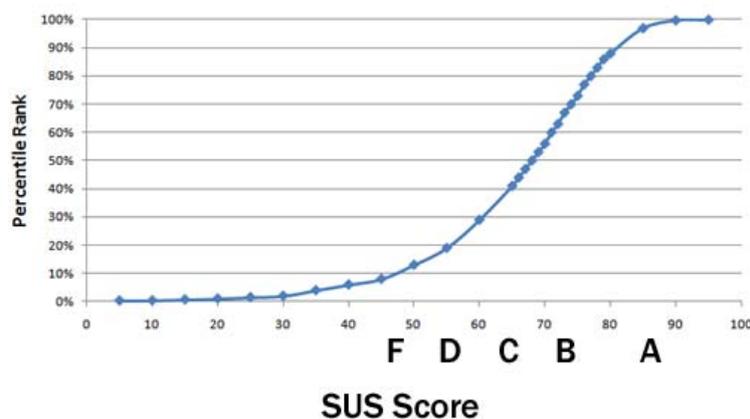


Figura 23 - Percentile SUS

L'equivalenza tra punteggio SUS e il termine che qualifica l'usabilità del sistema è espressa dal seguente grafico [BKM09]:

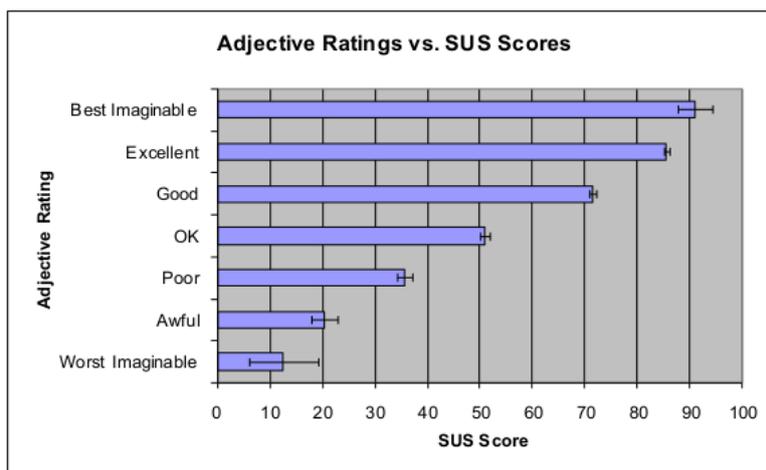


Figura 24 - Associazione percentile SUS aggettivo

5.1 Test di usabilità

Le funzionalità dell'interfaccia di GARTP sono state introdotte ai partecipanti al test con un colloquio personale, presso un luogo isolato dotato di un solo computer. Poi a ognuno di loro è stato chiesto di seguire ed eseguire la seguente successione di azioni:

- Visualizzare l'andamento della linea 671 in direzione Bologna – Vignola nel periodo compreso tra febbraio e giugno 2013 in una giornata di sole nella fascia di punta pomeridiana nel servizio feriale invernale scolastico.
- Trovare percorrenza dell'arco CRESP – BAZZV.
- Visualizzare la linea 656 nella direzione Bologna - Bazzano - Zocca.
- Trovare la percorrenza dell'arco BAZZFBV – ZICASS.
- Visualizzare l'andamento della linea 671 in direzione Vignola – Bologna nel periodo compreso tra febbraio e giugno 2013 in una giornata di pioggia nella fascia di punta mattinata nel servizio sabato invernale scolastico.
- Trovare percorrenza dell'arco CASAL2 – BOAUT.
- Portare la visualizzazione allo zoom iniziale.

E' stato cronometrato il tempo per completare l'intera sequenza.

Terminata l'ultima azione in elenco, a ogni addetto è stato somministrato il test SUS tradotto in lingua italiana, visibile nell'appendice B.

5.2 Risultato della SUS

La tabella 2 riassume il tempo impiegato per compiere la prova e il punteggio del test SUS per ciascuno dei 9 addetti.

Tabella 2 - Risultato test di usabilità

Addetto	Tempo completamento prova	di Punteggio della SUS
1	5 minuti 55 secondi	87,5
2	7 minuti 55 secondi	82,5
3	5 minuti 57 secondi	95
4	9 minuti 55 secondi	92,5
5	14 minuti 15 secondi	50
6	10 minuti 44 secondi	100
7	6 minuti 35 secondi	100
8	4 minuti 50 secondi	95
9	6 minuti 30 secondi	75
	Valori medi	
	8 minuti 4 secondi	86,4
	Deviazione standard	
	3 minuti 1 secondo	15,9

Il valore medio SUS di 86,4 può essere espresso con l'aggettivo "Excellent". [BKM09]

6. CONCLUSIONI

GARTP permette di analizzare i disallineamenti tra la programmazione e l'esercizio del servizio pubblico sulla sua intera estensione. La navigazione è comoda, intuitiva, fluida. Rispetto alle soluzioni individuate in letteratura, GARTP consente di aggiungere alcuni importanti parametri di analisi: la condizione meteorologica, gli eventi generanti criticità nella circolazione e la tipologia di alimentazione dell'autobus impiegato in linea. Una maggiore selettività sul dato degli anticipi e dei ritardi aiuta ad analizzarlo meglio per individuare criticità ricorrenti che vanno corrette.

GARTP è più flessibile delle soluzioni presentate. La sua strutturazione a livelli gli consente di accedere e di visualizzare una molteplicità di dati aggiungendone la sorgente al server cartografico o al WebGIS Client. L'analisi degli anticipi o dei ritardi è solo una delle tante informazioni gestibili da GARTP.

Il problema più immediato è rappresentato dalla visualizzazione di linee che utilizzano uno o più archi in comune o che si sovrappongono. Allo stato attuale è visibile solo l'ultimo disegnato in ordine di tempo. Il problema nasce dall'utilizzo dello stesso insieme di archi per tutte le linee e dalla difficoltà di individuare un modo chiaro di rappresentare l'informazione: disegnare linee colorate vicine tra loro può renderle indistinguibili o creare effetti ottici che inducono a interpretazioni errate dell'informazione.

GARTP può essere ulteriormente esteso e migliorato:

- dal censimento dei problemi alla circolazione non segnalati dalla stampa, ad esempio banali tamponamenti possono generare importanti congestioni del traffico;
- dai rapporti e dalle segnalazioni inviati dagli addetti della gestione e del controllo, ad esempio situazioni a bordo dell'autobus possono averne rallentato la marcia;
- dalle segnalazioni inviate dai clienti del trasporto.

GARTP potrebbe visualizzare le rilevazioni in tempo reale degli AVL degli autobus contestualmente alle analisi storiche presentate dall'applicazione. Utilizzando la stessa leggenda di colori, per il dato sia storico sia in tempo reale, si avrebbe un immediato confronto tra le medie storiche e la condizione di circolazione attuale. Più i colori si avvicinano più la situazione attuale rispecchia i trend storici.

7. BIBLIOGRAFIA

- [MBPTT04] G. Mintsis, S. Basbas, P. Papaioannou, C. Taxiltaris, I.N. Tziavos *Applications of GPS technology in the land transportation system*
European Journal of Operational Research 152 (2004) 399–409
- [W04] C. Ware, *Information Visualization: Perception for Design, Second Edition*
Morgen Kaufman, USA, 2004.
- [C01] L. Chittaro *Information visualization and its application to medicine*
Artificial Intelligence in Medicine 22 (2001) 81-88
- [HMM00] I. Herman , G. Melanc,on, MS Marshall. *Graph visualization and navigation in information visualization: a survey.*
IEEE Transactions on Visualization and Computer Graphics 2000; 6: 24–43.
- [K02] Daniel A. Keim *Information Visualization and Visual Data Mining*
IEEE Transactions on Visualization and Computer Graphics, Vol. 8, NO. 1, January-March 2002
- [K03] Menno-Jan Kraak *Geovisualization illustrated*
ISPRS Journal of Photogrammetry & Remote Sensing 57 (2003) 390–399
- [S07] R. Spence. *Information Visualization: Design for Interaction, Second Edition*
Englewood Cliffs, NJ: Prentice Hall, 2007.
- [T14] Alexandru C. Telea *Data Visualization Principles And Practice, Second Edition*
CRC Press, 2014
- [S05] John C. Sutton *GIS Applications in Transit Planning and Operations: A Review of Current Practice, Effective Applications and Challenges in the USA*
Transportation Planning and Technology, August 2005 Vol. 28, No. 4, pp. 237? / 250
- [DBPEMS15] Fan Du, Joshua Brulé, Peter Enns, Varun Manjunatha and Yoav Segev *MetroViz: Visual Analysis of Public Transportation Data*
https://wiki.cs.umd.edu/cmsc734_f13/images/6/60/MetroViz_final_paper.pdf
ultimo accesso 06/03/2015
- [K07] Thomas J. Kimpel *Data Visualization as a Tool for Improved Decision Making within Transit Agencies Final*
Technical Report TNW2006-14 Research Project Agreement No. 821528-5, February 2007
- [PMAR13] Stefano Pensa, Elena Masala, Maurizio Arnone, Andrea Rosa *Planning local public transport: a visual support to decision-making*
EWGT2013 – 16 th Meeting of the EURO Working Group on Transportation © 2013 The Authors.
Published by Elsevier Ltd
- [H08] Ricardo Hoar *Visualizing Transit Through a Web Based Geographic Information System*
World Academy of Science, Engineering and Technology

International Journal of Environmental, Ecological, Geological and Marine Engineering Vol:2
No:10, 2008

[BCDLMPT13] Benedetto Barabino, Carlino Casari, Roberto Demontis, Cristian Lai, Sara Mozzoni,
Antonio Pintus, Proto Tilocca

A Web Portal for Reliability Diagnosis of Bus Regularity

Proceedings of the 7th International Workshop on Information Filtering and Retrieval co-located
with the 13th Conference of the Italian Association
for Artificial Intelligence (AI*IA 2013), Turin, Italy, December 6, 2013.

[MCLN12] Mahmoud Mesbah , Graham Currie, Claudia Lennon, Trevor Northcott *Spatial and
temporal visualization of transit operations performance data at a network level*

Journal of Transport Geography 25 (2012) 15–26

[BKM09] Aaron Bangor, Philip Kortum, James Miller, *Determining What Individual SUS Scores
Mean: Adding an Adjective Rating Scale*

Journal of Usability Studies Vol. 4, Issue 3, May2009, pp. 114-123

[N15] Netcraft <http://news.netcraft.com/archives/2015/02/24/february-2015-web-server-survey.html>

ultimo accesso 04/03/2015

Appendice A – Servizio giornaliero

Una corsa di una linea è definita dal suo percorso e dai suoi orari di transito.

L'insieme delle corse da esercire in una stessa giornata è detto servizio giornaliero.

1. Copertura del servizio giornaliero

Affinché il servizio giornaliero programmato divenga una realtà fruibile dai clienti del servizio pubblico, è necessario che le corse programmate divengano effettivi transiti di bus alle fermate.

Per raggiungere il massimo contenimento dei costi, deve essere individuato il numero minimo di veicoli da impiegare. Il processo per arrivare a questo sia articolato di alcune fasi.

La prima consiste nel partizionamento dell'insieme delle corse in sottoinsiemi detti viaggi.

Un viaggio è una sequenza di corse $c_1, c_2, c_3, \dots, c_k$ della stessa linea nel quale il nodo finale della corsa c_i coincide con il nodo iniziale della corsa c_{i+1} con $1 < i < k-1$ e l'orario di termine della corsa c_i è minore dell'orario d'inizio della corsa c_{i+1} .

La seconda consiste nel generare una serie di sequenze dei viaggi $v_1, v_2, v_3, \dots, v_k$, creati nel passo precedente, aventi le seguenti caratteristiche:

- ogni viaggio deve appartenere a una sola sequenza;
- per ogni v_i e v_{i+1} appartenenti alla stessa sequenza, v_i deve terminare cronologicamente prima dell'inizio di v_{i+1} ;
- v_{i+1} deve iniziare nello stesso nodo in cui termina v_{i+1} oppure deve esistere una sequenza di uno o più archi a vuoto, che collegano il nodo di termine f di v_i e il nodo d'inizio i di v_{i+1} , la cui somma delle percorrenze sia minore o uguale alla differenza tra il tempo di transito di v_{i+1} in i e di v_i in f . Ad esempio, qualora all'interno della sequenza il tempo intercorrente t tra il termine del viaggio v_i e l'inizio del viaggio v_{i+1} fosse superiore al tempo t_i somma delle percorrenze degli archi a vuoto da e per un qualche nodo di sosta, si può considerare di portare il bus in sosta su quel nodo per il tempo $t - t_i$;
- poiché ogni bus esce da un deposito per entrare in servizio e rientra in un deposito al termine del servizio, nella sequenza da generare il viaggio v_i dovrà seguire un arco a vuoto che ha come nodo origine un nodo di sosta m e come nodo destinazione il

nodo iniziale di v_l e v_k sarà seguito da un arco a vuoto che ha come nodo origine il nodo di termine del viaggio v_k e come nodo destinazione un nodo di sosta N . Di solito m ed n sono lo stesso nodo.

Le possibili serie sequenze generabili sono diverse. Per scegliere la più adeguata, un algoritmo stima il costo di ognuna di esse secondo alcuni parametri e individua la più conveniente.

2. Turno macchina

Ogni sequenza prende il nome di turno macchina ed ha un numero che la identifica univocamente.

Il parco veicolare dell'azienda è costituito da modelli diversi di veicoli. Per comodità l'insieme dei mezzi adibiti al trasporto pubblico è stato suddiviso in tipologie. I veicoli appartenenti alla stessa tipologia condividono alcune caratteristiche ad esempio hanno la medesima lunghezza o tipo di alimentazione o abilitazione d'impiego.

Una delle informazioni indicate sul turno macchina è la tipologia di veicolo da impiegare. La scelta dalla tipologia dipende da alcuni fattori tra i quali il previsto numero massimo di viaggiatori da trasportare e le corse effettuare.

Esistono un certo numero di depositi, ognuno con una propria dotazione di veicoli. Ogni giorno è realizzato dal personale del deposito l'abbinamento tra turno macchina e la matricola di un bus disponibile. Il bus scelto deve essere di una tipologia compatibile con quella indicata sul turno macchina.