

**TOR VERGATA**  
UNIVERSITÀ DEGLI STUDI DI ROMA

FACOLTÁ DI SCIENZE MATEMATICHE, FISICHE E NATURALI

---

Dipartimento di Fisica

Tesina di Complex and Neural Networks

## Studio dei flussi aerei e dei flussi ferroviari italiani

Prof. Gaetano Salina

Gruppo 1:  
Chiara Calascibetta  
Maria Elena Perruzza  
Martino Scarpolini  
Matteo Veglianti

---

ANNO ACCADEMICO 2019/2020

# Indice

<b>1 FLUSSI AEROPORTUALI</b>	<b>3</b>
1.1 Analisi qualitativa . . . . .	3
1.2 Analisi dei flussi in ingresso per ogni aeroporto . . . . .	5
1.2.1 Probabilità in entrata . . . . .	5
1.2.2 Mappa dei link entranti in ogni nodo . . . . .	5
1.3 Analisi dei flussi in uscita per ogni aeroporto . . . . .	6
1.3.1 Probabilità in uscita . . . . .	6
1.3.2 Mappa dei link uscenti in ogni nodo . . . . .	6
1.4 Analisi quantitativa . . . . .	7
1.4.1 Matrice dei collegamenti in entrata . . . . .	8
1.4.2 Matrice dei collegamenti in uscita . . . . .	9
<b>2 FLUSSI FERROVIARI</b>	<b>10</b>
2.1 Analisi qualitativa . . . . .	10
2.1.1 Analisi dei flussi dei Freccia-Rossa . . . . .	10
<b>3 IMPLEMENTAZIONE DELLE RETI NELLA RETE GLOBALE</b>	<b>12</b>
<b>4 Conclusioni</b>	<b>13</b>

# Introduzione

Questo lavoro è finalizzato allo studio della diffusione di un'epidemia su una rete globale che comprende diversi layer, tra i quali sottoreti relative al flusso aereo e ferroviario, su cui ci soffermiamo. Dunque, per simulare l'espansione dell'epidemia, sono stati analizzati, mediante un codice scritto con il linguaggio di programmazione "Python", i flussi aerei e i flussi ferroviari tra le province italiane e ne è stata data una rappresentazione sia in modo quantitativo che qualitativo.

\* PARTE 1:

Nella prima parte è stata creata una rete che ha come nodi gli aeroporti italiani. Tra di essi sono stati associati dei link pesati in base al flusso di passeggeri annuali riferito all'anno 2017. Successivamente i link sono stati normalizzati in base ad un processo di normalizzazione che spiegheremo più avanti.

Con questi dati sono state poi create tre mappe (1.2, 1.3, 1.4) e tre matrici (1.5, 1.6, 1.7) .

\* PARTE 2:

Nella seconda parte è stata creata una rete analoga, per le stazioni ferroviarie della linea alta velocità della Freccia Rossa.

# 1 FLUSSI AEROPORTUALI

Il flusso aereo è stato studiato inizialmente tra l'Italia e la Cina, ma per mancanza di dati lo studio dettagliato è stato proseguito solo per l'Italia. Per i voli internazionali è stato trovato online un database contenente tutte le tratte a livello mondiale. In base ad esso è stata creata una rete, inserendo link tra nodi collegati da tratte trovate nel database. In questo database, inoltre, le tratte erano ordinate anche per compagnia aerea. Una prima stima del flusso di passeggeri poteva essere proporzionale al numero di tratte trovate tra due nodi. Si è preferito, però, per via dell'incertezza sulla qualità delle informazioni sul numero di tratte, restringere lo studio dettagliato dei flussi all'interno della sola Italia.

## 1.1 Analisi qualitativa

In questa sezione vengono realizzate le mappe relative alle tratte aeroportuali italiane mediante l'utilizzo dei dati estratti da un database trovato in rete. Per l'analisi sono stati importati due file: "airports.dat" per gli aeroporti e "routes.dat" per le rotte, dal sito "[openflights.org](http://openflights.org)".

Il file "airports.dat" fornisce per ogni aeroporto, tra le varie informazioni, il nome, la città in cui è situato, i codici IATA e ICAO e la latitudine e longitudine. Il file "routes.dat" fornisce per ogni tratta aeroportuale le città di partenza e di destinazione ed i codici IATA e ICAO degli aeroporti di partenza e di destinazione.

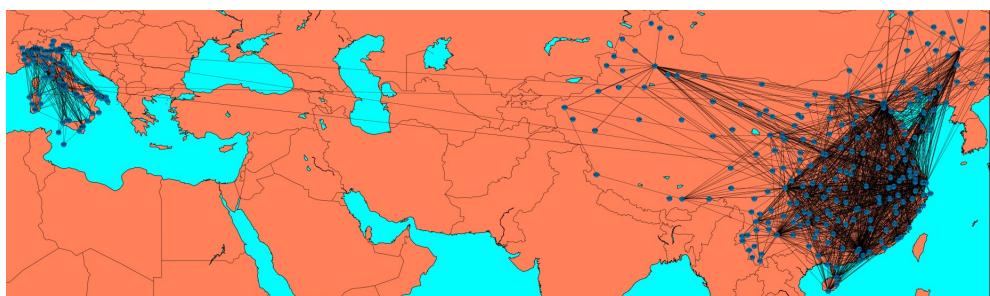


Figura 1.1: Rete dei flussi tra Italia e Cina. Lo spessore dei link è uniforme e non sono stati usati criteri per stimare il flusso tra aeroporti.

Successivamente, mediante un'analisi di scraping, sono stati estrapolati da un [file](#) dell'ENAC i dati relativi ai flussi aeroportuali italiani interni. Questo passaggio ha richiesto delle modifiche manuali in quanto i flussi nel file dell'ENAC erano raggruppati per città, invece che per aeroporto,

come nel primo database.

E' stata creata allora la mappa con gli aeroporti italiani e le rotte. Nella

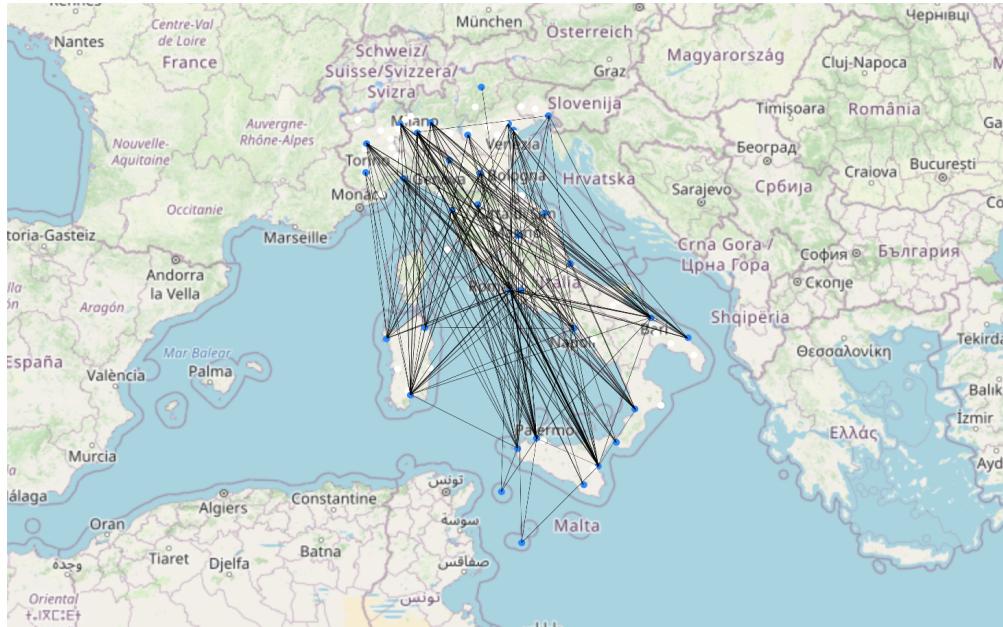


Figura 1.2: Mappa degli aeroporti italiani. Lo spessore dei link è proporzionale al flusso tra i nodi.  
[Link github](#)

figura 1.2 sono mostrati i link tra i vari aeroporti in cui lo spessore è indice del flusso tra gli stessi. La differenza di colore è stata inserita per poter differenziare gli aeroporti con tratte ( in blu) e quelli senza (in bianco); questi ultimi fanno riferimento ad aeroporti non aperti al traffico civile.

## 1.2 Analisi dei flussi in ingresso per ogni aeroporto

In questa sezione viene sviluppata la mappa dei link entranti per ogni nodo mediante una analisi delle rispettive probabilità.

### 1.2.1 Probabilità in entrata

La probabilità per un nodo B di essere raggiunto da un nodo A verrà calcolata come il rapporto tra il flusso tra A e B e la somma dei flussi estesa a tutti i nodi "primi vicini" entranti in B:

$$P_{\text{entrata}:A \rightarrow B} = \frac{F_{A \rightarrow B}}{\sum_{C \in V_B} F_{C \rightarrow B}}$$

### 1.2.2 Mappa dei link entranti in ogni nodo

Per poter creare la mappa dei link entranti è stata fatta una analisi specifica per ogni nodo. Sono state quindi contate tutte le tratte aveni come destinazione l'aeroporto in questione, visualizzandole sulla mappa come link di vario spessore. Lo spessore è stato pesato con la probabilità P descritta in precedenza.

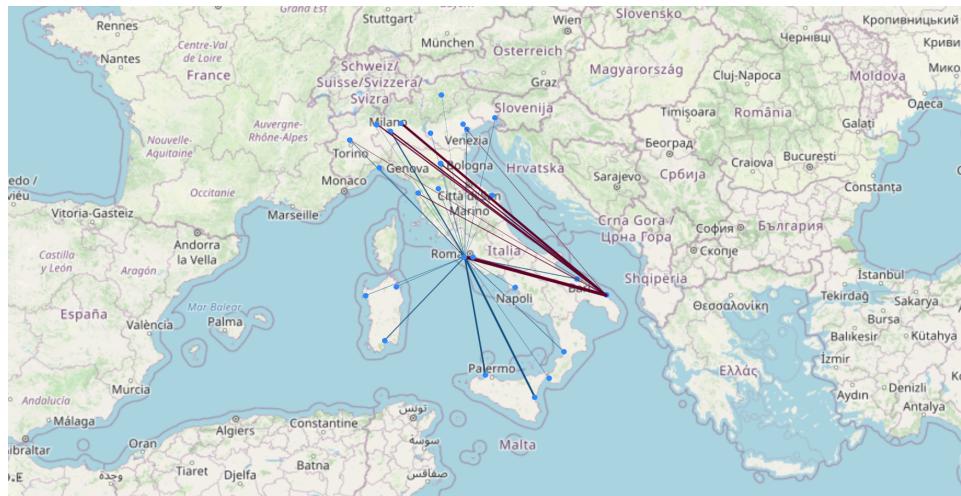


Figura 1.3: Mappa degli aeroporti italiani. Lo spessore dei link è proporzionale alla probabilità in entrata per ogni nodo. [Link github](#)

Nella figura 1.3 sono mostrati i link tra i vari aeroporti che arrivano nel nodo in questione in cui lo spessore è indice del flusso tra gli stessi.

## 1.3 Analisi dei flussi in uscita per ogni aeroporto

In questa sezione viene sviluppata la mappa dei link uscenti per ogni nodo mediante una analisi delle rispettive probabilità.

### 1.3.1 Probabilità in uscita

La probabilità per un nodo A di raggiungere un nodo B verrà calcolata come il rapporto tra il flusso tra A e B e la somma dei flussi estesa a tutti i nodi "primi vicini" uscenti da A:

$$P'_{uscita:A \rightarrow B} = \frac{F_{A \rightarrow B}}{\sum_{C \in V_A} F_{A \rightarrow C}}$$

### 1.3.2 Mappa dei link uscenti in ogni nodo

Per l'analisi dei link uscenti vengono contate tutte le tratte aventi come partenza l'aeroporto in questione, visualizzandole sulla mappa come link di vario spessore. Lo spessore è stato pesato con la probabilità P' descritta in precedenza.

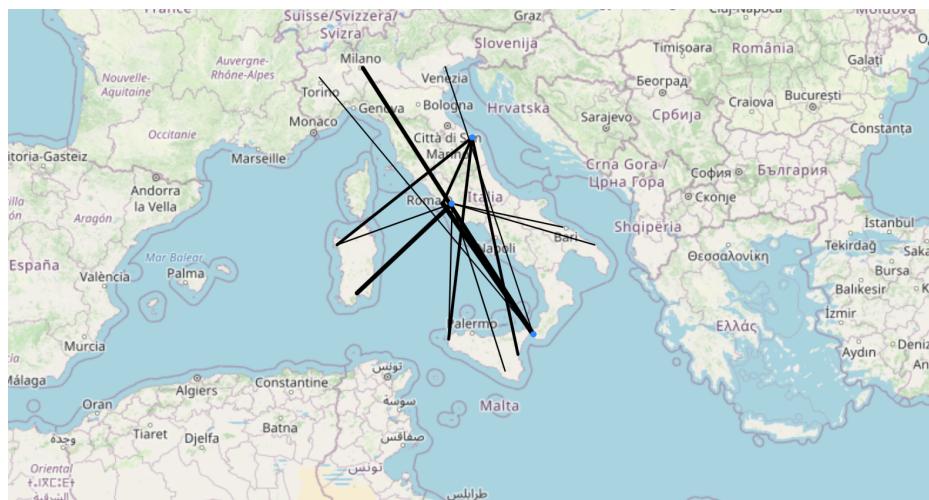


Figura 1.4: Mappa degli aeroporti italiani. Lo spessore dei link è proporzionale alla probabilità in uscita per ogni nodo.[Link github](#)

Nella figura 1.4 sono mostrati i link in cui l'aeroporto in questione è la partenza. Anche in questo caso lo spessore è indice del flusso tra gli stessi.

## 1.4 Analisi quantitativa

In questa sezione vengono costruite le matrici relative ai flussi aerei raggruppati per regioni o per province italiane. In seguito verranno riportate le immagini relative al solo raggruppamento per regioni per motivi pratici di visualizzazione (matrici più piccole e facilmente visualizzabili). Informazioni analoghe per le province sono riportate in file csv in [Link github](#). In questi file csv le province sono identificate dal proprio codice istat, le regioni dal codice istat del proprio capoluogo.

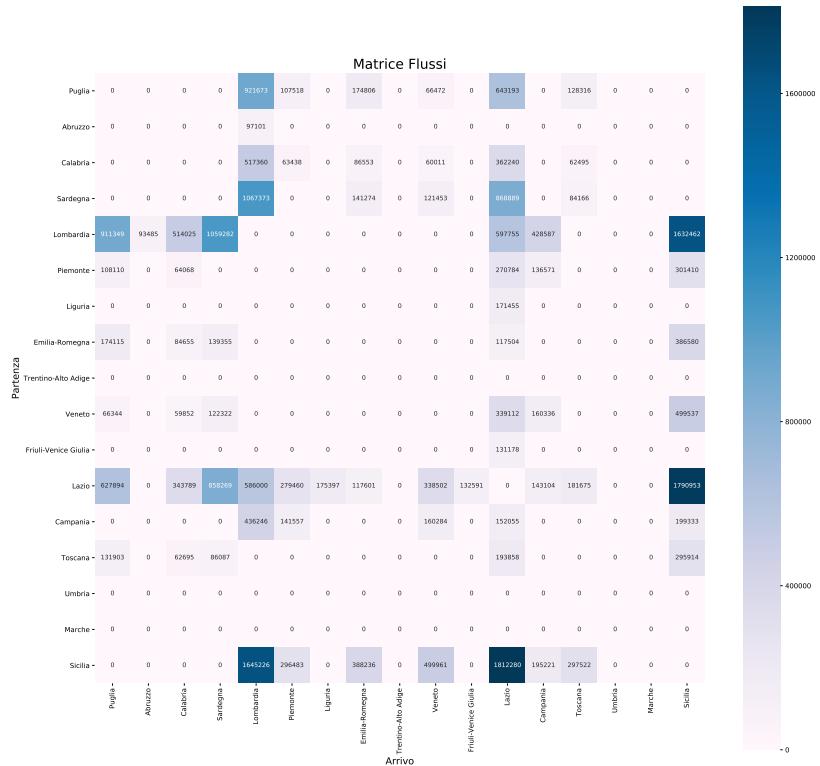


Figura 1.5: Matrice dei flussi aeroportuali tra le varie regioni italiane. Qui i flussi non sono normalizzati ed i valori rappresentano il numero di passeggeri annuali tra le varie regioni, riferito all'anno 2017. I dati sono presi dal file dell'ENAC descritto nel testo.

### 1.4.1 Matrice dei collegamenti in entrata

Per la costruzione della matrice dei collegamenti in entrata, fissata una regione (o provincia) di arrivo, sono state analizzate le regioni (o province) ad essa collegate. La matrice, quindi, fornisce i valori delle probabilità di partire da una regione (o provincia) B, per arrivare ad una regione (o provincia) A fissata. Se la matrice viene impostata con la partenza sulle righe e l'arrivo sulle colonne, si può dire che in questo caso la matrice sarà normalizzata per colonne.

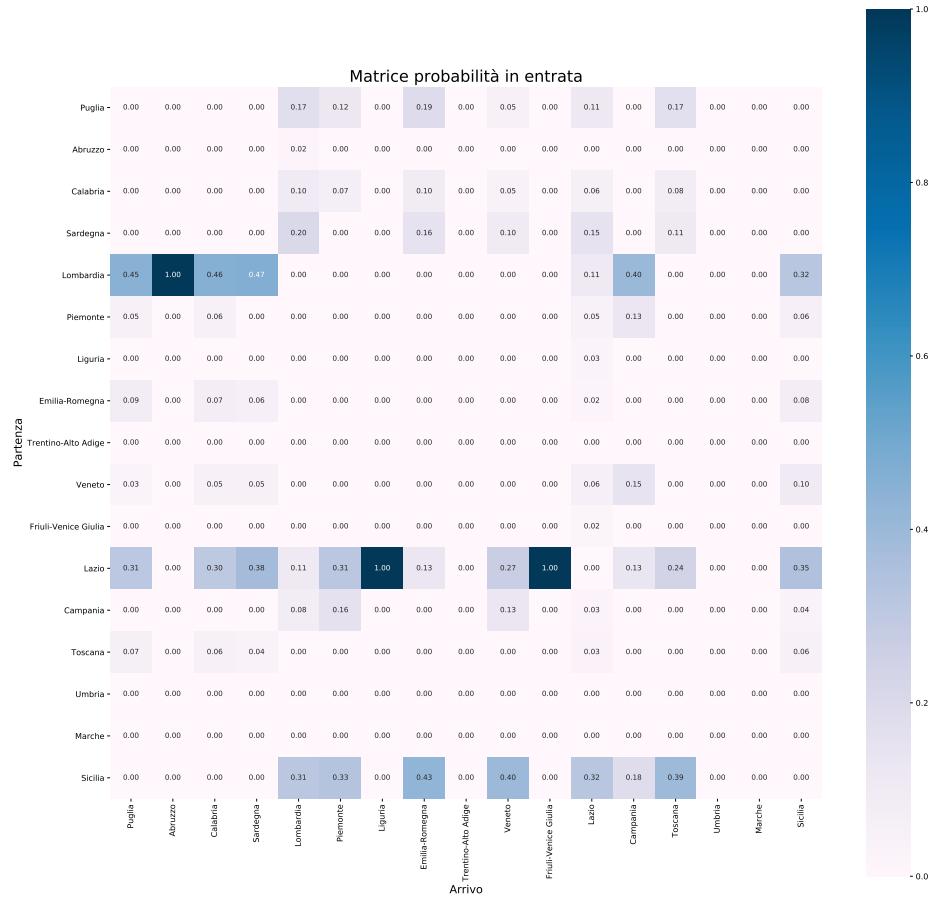


Figura 1.6: Matrice dei flussi aeroportuali tra le varie regioni italiane. Qui i flussi sono normalizzati e rappresentano le probabilità in entrata per ogni regione.

### 1.4.2 Matrice dei collegamenti in uscita

Per la costruzione della matrice dei collegamenti in uscita è stata fatta un'analisi analoga a quella per i collegamenti in entrata. Qui, però, ad essere fissata è la regione (o provincia) di partenza. La matrice, quindi, fornisce i valori delle probabilità di arrivare in una regione (o provincia) B, partendo da una regione (o provincia) A fissata. Se la matrice viene impostata con la partenza sulle righe e l'arrivo sulle colonne, si può dire che in questo caso la matrice sarà normalizzata per righe.

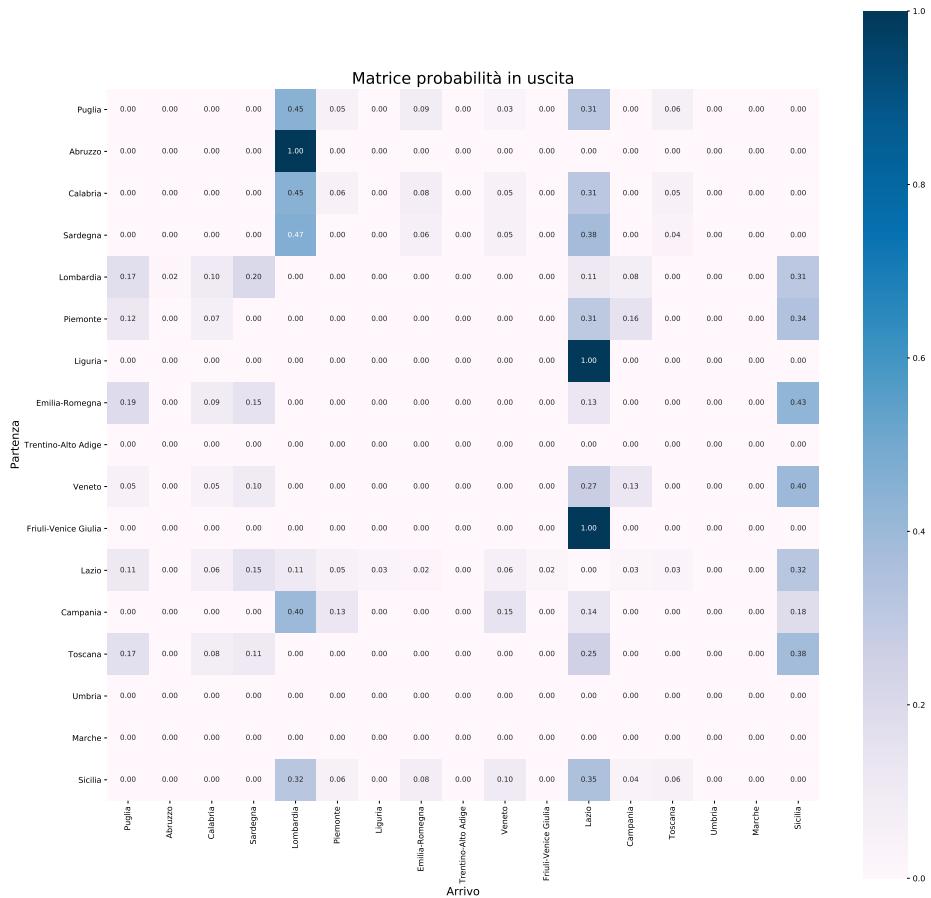


Figura 1.7: Matrice dei flussi aeroportuali tra le varie regioni italiane. Qui i flussi sono normalizzati e rappresentano le probabilità in uscita per ogni regione.

## 2 FLUSSI FERROVIARI

### 2.1 Analisi qualitativa

In questa sezione vengono studiate le tratte ferroviarie mediante l'utilizzo di file trovati in rete sul sito della compagnia Trenitalia. I dati sono stati estratti da un [file](#) (riportato in figura 2.1) messo a disposizione annualmente da Trenitalia. Quest'ultimo contiene gli orari e le fermate di ogni singolo treno della compagnia che viaggia all'interno del suolo italiano. Per ottenere dati utilizzabili nel codice è stato analizzato e trascritto

26

QUADRO ORARIO 1															
	FR 9501	FR 9581	FR 9691	FR 9503	FR 9505	FR 9507	FR 9591	FR 9595	FR 9597	FR 9601	FR 9303	FR 9603	FR 9511	FR 9605	FR 9607
Brescia	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Torino Porta Nuova	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Torino P. Susa	-	-	-	-	-	-	-	-	-	-	-	-	-	-	06:50
Rho Fiera	-	-	-	-	-	-	-	-	-	-	-	-	-	-	06:50
Milano Porta Garibaldi a	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Milano Porta Garibaldi b	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Milano Centrale	a	-	-	-	-	-	-	-	-	-	-	-	-	-	06:50
Milano Rogoredo	-	-	-	-	-	-	-	-	-	-	-	-	-	-	07:10
Piacenza	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Parma	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Pistoia Emilia	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Poggio Emilia AV	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Modena	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Bologna Centrale	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Firenze S.M.Novella	a	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Firenze S.M.Novella	-	-	-	-	-	-	x	06:00	06:59	x	07:05	07:14	-	08:59	-
Perugia	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Roma Tiburtina	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Roma Termini	a	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Roma Termini	-	-	-	-	-	-	-	-	-	-	-	-	-	-	10:11
Napoli Afragola	a	07:00	07:00	07:00	07:25	08:00	09:15	-	-	09:15	08:53	-	09:25	10:00	09:53
Napoli Afragola	-	07:58	07:58	07:58	08:33	09:13	10:28	-	-	10:28	10:03	-	10:33	11:12	11:03
Napoli Centrale	a	08:12	08:12	08:33	x	09:13	10:28	-	-	10:50	-	-	11:25	-	-
Napoli Centrale	-	-	-	-	-	-	-	-	-	11:27	-	-	12:00	-	-
Salerno	a	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Salerno	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Destinazione	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

ⓘ Per l'offerta completa delle Freccce che percorrono almeno un tratto Bologna e Salerno vedi quadro 9  
 ⓘ Stagione bin 15/06/2011 - 15/09/2011  
 ⓘ Ferma anche a Arezzo (06:36)  
 ⓘ Si effettua fino al 7/1. Sospeso nei festivi fino al 6/1.  
 ⓘ Ferma anche a Arezzo (07:36). Chiusi-Chianciano T. (08:02)  
 ⓘ Si effettua dal 8/1. Sospeso nei festivi dal 12/1.  
 ⓘ Ferma anche a Arezzo (07:30)  
 ⓘ Sospeso nei festivi e nei lavorativi pre-festivi.

ⓘ For complete offer of Freccce that run at least one stretch between Bologna and Salerno see timetable 9  
 ⓘ Stagione bin 15/06/2011 - 15/09/2011  
 ⓘ Stopped at points 15 thru 19  
 ⓘ Also stops in Arezzo (06:36)  
 ⓘ Runs until 7/1. Suspended service on public holidays until 6/1.  
 ⓘ Also stops in Arezzo (07:36). Chiusi-Chianciano T. (08:02)  
 ⓘ Runs from 8/1. Suspended service on public holidays from 12/1.  
 ⓘ Also stops in Arezzo (07:30)  
 ⓘ Suspended service on public holidays and on working days before public holidays

Figura 2.1: Esempio di quadro orario che è stato trascritto manualmente in file csv. [Link al file](#).

manualmente ogni quadro presente nel file. Sono stati costruiti dei file csv guardando singolarmente ogni orario in quanto un'analisi di scraping non è stata possibile per il formato del PDF. Ogni file costruito ha tre colonne: città di partenza, città di arrivo e numero di treni giornalieri. Sono stati analizzati solamente i treni Freccia-Rossa.

In figura 2.2 è stata riportata una visualizzazione del risultato finale della trascrizione dei quadri orari.

#### 2.1.1 Analisi dei flussi dei Freccia-Rossa

Dal file csv costruito contenente ogni singola tratta è stata svolta un'analisi analoga a quella svolta per i flussi aeroportuali. I nodi della rete ed i rispettivi link sono stati raggruppati per regioni e per province. Come nel caso dei flussi aeroportuali il raggruppamento per province non ha cambiato quasi nulla rispetto alla struttura originale del file csv, in quanto le stazioni ferroviarie servite dalla Freccia-Rossa sono quasi sempre a

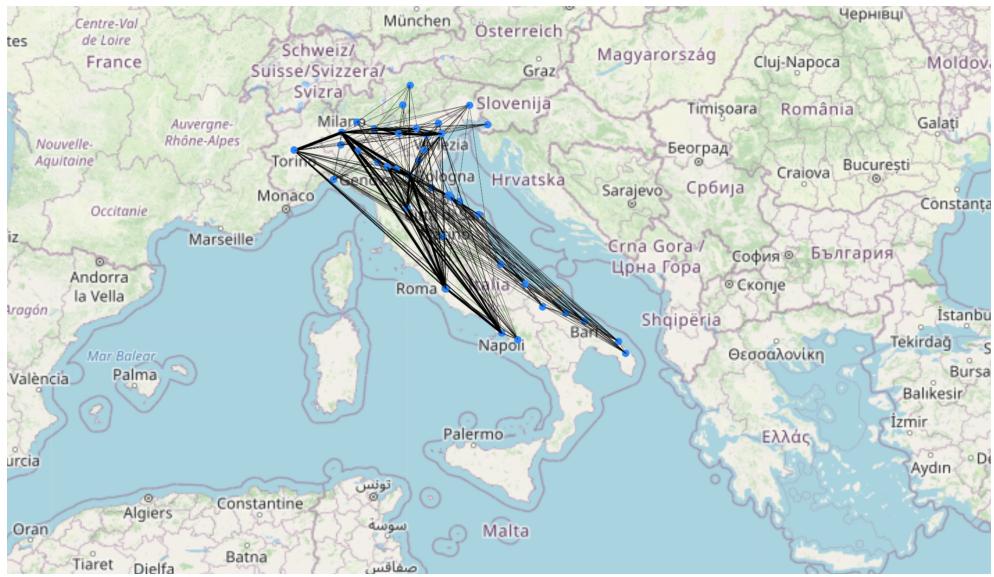


Figura 2.2: Collegamenti Freccia-rossa tra le città. Lo spessore di ogni link è proporzionale al numero di treni al giorno in quella specifica tratta. [Link github del file per realizzare la mappa.](#)

livello provinciale. Per ognuno dei raggruppamenti è stata svolta l'analisi delle probabilità in entrata ed in uscita, come per i flussi aeroportuali, [Link github](#). Nelle didascalie delle figure sono riportati i link su github dei file con cui è stata effettuata l'analisi.

### 3 IMPLEMENTAZIONE DELLE RETI NELLA RETE GLOBALE

Come detto, questo lavoro è finalizzato allo studio della diffusione di un'epidemia su una rete globale. Questo studio è stato fatto generalizzando il modello SIR tramite le caratteristiche topologiche della rete stessa e studiando il processo diffusivo come una catena di Markov.

La rete globale è costruita dall'unione di diversi layer, dove ognuno di essi rappresenta un diverso livello di dettaglio delle connessioni tra i nodi.

In totale si sono ottenuti 177918 nodi e 1020869 link. Si hanno, inoltre, 7904 sotto-reti (Comuni) connesse con 6 reti (layer) distinte e opportunamente pesate:

1. Connessione tra Comuni con una matrice di contiguità a primi vicini;
2. Connessione dei Comuni con una matrice gerarchica con il Capoluogo di Provincia;
3. Connessione tra Province a primi vicini;
- 4. Rete Ferroviaria (AV);**
- 5. Rete Aerea;**
6. Rete generata da flussi economici (4 Classi Ateco).

Infine, sulla rete globale è stato anche studiato l'effetto del lockdown in funzione dell'inibizione dei diversi layer. Per dettagli più approfonditi si rimanda a "*Effetto del Lockdown sullo spread epidemico su rete*", *F.Fossella, M.Lacal, C.Marin, S.Ascione, F.Monnati, E.Mariani*.

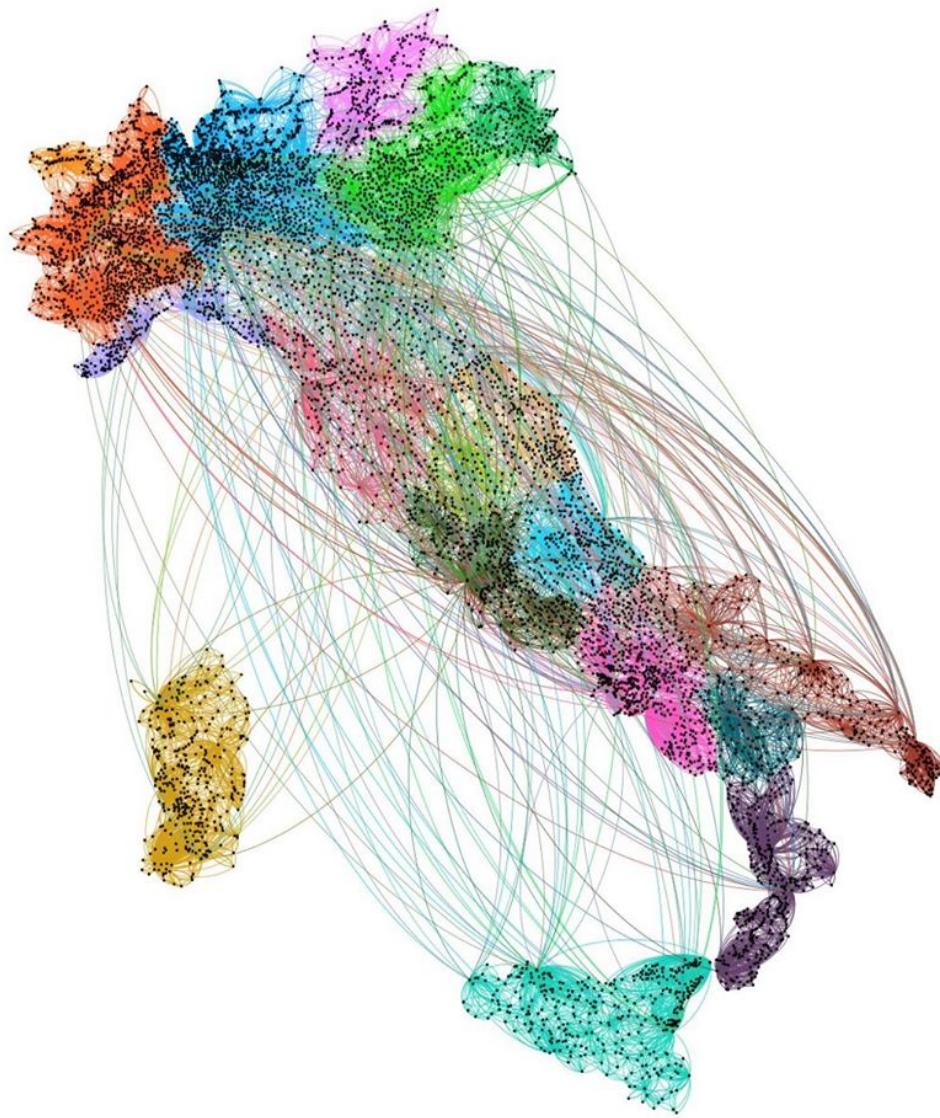


Figura 3.1: Semplice rappresentazione della rete globale con tutti i layer di connessione ugualmente pesati.

## 4 Conclusioni

Le Reti Aeree e ferroviarie (AV) sono state costruite a partire da dati con una diversa struttura. Se è stato possibile analizzare direttamente il flusso di passeggeri aereo messo a disposizione dell'ENAC, altrettanto non si può dire per il flusso ferroviario. In questo secondo caso, è stato possibile dare una stima del flusso solo tramite il conteggio del numero di tratte e non tramite il numero di passeggeri. Abbiamo quindi solo una

rappresentazione indiretta del dato che vorremmo utilizzare. Pertanto, bisogna tenere in considerazione che le reti ottenute sono parziali e che possono essere affette da una sistematica non banale.

Inoltre, il flusso di passeggeri aerei, seppur correttamente a disposizione, rimane significativo solo per gli spostamenti da nord a sud e quindi ricopre un tasso di influenza minore tra le regioni contigue.

In definitiva, bisogna tenere in considerazione che nella simulazione finale sarà presente questa disomogeneità nella costruzione dei pesi dei vari link delle due reti.

Si potrebbe pensare di fare di meglio nel caso in cui si avessero a disposizione i dati relativi al flusso di passeggeri per ogni tipologia di mezzo (Ferrovie (AV) e non, tratte aree e autostradali, autobus, traghetti ecc.), dando maggior peso al mezzo più utilizzato in relazione alla tratta.

## Elenco delle figure

1.1	Rete dei flussi tra Italia e Cina. Lo spessore dei link è uniforme e non sono stati usati criteri per stimare il flusso tra aeroporti. . . . .	3
1.2	Mappa degli aeroporti italiani. Lo spessore dei link è proporzionale al flusso tra i nodi. <a href="#">Link github</a> . . . . .	4
1.3	Mappa degli aeroporti italiani. Lo spessore dei link è proporzionale alla probabilità in entrata per ogni nodo. <a href="#">Link github</a> . . . . .	5
1.4	Mappa degli aeroporti italiani. Lo spessore dei link è proporzionale alla probabilità in uscita per ogni nodo. <a href="#">Link github</a> . . . . .	6
1.5	Matrice dei flussi aeroportuali tra le varie regioni italiane. Qui i flussi non sono normalizzati ed i valori rappresentano il numero di passeggeri annuali tra le varie regioni, riferito all'anno 2017. I dati sono presi dal file dell'ENAC descritto nel testo. . . . .	7
1.6	Matrice dei flussi aeroportuali tra le varie regioni italiane. Qui i flussi sono normalizzati e rappresentano le probabilità in entrata per ogni regione. . . . .	8
1.7	Matrice dei flussi aeroportuali tra le varie regioni italiane. Qui i flussi sono normalizzati e rappresentano le probabilità in uscita per ogni regione. . . . .	9
2.1	Esempio di quadro orario che è stato trascritto manualmente in file csv. <a href="#">Link al file</a> . . . . .	10
2.2	Collegamenti Freccia-rossa tra le città. Lo spessore di ogni link è proporzionale al numero di treni al giorno in quella specifica tratta. <a href="#">Link github del file per realizzare la mappa</a> . . . . .	11
3.1	Semplice rappresentazione della rete globale con tutti i layer di connessione ugualmente pesati. . . . .	13