

Esame Complex & Neural Networks

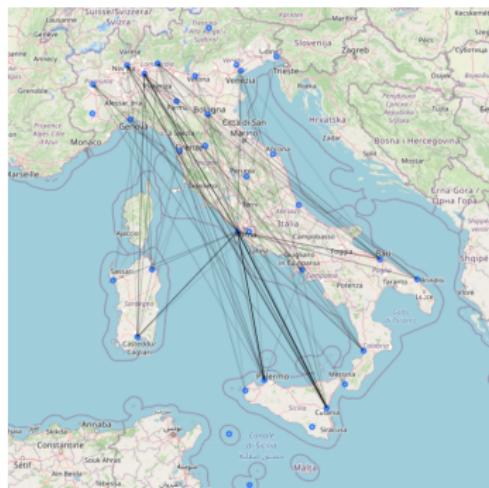
Chiara Calascibetta,
Martino Scarpolini,
Maria Elena Perruzza,
Matteo Veglianti

25 Giugno 2020

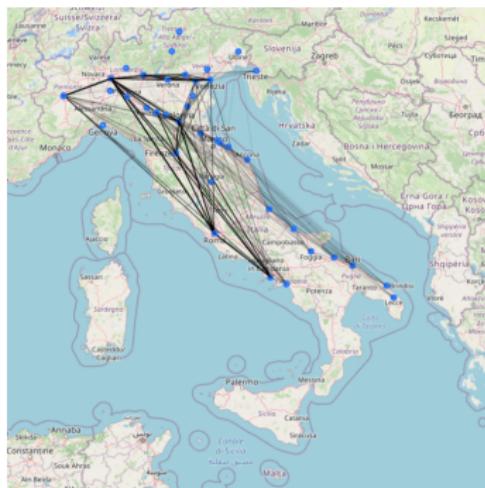
Introduzione al progetto

Studio dei flussi aeroportuali e ferroviari italiani

Realizzazione di due reti basate su i collegamenti aerei e ferroviari sul suolo italiano. Lo scopo sarà aggiungere queste due reti ad una rete globale italiana e simulare l'effetto del lockdown sulla diffusione di un'epidemia in Italia.



(a) Rete aeroportuale



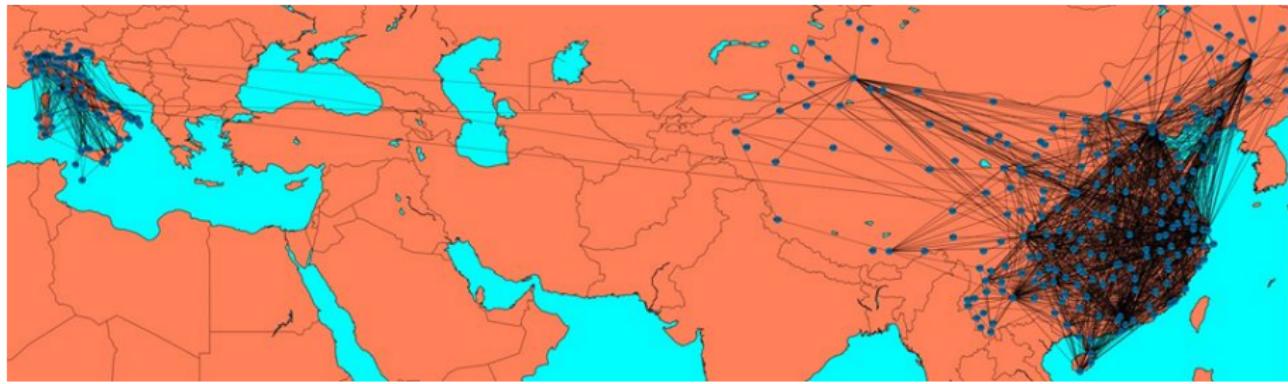
(b) Rete ferrovaria

Database usati

- **Database 0** - Aeroporti e tratte da www.openflights.org ([link](#))
- **Database 1** - File ISTAT con comuni italiani ([link](#))
- **Database 2** - Riepilogo flussi aeroportuali italia ENAC ([link](#))
- **Database 3** - Quadro orari Trenitalia ([link](#))
- **Archivio del progetto** - tutti i codici ed i documenti sono stati caricati su una pagina [github](#) ([link](#)) e su un sito ([link](#))

Idea iniziale: studio dei flussi aerei Italia - Cina

Realizzazione di una rete, tramite Database 0, basata sui collegamenti aerei (diretti e non) tra l'Italia e la Cina. Tuttavia, i link rappresentati nella mappa sono privi di un peso: non è stato possibile, per mancanza di dati, stimare il flusso di passeggeri per ogni tratta.
Questo rende la rete poco significativa, dunque nel corso del lavoro ci concentriamo unicamente sul suolo italiano.



Database 2 (ENAC): estrazione dei flussi di passeggeri annuali anno 2017

Only routes with more than 50.000 pax/year

Route Origin-Destination	Passengers (no.)
Catania	1.016.262
Roma Fiumicino	991.516
Roma Fiumicino	799.435
Palermo	796.016
Milano Linate	597.754
Roma Fiumicino	585.999
Catania	477.510
Milano Malpensa	473.521
Cagliari	467.938
Roma Fiumicino	457.099
Bari	380.980
Roma Fiumicino	370.653
Cagliari	315.440
Milano Linate	311.859
Catania	295.959
Milano Linate	289.850
Roma Fiumicino	279.460
Torino	270.784
Brindisi	262.211
Napoli	258.545
Roma Fiumicino	257.239
...	...

Tramite un'analisi di scraping è stato convertito in formato csv il database 2 trovato sul sito dell'ENAC. I nomi degli aeroporti sono stati convertiti nei nomi delle città relative e nel caso di doppie corrispondenze è stato considerato il flusso totale.

Database 3 (Trenitalia): estrazione del numero di tratte giornaliere nell'anno 2019-2020

Per la rete ferroviaria non sono stati trovati i dati sui flussi di passeggeri da stazione a stazione. Per questo motivo è stato usato il database 3 contenente gli orari di ogni treno in ogni stazione italiana. È stato necessario estrarre a mano il numero di treni totali tra ogni coppia di stazioni e dunque ci siamo limitati ai soli treni Frecciarossa.

QUADRO ORARIO 1%													
	FR 9690 B	FR 9326	FR 9692	FR 9590 B	FR 9693 B	FR 9598	FR 9330	FR 9594 B	FR 9594 B	FR 9672	FR 9568 B	FR 9574 B	
Provenienza	-	-	-	-	-	-	-	-	-	-	-	-	-
Salerno	a	-	-	-	-	-	-	-	-	-	-	-	-
Bologna	-	-	-	-	-	-	-	-	-	-	-	-	-
Napoli Centrale	a	-	-	-	-	-	-	-	-	-	-	-	-
Napoli Centrale	-	16.55	17.25	17.40	18.26	18.30	-	-	-	18.30	18.41	-	20.39
Napoli Centrale	-	-	-	17.25	17.40	18.26	-	-	-	18.40	18.50	-	-
Napoli Afragola	a	-	17.10	-	17.55	-	-	-	-	18.55	18.58	-	-
Roma Termini	a	-	18.05	18.35	18.55	19.40	-	-	-	19.55	20.40	20.55	21.40
Roma Tiburtina	a	18.20	-	-	-	-	-	-	-	19.55	20.40	20.55	21.40
Roma Tiburtina	-	18.35	-	19.20	19.40	19.55	19.60	19.65	19.70	20.20	21.20	21.35	21.40
Roma Tiburtina	-	-	-	-	-	-	-	-	-	20.00	20.05	20.20	-
Roma Tiburtina	-	-	-	-	-	-	-	-	-	-	-	-	-
Firenze S.M.Novella	a	-	-	20.01	-	20.46	-	-	-	21.27	21.46	21.46	-
Firenze S.M.Novella	-	-	-	20.10	-	20.55	-	-	-	21.30	21.65	21.65	-
Bologna Centrale	-	12.02	12.05	12.53	-	12.58	-	-	-	12.20	12.38	12.38	-
Milano Centrale	-	-	-	-	-	-	-	-	-	22.40	-	-	-
Reggio Emilia AV	-	-	21.15	-	22.00	-	-	-	-	23.00	23.00	-	-
Reggio Emilia AV	-	-	-	-	-	-	-	-	-	22.40	-	-	-
Parmigiano	-	-	-	-	-	-	-	-	-	23.14	-	-	-
Parmigiano	-	-	-	-	-	-	-	-	-	23.41	-	-	-
Milano Rogoredo	a	-	-	21.55	21.49	22.36	-	-	-	23.39	23.39	23.39	-
Milano Centrale	a	21.30	-	-	22.00	22.50	22.55	22.55	00.15	23.50	23.50	00.00	-
Milano Centrale	-	-	-	-	22.10	-	23.07	-	-	-	-	-	-
Milano Porta Garibaldi	-	-	-	-	-	-	-	-	-	-	-	-	-
Milano Porta Garibaldi	-	-	-	-	-	-	-	-	-	-	-	-	-
Torino P. Susa	a	-	-	23.08	22.59	-	-	-	00.01	-	-	-	-
Torino Porta Nuova	-	-	-	-	23.20	23.50	-	-	00.01	-	-	-	-
Brescia	-	-	-	-	-	-	-	-	-	-	-	-	-

● Per l'offerta completa delle Freccie che percorrono almeno un tratto tra Salerno e Bologna vedi quadro 9

● A Ferrara bin. 16-18

● Si effettua solo nei giorni pre-festivi nonché il 25/12, il 5/1, il 12/4 e il 25/4.

● Si effettua dal lunedì al giovedì nonché il 15/12, il 5/1, il 12/4 e il 25/4.

● Si effettua dal lunedì alle domeniche a partire dal 20/12.

● Per complete offer of Freccie that run at least one stretch between Salerno and Bologna see timetable 9

● At Ferrara bin. 16-18

● Offer suspended on working days before public holidays as well as on

25/12, 5/1, 12/4 and on 25/4

● Offer runs from Monday to Thursday as well as on 15/12.

Figure: Esempio di quadro orario del database 3

Struttura delle reti

Con i database appena importati sono state create in totale 4 reti (2 aeroportuali e 2 ferroviarie), costruite su due diversi livelli: regionale e provinciale. A tale scopo, i database 2 e 3 sono stati utilizzati per costruire i link tra i nodi del database 1 messo a disposizione dall'ISTAT.

Quest'ultimo consiste in un file "csv" in cui ogni riga è composta da:

- codice istat
- nome della provincia
- coordinate geografiche
- regione di appartenenza

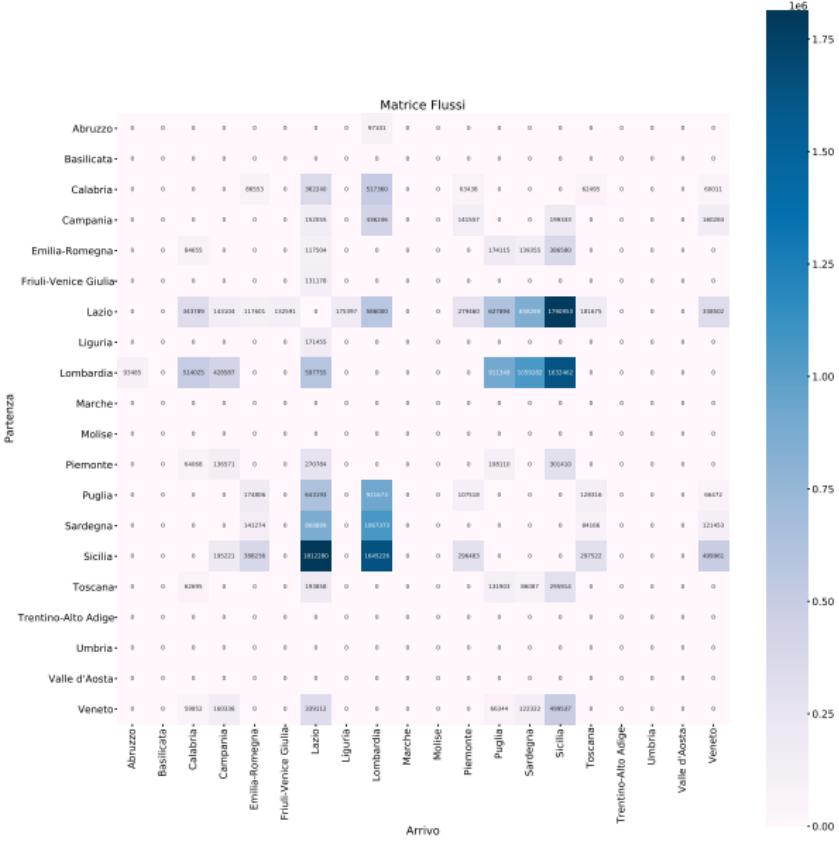


Figure: Matrice aeroportuale con colorazione relativa al flusso. (link)

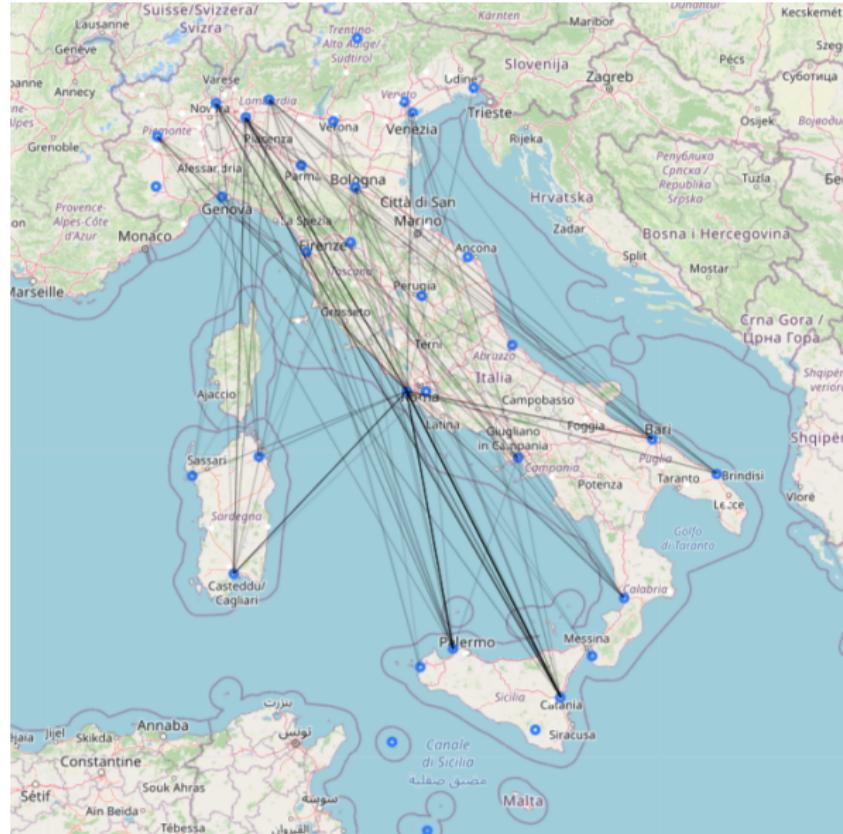


Figure: Rete aeroportuale con spessore dei link proporzionale al flusso. (link)

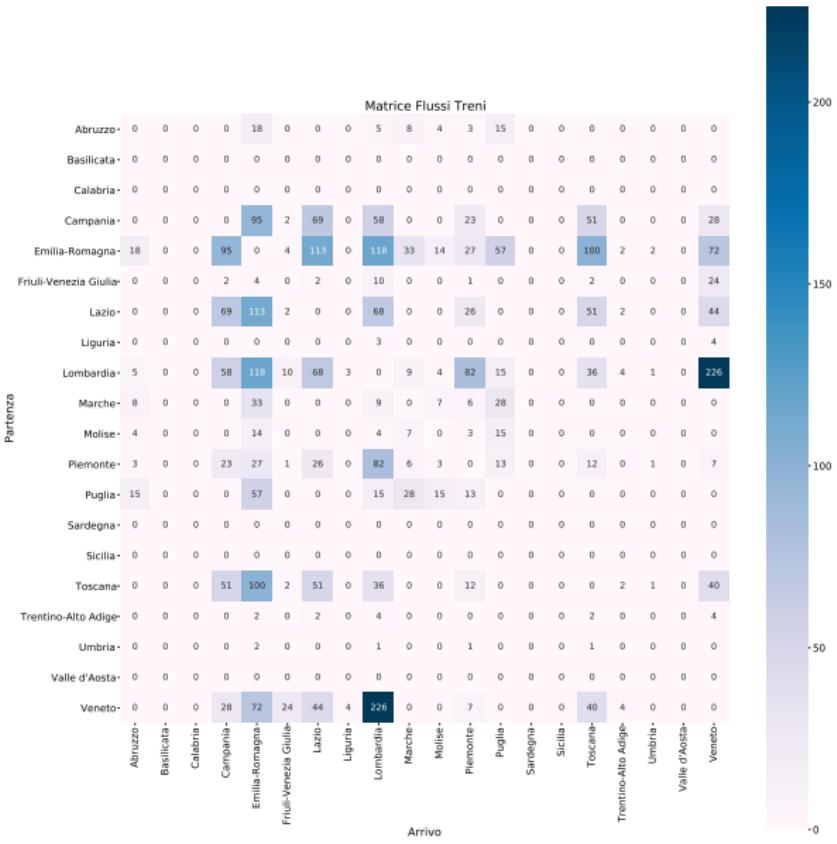


Figure: Matrice ferroviaria con colorazione relativa al flusso. (link)

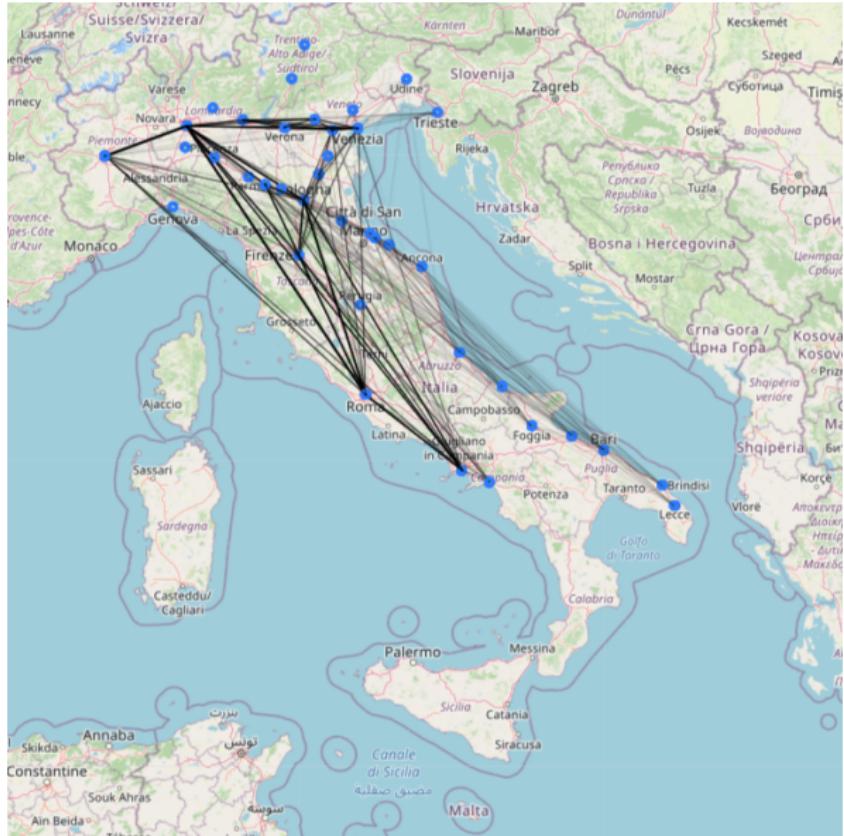


Figure: Rete ferroviaria con spessore dei link proporzionale alle tratte. (link)

Peso dei link: rete aeroportuale

Per dare un senso probabilistico al peso del link (per poi utilizzarlo nel modello di diffusione dell'epidemia) sono stati normalizzati tutti i pesi dei link di ogni singolo nodo in base al flusso totale in uscita dal nodo stesso.

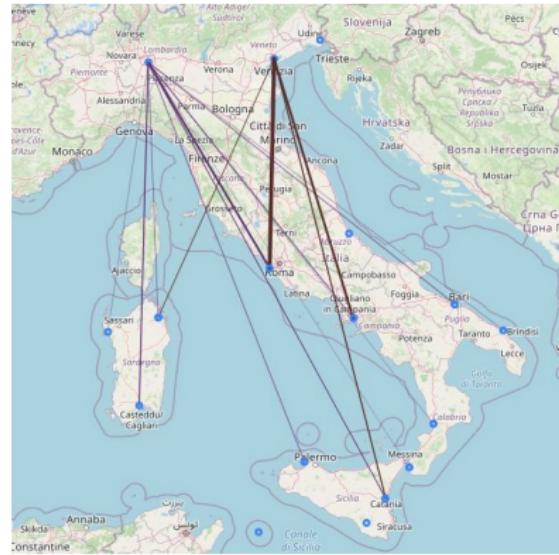
$$P_{\text{uscita}:A \rightarrow B} = \frac{F_{A \rightarrow B}}{\sum_{C \in V_A} F_{A \rightarrow C}}$$

Un'alternativa a questo consiste nel normalizzare rispetto al flusso totale in entrata su ogni nodo. La formula sarebbe

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

dove V_A sono tutti i primi vicini del nodo A e $F_{A \rightarrow B}$ è il flusso di passeggeri tra i nodi A e B .

Peso dei link: rete aeroportuale



(a) Link pesati con $P_{uscita:A \rightarrow B}$ (link)



(b) Link pesati con $P_{entrata:B \rightarrow A}$ (link)

Peso dei link: rete ferroviaria

Come anticipato in precedenza, non avendo a disposizione il flusso di passeggeri per la rete ferroviaria, è stato usato il numero di treni per ogni tratta. Ipotizzando che il flusso di passeggeri sia linearmente proporzionale al numero di treni tra due stazioni, una volta normalizzati, i pesi delle due reti hanno lo stesso significato.

$$P_{\text{uscita}:A \rightarrow B} = \frac{T_{A \rightarrow B}}{\sum_{C \in V_A} T_{A \rightarrow C}}$$

dove V_A sono tutti i primi vicini del nodo A e $T_{A \rightarrow B}$ è il numero di tratte tra i nodi A e B .

Ovviamente l'ipotesi fatta è solo un'approssimazione e ci espone ad un errore sistematico.

Peso dei link: rete ferroviaria

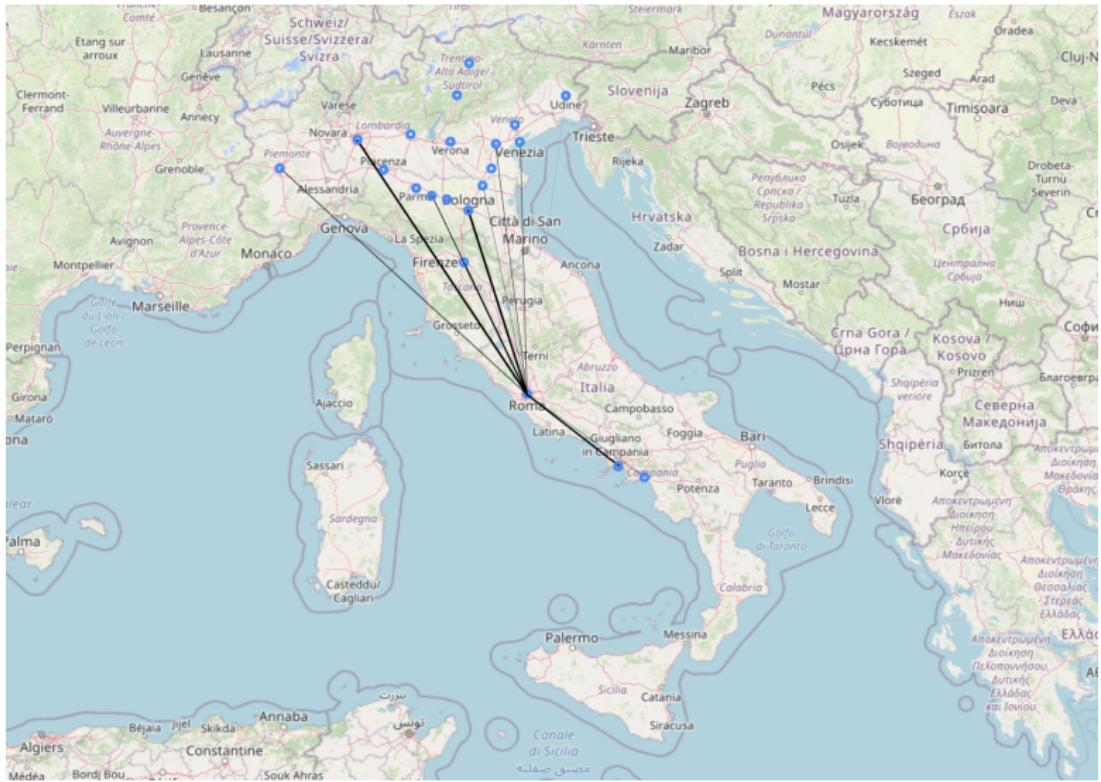


Figure: Link pesati con $P_{uscita:A \rightarrow B}$ (link)

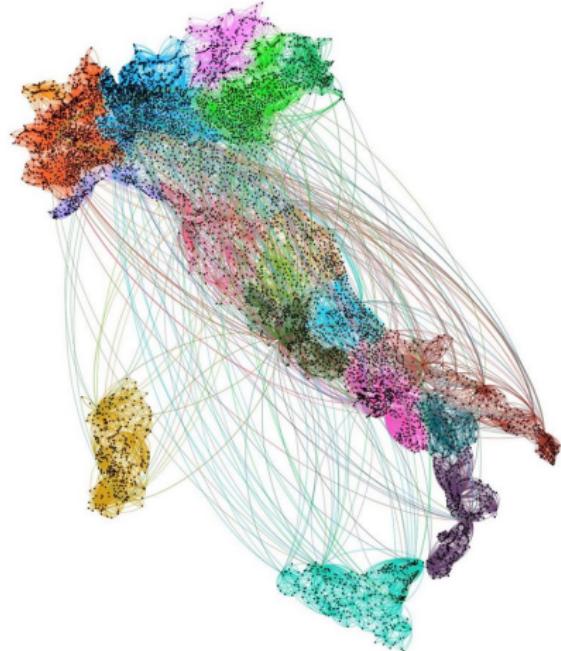
Composizione dei database finali

Le reti finali sono state messe a disposizione degli altri gruppi tramite quattro file csv: due per la rete aeroportuale, due per quella ferroviaria. I due file sono strutturati nel seguente modo:

- Provinciale (aerei) (treni):
 - codice istat della provincia di partenza
 - codice istat della provincia di arrivo
 - peso del link normalizzato con $P_{uscita:A \rightarrow B}$
- Regionale (aerei) (treni):
 - codice istat del capoluogo di regione di partenza
 - codice istat del capoluogo di regione di arrivo
 - peso del link normalizzato $P_{uscita:A \rightarrow B}$

Rete Globale

La rete globale è stata dal gruppo 3 costruita dall'unione di diversi layer:



- Connessione tra Comuni contigui;
- Connessione dei Comuni con il Capoluogo di Provincia;
- Connessione tra Province contigue;
- Rete Ferroviaria (AV);
- Rete Aerea;
- Rete generata da flussi economici (4 Classi Ateco).

Conclusioni e prospettive future

- Le Reti Aeree e ferroviarie (AV) sono state costruite a partire da dati con una diversa struttura;
- Bisogna tenere in considerazione che le reti ottenute sono parziali e che possono essere affette da una sistematica non banale;
- Si potrebbe pensare di fare meglio nel caso in cui si avessero a disposizione i dati relativi al flusso di passeggeri per ogni tipologia di mezzo, dando maggior peso al mezzo più utilizzato in relazione alla tratta.