INDIA BUS ROUTES (2016)

Leonardo Massaro
N.mat:27000/410
Roma,Italia
l.massaro4@lumsastud.it

Riccardo Fidanza
N.mat:26814/410
Roma,Italia
r.fidanza1@lumsastud.it

Davide Di Marco
N.mat:27017/410
Roma, Italia
d.dimarco@lumsastud.it

I. Introduzione

La Social Network analysis (SNA) è composta da strumenti utilizzati per analizzare i modelli di relazioni tra le persone nei gruppi e quindi di individuare e studiare i rapporti nel sociale. La SNA implica la raccolta di dati da più fonti (come sondaggi, e-mail, blog), l'analisi dei dati per identificare le relazioni e l'estrazione di nuove informazioni, come la qualità o l'efficacia di una relazione, inoltre copre varie discipline dalla matematica all'economia ed è una branca della Network Science. Nella Social Network analysis la caratteristica principale sono le persone e le relazione tra esse.
In molti casi lo studio di quest'ultima ha reso possibile l'individuazione di celle terroristiche evitando attentati.
Come nel nostro caso è stata utilizzata per lo studio delle tratte degli autobus di alcune città indiane.

II. India bus routes (2016)

A. Dati

I dati da noi utilizzati provengono da uno studio statistico sulle rete degli autobus di sei grandi città indiane pubblicato sul sito Icon.colorado.edu: Ahmedabad, Chennai, Delhi, Hyderabad, Kolkata, Mumbai.
In queste reti indirette i nodi rappresentano le fermate degli autobus e i link connettono le fermate consecutive, il peso del link indica che il medesimo percorso viene ripetuto più di una volta dai bus.

B. Sviluppo

Lo scopo della nostra Analisi è stato quello di analizzare le 6 differenti Network cercando di capire e indivuare le loro caratteristiche e il loro funzionamento.

III. Ahmedabad

Ahmadâbâd è una suddivisione dell'India, classificata come municipal corporation, di 3 515 361 abitanti, capoluogo del distretto di Ahmedabad, nello stato federato del Gujarat. In base al numero di abitanti la città rientra nella classe I (da 100 000 persone in su) ed è la città più grande dello stato.

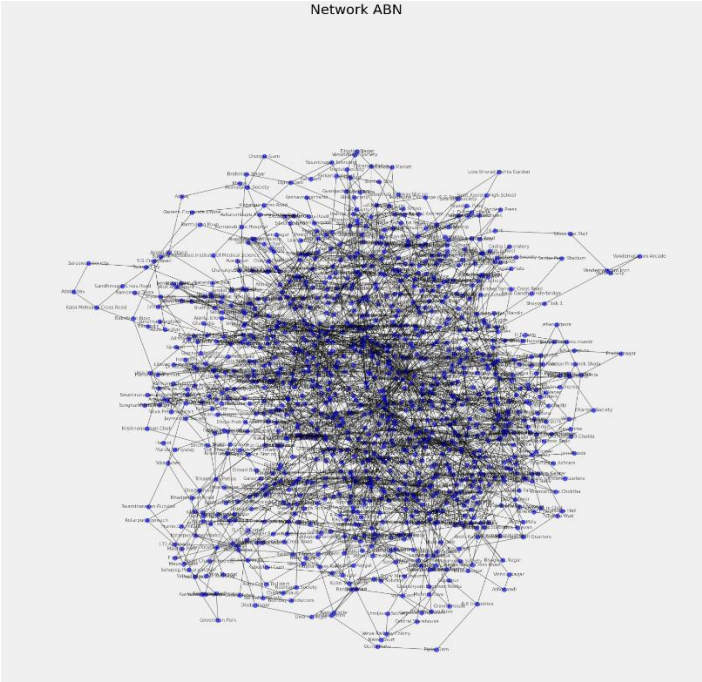
A. Analisi della rete

Il punto iniziale è stato il caricamento della Network e il calcolo dei nodi e link totali con successivo sviluppo di una matrice di adiacenza per una miglior visualizzazione dei dati.

```
NODI:
 1103

LINKS:
 2157
```

Grafico della Network
Una volta individuati i nodi e i link abbiamo graficato a schermo la Network attraverso la libreria networkx di Python



Da come si può vedere la Network è connessa (con avvenuta conferma attraverso networkx).

B. Large Scale

Ci permettono di comprendere la struttura generale della rete. Le metriche sono state calcolate considerando la rete pesata e non pesata.

- Average Path Length: shortest-path medio tra tutti i nodi della Network. Diverso tra rete pesata e non pesata poichè nel calcolo degli shortest path si tiene conto dei pesi.
- Diametro: shortest-path più grande della Network
- Density: rapport tra il numero di link che la rete ha e quelli che potrebbe avere.
- Average Local Clustering: coefficente di clustering medio calcolato per i singoli nodi
- Global Clustering: è il numero di triangoli sul numero di triadi presenti nella Network.

	Network non pesata	Network pesata
Average Path Length	5.5930	6.4630
Diametro	16.0000	16.0000
Density	0.0035	0.0035
Average Local Clustering	0.1891	0.0154
Global Clustering	0.1270	0.1270

All'interno di questa rete il livello di density è abbastanza basso indicandoci che la rete è relativamente poco compatta.

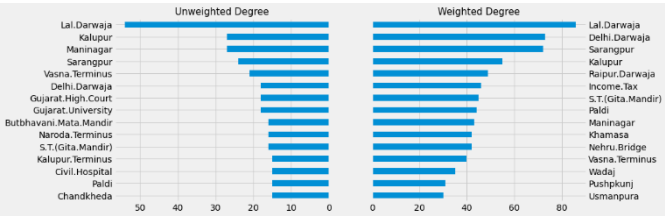
C. Small Scale

Ci consentono di valutare il ruolo dei nodi individuali all'interno della rete.

- Degree: grado di ogni nodo
- Degree centrality: numero di vicini che ha un nodo
- Betweenness: nodi che permettono un collegamento tra nodi separati.
- Closeness: quanto più un nodo arriva con pochi passaggi ad altri nodi della rete.
- Eigenvector: estensione della degree centrality, calcola la centralità di un nodo basandosi sulla centralità dei suoi vicini.
- Page Rank: algoritmo introdotto da Google che utilizza come fattori di calcolo il numero di link che un nodo riceve in ingresso, la propenzione che i nodi collegati a questo hanno altri link, centralità dei nodi che linkano.

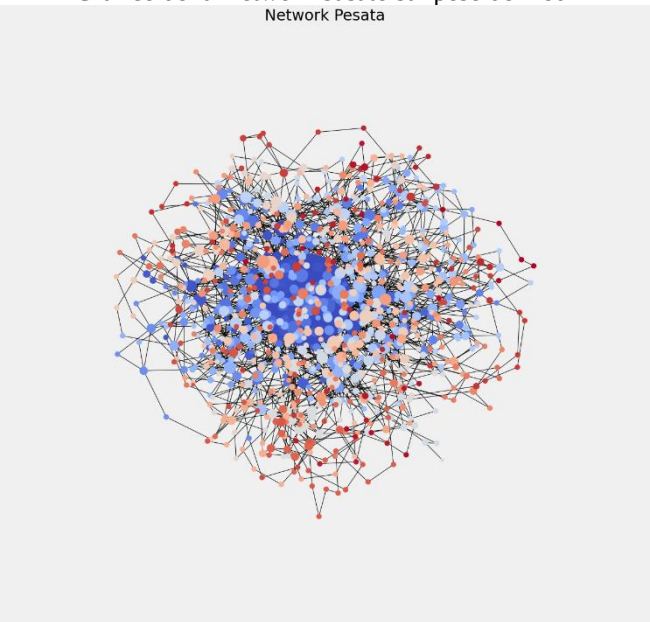
Per ogni metrica è stato eseguito un grafico rappresentante i 15 nodi con valore più alto relativo alla misura presa in considerazione.

DEGREE

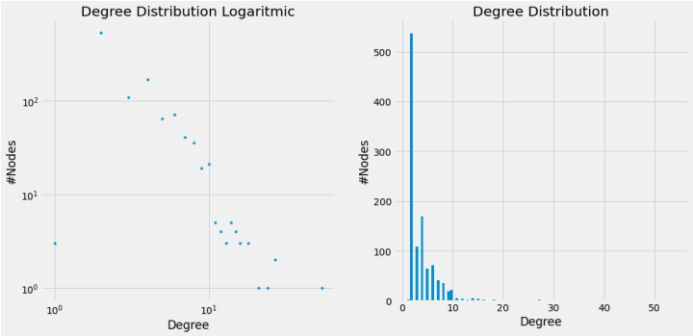


Notiamo come sia nella rete pesata che non Lal.Darwaja sia il nodo con grado maggiore. Questo nodo si può definire hub, la sua rimozione comporterebbe una compromissione della network.

Grafico della Network basato sul peso dei nodi

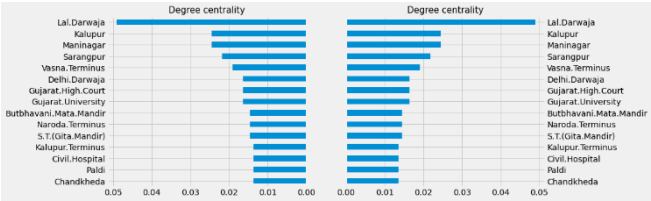


Istogramma della Degree distribution (fornisce la probabilità che un nodo selezionato casualmente nella rete ha grado K)



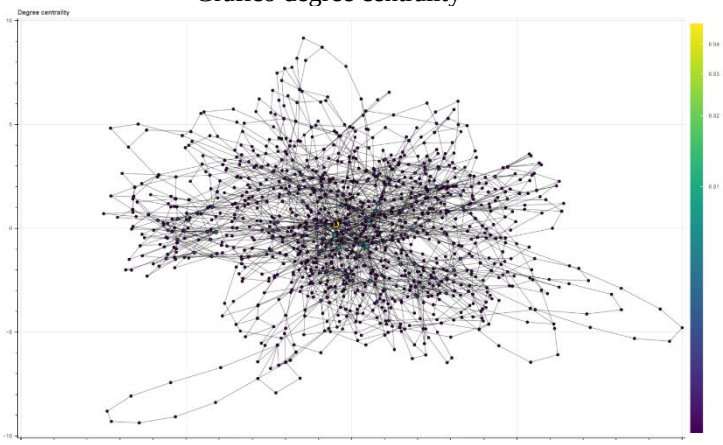
Da questo osserviamo che sono presenti molti nodi con grado basso poichè sono collegati ad altri nodi con grado elevato che rappresentano gli hubs.

DEGREE CENTRALITY

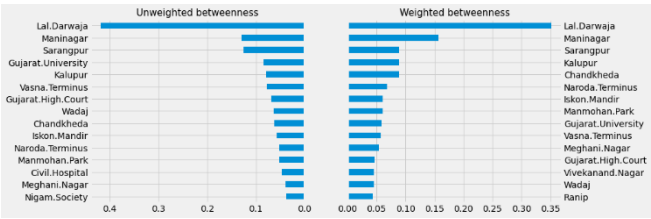


In questo caso Lal.Darwaja è il nodo con maggior Degree centrality e non ci sono differenze tra la rete pesata e non (nel calcolo della degree centrality non si tiene conto dei pesi). Siamo sempre più in grado di confermare l'elevata importanza di questo nodo nella network.

Grafico degree centrality



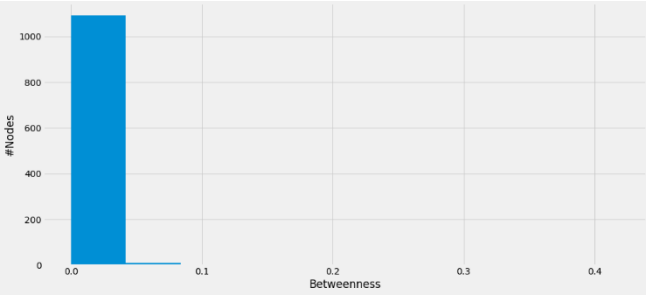
BETWEENNESS



Come per le metriche precedenti anche in questo caso Lal.Darwaja mantiene il valore più elevato anche se assume

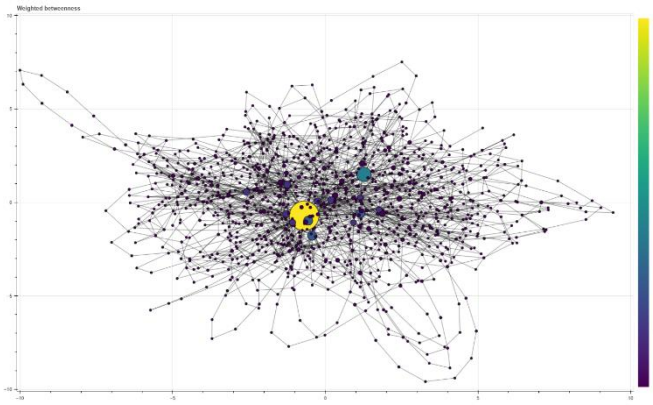
un valore diverso per la rete pesata e non, ciò significa che rappresenta un punto di elevata importanza all'interno della rete di autobus di Ahmedabad, facendoci intuire che possa essere una stazione centrale della città.

Betweenness distribution

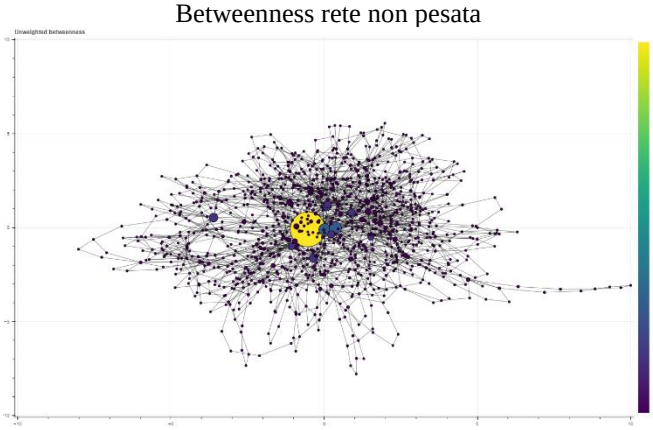


Il grafico conferma che nell'intera rete la maggioranza dei nodi ha un basso valore di betweenness.

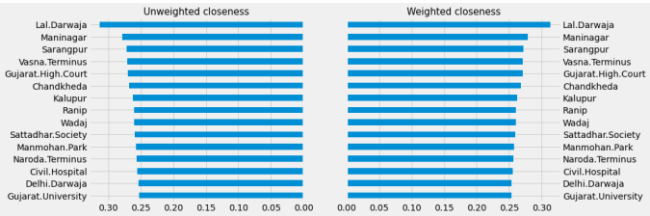
Betweenness rete pesata



Betweenness rete non pesata

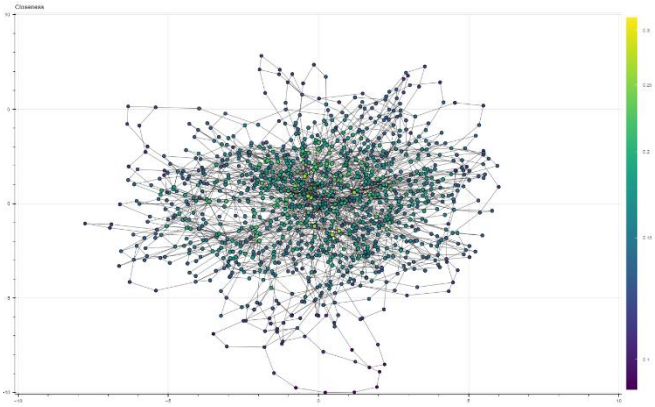


CLOSENESS

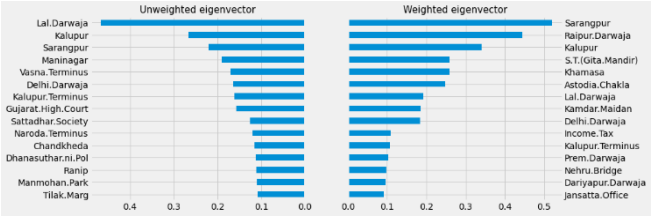


La closeness ci permette di stabilire la centralità di un nodo ovvero quanti nodi si possono raggiungere tramite gli shortest-path del nodo di partenza e non differisce tra rete pesata e non. Differentemente dalle altre metriche il valore di closeness di Lal.Darwaja è comunque il più elevato ma non distaccandosi di molto dagli altri.

Closeness

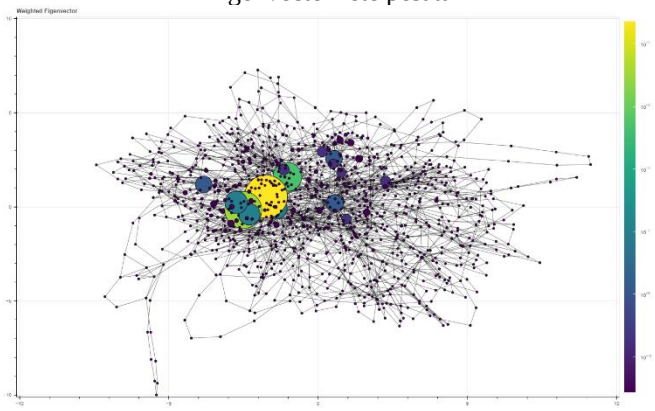


EIGENVECTOR

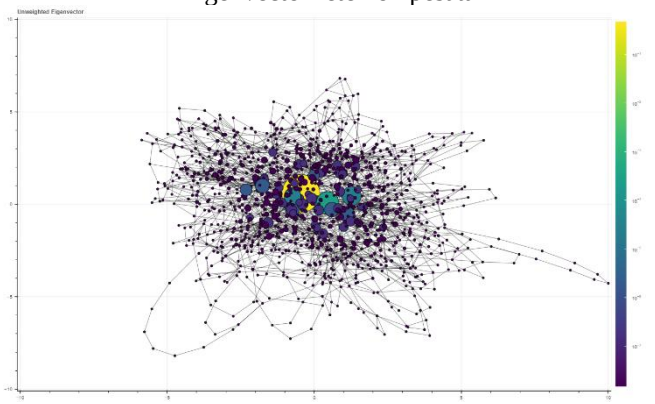


Per la rete non pesata il valore maggiore è occupato sempre da Lal.Darwaja mentre in questa metrica per la rete pesata è Sarangpur, probabilmente perchè è collegato a nodi con alta centralità.

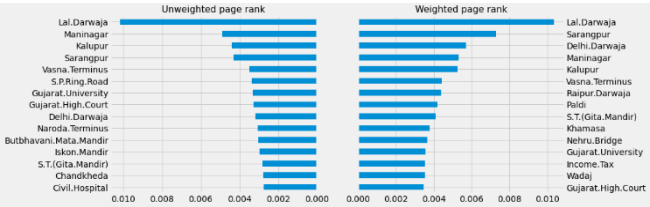
Eigenvector rete pesata



Eigenvector rete non pesata

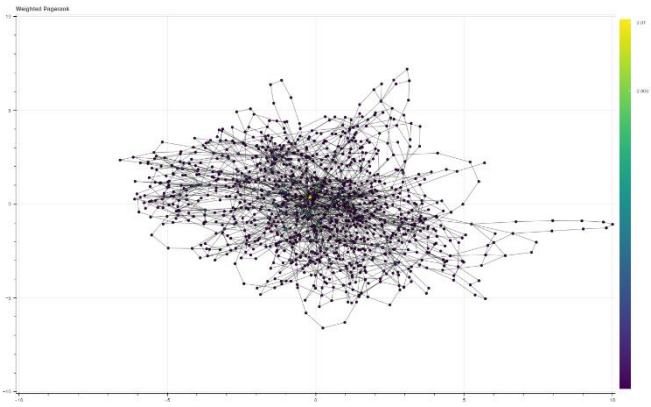


PAGE RANK

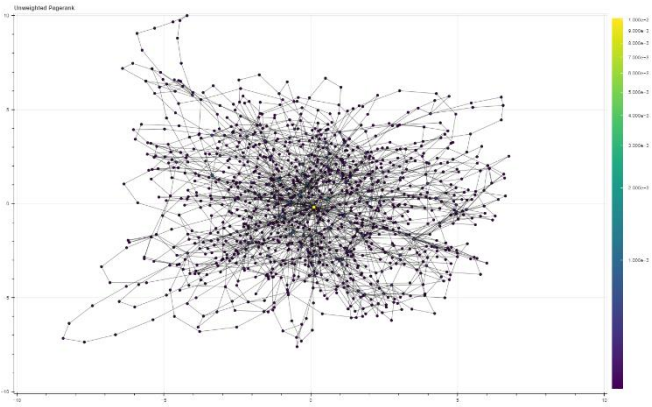


Come nelle metriche precedenti Lal.Darwaja rimane il nodo di maggiore importanza poiché attrae più collegamenti anche da nodi parsimoniosi cioè linkano poco.

Page rank rete pesata



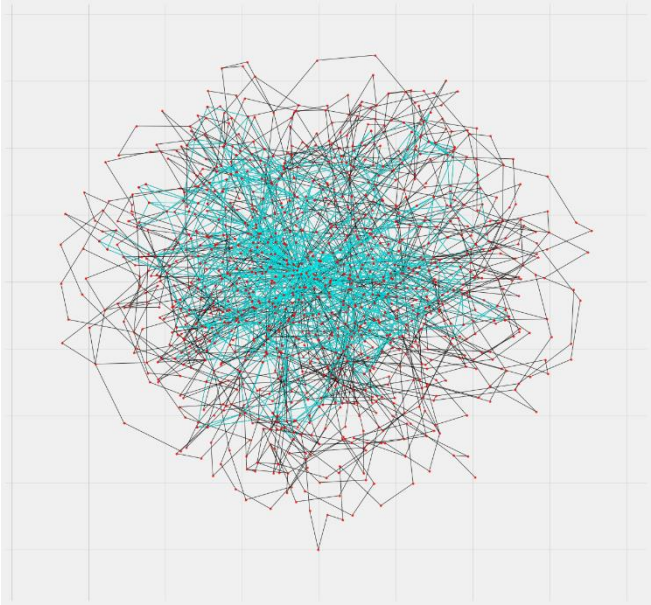
Page rank rete non pesata



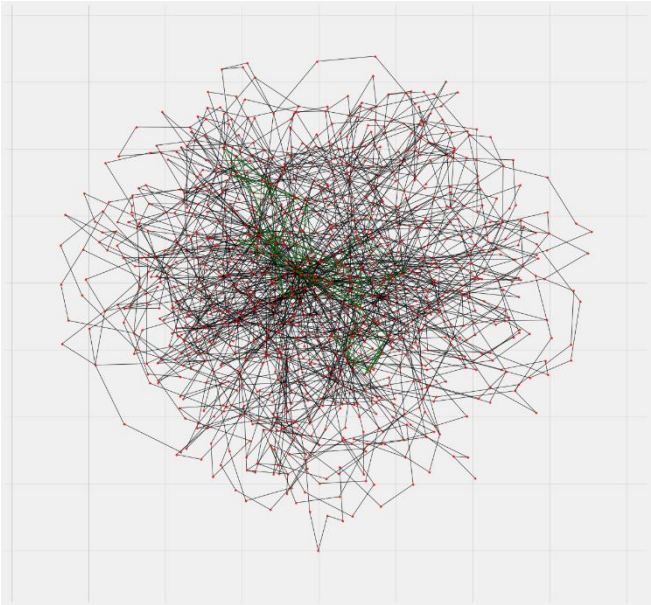
D. Medium Scale

Abbiamo cercato di capire se all'interno della rete fossero presenti delle communities attraverso il metodo K-core e facendolo iterare per tutti i valori possibili compresi tra 0 e 100. Per questa rete i valori più alti sono 3 e 4.

K-core con K=3



K-core con K=4



Come si può ben notare dai grafici soprastanti le differenze sono ben evidenti, il primo grafico ha un raggruppamento più numeroso poiché evidentemente i nodi sono più raggruppati per un core di grado 3.

WEAK TIE

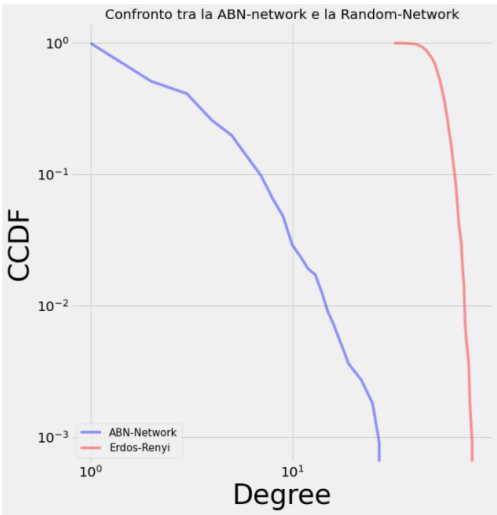
I legami deboli sono molto importanti all'interno di una rete poiché collegano diverse comunità.

Edge	Strength	Span
('Lal.Darwaja', 'Bhadra')	1	4
('Lal.Darwaja', 'Khanpur')	1	4
('Lal.Darwaja', 'Ratan.Park')	1	5
('Lal.Darwaja', 'Chosar')	1	7
('Lal.Darwaja', 'Vivekanand.Nagar')	1	3
('Lal.Darwaja', 'Civil.Hospital')	1	3
('Lal.Darwaja', 'Kalapi.Nagar')	1	3
('Lal.Darwaja', 'Chandkheda.Gam')	1	3
('Lal.Darwaja', 'Gujarat.University')	1	3
('Lal.Darwaja', 'Mirzapur.Road')	1	3
('Lal.Darwaja', 'Makarba.Gam')	1	3
('Lal.Darwaja', 'Anand.Flats')	1	4
('Lal.Darwaja', 'Iskon.Mandir')	1	3
('Lal.Darwaja', 'Judges.Bunglow')	1	3
('Lal.Darwaja', 'Vaishnodevi.Mandir')	1	3
('Lal.Darwaja', 'Jodhpur.Gam')	1	3
('Lal.Darwaja', 'Ghuma.Gam')	1	3
('Lal.Darwaja', 'Vanch')	1	8
('Lal.Darwaja', 'Gatrad.Gam')	1	6
('Lal.Darwaja', 'Kamod.Gam')	1	3
('Lal.Darwaja', 'Sitarambapa.Chowk')	1	3
('Lal.Darwaja', 'Sardar.Park')	1	4
('Lal.Darwaja', 'Swami.Vivekanand.Bridge')	1	3

E. Networks models

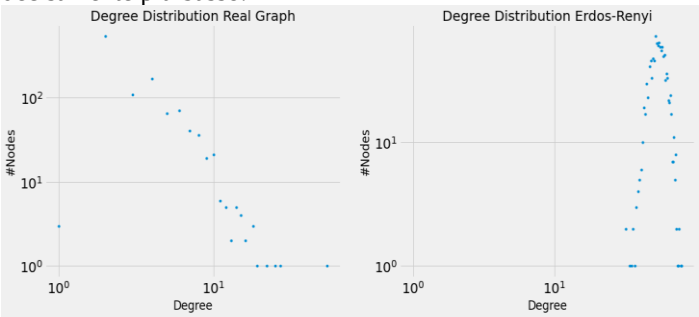
In ultima analisi abbiamo sviluppato 3 diversi modelli per simulare la nostra rete partendo da regole ben definite. Per ognuno di questi si è proceduto ad un confronto grafico della CCDF (distribuzione cumulativa, probabilità che un nodo abbia un certo grado o che sia maggiore di quell grado), ad un confronto su misure a larga scala e a un confronto della degree distribution.

RANDOM NETWORK ERDOS-RENYI



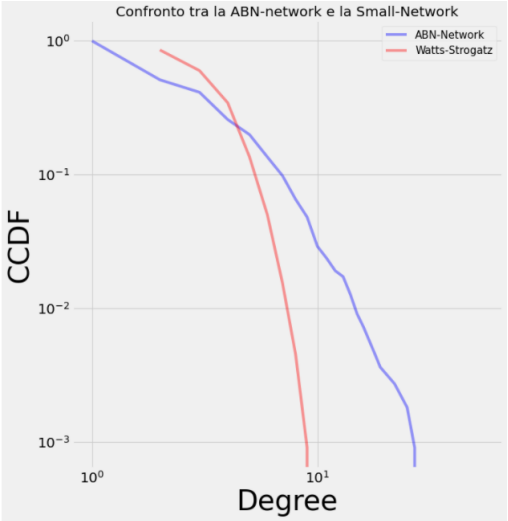

```
Average degree Random Network: 54.97733454215775
Average degree ABN Network: 3.898458748866727
Il global clustering della network ABN è:  0.127
Il global clustering della random network è:  0.05
L'APL della network ABN è:  6.463
L'APL della random network è:  2.011
```

Come si può notare dal grafico la CCDF si discosta molto dal similuzione della nostra rete. Questo modello simula una rete con average degree molto più alto della rete reale e APL decisamente più basso.



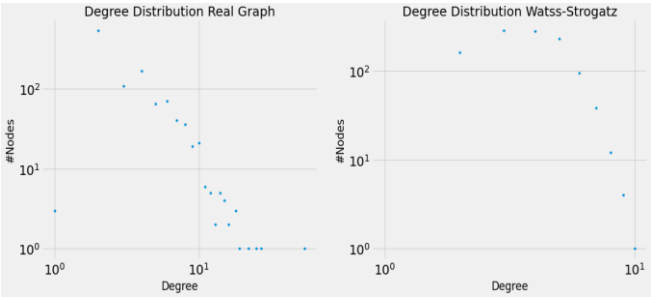
Anche la degree distribution è totalmente diversa. Possiamo concludere che il modello di Erdos-Renyi non è adatto a simulare la nostra rete.

WATTS-STROGATZ



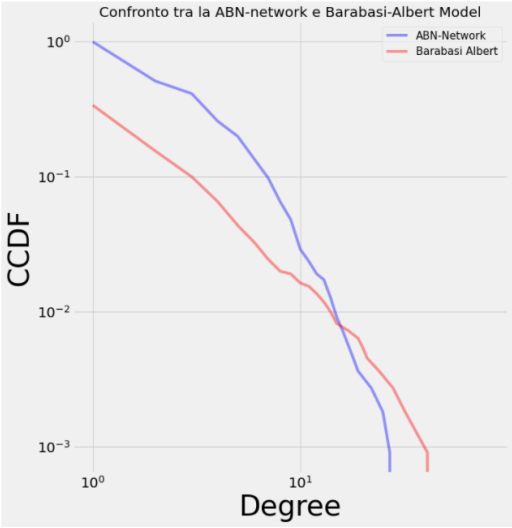

```
Average degree Small word: 4.0
Average degree ABN Network: 3.898458748866727
Il global clustering della network ABN è:  0.127
Il global clustering della Small word network è:  0.002
L'APL della network ABN è:  6.463
L'APL della small word network è:  5.403
```

In questo caso sia la CCDF che le misure a larga scala si avvicinano molto di più alla nostra rete.



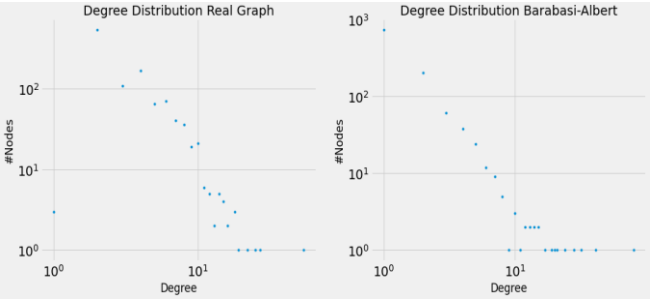
La degree distribution non si avvicina alla rappresentazione della nostra rete e anche in questo caso possiamo concludere affermando che questo modello non è adatto.

BARABASI ALBERT



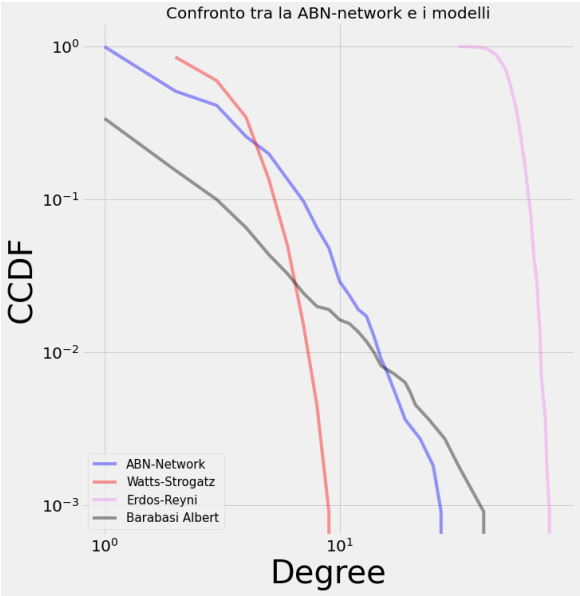

```
Average degree Barabasi-Albert: 1.99818676337262
Average degree ABN Network: 3.898458748866727
Il global clustering della network ABN è:  0.127
Il global clustering del modello di Barabasi-Albert è:  0
L'APL della network ABN è:  6.463
L'APL del modello di Barabasi- Albert è:  7.049
```

A differenza dei modelli precedenti la CCDF si avvicina molto di più mentre per alcune large scale come, l'average degree, il valore è più basso.



Per quanto riguarda la degree distribution segue abbastanza l'andamento della nostra rete.

Infine riportiamo un grafico di confornto fra tutte le CCDF.



Possiamo concludere che il modello di Barabbasi Albert è il migliore per simulare la nostra rete.

F. Conclusioni

Lo studio della network è stato fondamentale per individuare l'importanza del nodo Lal Darwaja all'interno della rete dei bus di Ahmedabad, il quale attraverso alcune ricerche sul web si è scoperto essere un deposito al centro della città a conferma delle nostre intuizioni. Questo ci porta a dire che molti nodi all'interno della network sono collegati ad altri nodi passando per degli hub, la rimozione di quest'ultimi potrebbe compromettere la rete stessa.

IV. CHENNAI

Chennai è una suddivisione dell'India, classificata come municipal corporation, di 4 216 268 abitanti, capoluogo del distretto di Chennai, nello stato federato del Tamil Nadu di cui è la capitale. Sino al 1996 la città era conosciuta col nome di Madras, che all'interno della conurbazione odierna è il nome del nucleo centrale di Chennai. La città, importante centro commerciale ed industriale, vanta un ricco patrimonio culturale ed artistico.

A. Analisi della rete

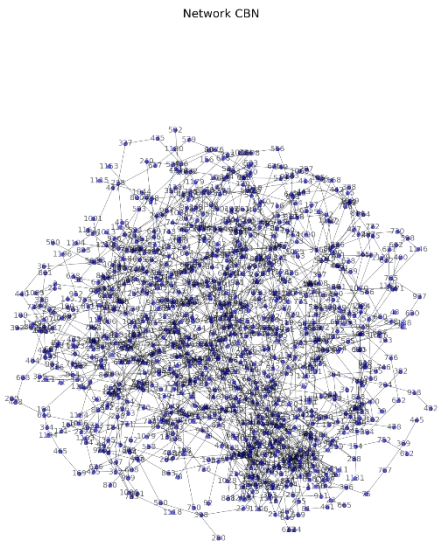
Il punto iniziale è stato il caricamento della Network e il calcolo dei nodi e link totali con successivo sviluppo di una matrice di adiacenza per una miglior visualizzazione dei dati.

```
NODI:
 1009

LINKS:
 1610
```

Grafico della Network

Una volta individuati i nodi e i link abbiamo graficato a schermo la Network attraverso la libreria networkx di Python



B. *Large Scale*

Ci permettono di comprendere la struttura generale della rete. Le metriche sono state calcolate considerando la rete pesata e non pesata.

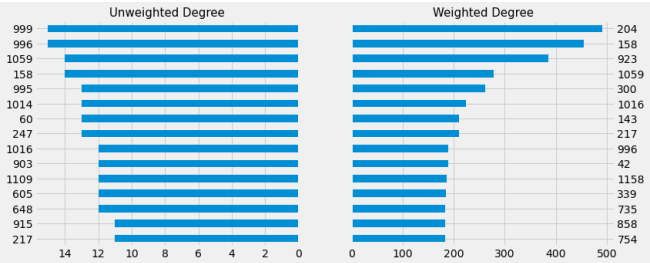
	Network non pesata	Network pesata
Average Path Length	8.7300	23.6050
Diametro	27.0000	27.0000
Density	0.0032	0.0032
Average Local Clustering	0.0678	0.0021
Global Clustering	0.1050	0.1050

Anche all'interno di questa rete il livello di density è abbastanza basso indicandoci che la rete è anch qui poco compatta, il diametro e l'APL sono decisamente più grandi rispetto alla rete precedente indicandoci che i nodi sono più distanti tra di loro.

C. *Small Scale*

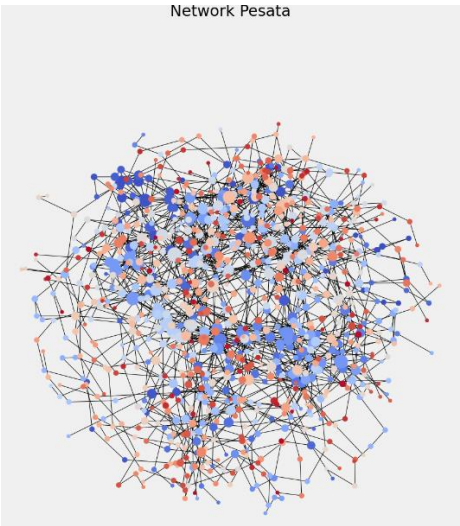
Ci consentono di valutare il ruolo dei nodi individuali all'interno della rete.
Per ogni metrica è stato eseguito un grafico rappresentante i 15 nodi con valore più alto relativo alla misura presa in considerazione.

DEGREE

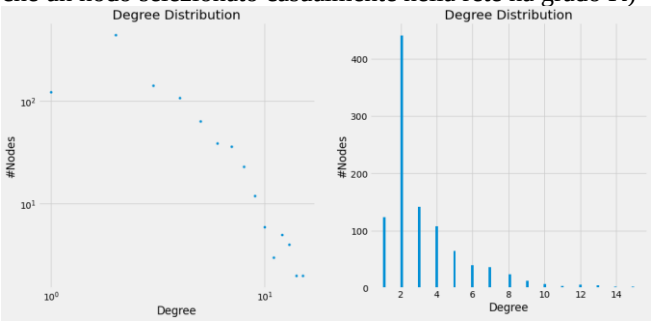


Nella rete non pesata i nodi con grado più alto sono il 999 e il 996 ma con un valore non troppo distaccato dagli altri.
Per la rete pesate i nodi 204, 158 e 923 si distanziano dagli altri avendo un peso molto alto

Grafico della Network basato sul peso dei nodi



Istogramma della Degree distribution (fornisce la probabilità che un nodo selezionato casualmente nella rete ha grado K)

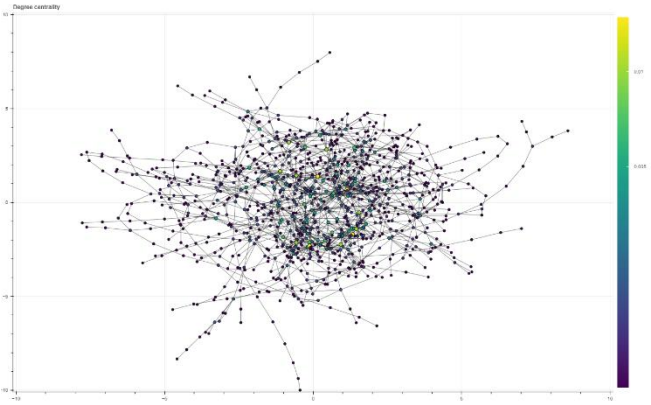


Da questo osserviamo che sono presenti molti nodi con grado basso soprattutto quelli con grado 2 che superano i 400

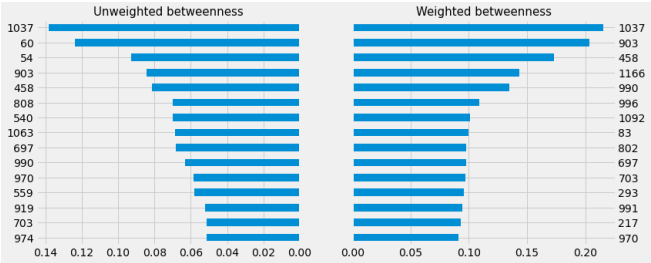
DEGREE CENTRALITY

Per la degree centrality i nodi 999 e 996 sono quelli con valore più elavato senza differenze tra rete pesata e non.

Da qui si inizia a vedere che questi due nodi possano risultare i più importanti nella rete

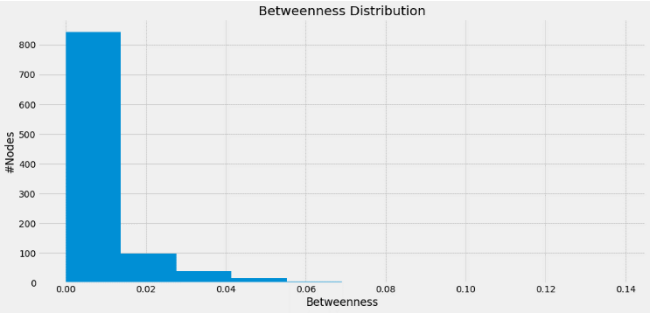


BETWEENNESS



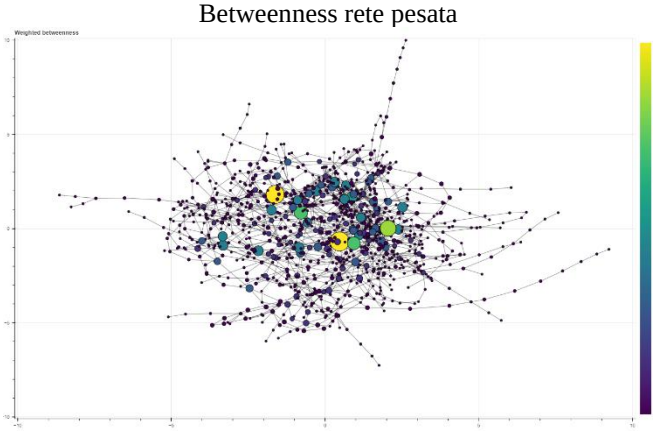
A differenza delle metriche precedenti il nodo con valore più elevato è il 1037 in tutte e due le reti, questo ci permette di affermare l'importanza di questo nodo che permette un

collegamento tra nodi separati trovandosi spesso al centro degli shortest-path tra due nodi.

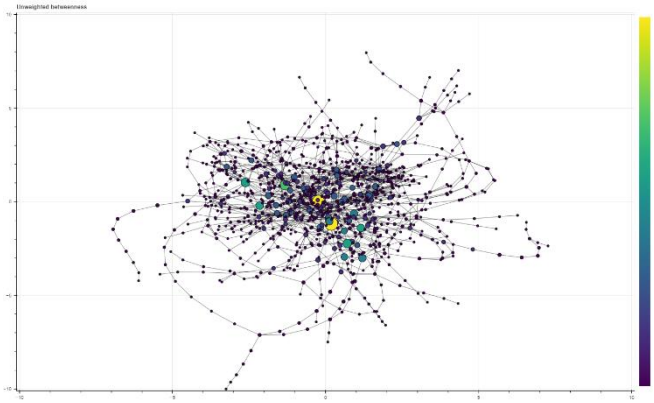


Il grafico conferma che nell'intera rete la maggioranza dei nodi ha un basso valore di betweenness, a differenza della città precedente abbiamo un più alto numero di nodi con betweenness maggiore.

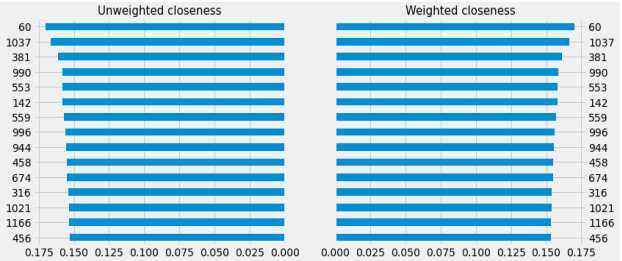
Betweenness rete pesata



Betweenness rete non pesata

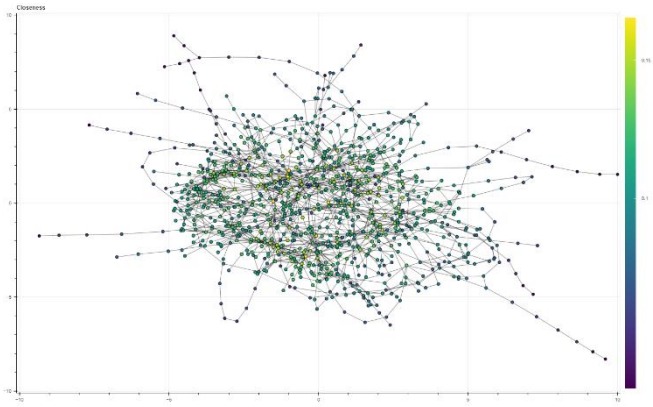


CLOSENESS

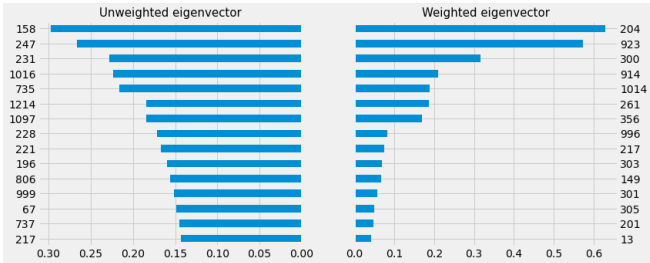


La closeness ci permette di stabilire la centralità di un nodo ovvero quanti nodi si possono raggiungere tramite gli shortest-path del nodo di partenza e non differisce tra rete pesata e non. Il valore di closeness dei 15 nodi con punteggio più alto è abbastanza omogeneo, in questo caso i primi sono il 60 e il 1037.

Closeness

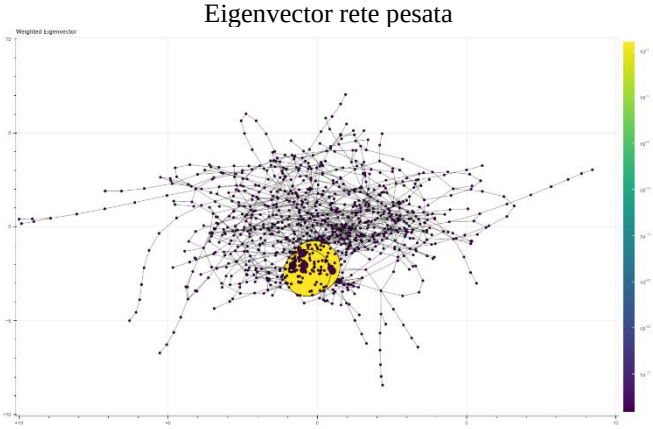


EIGENVECTOR

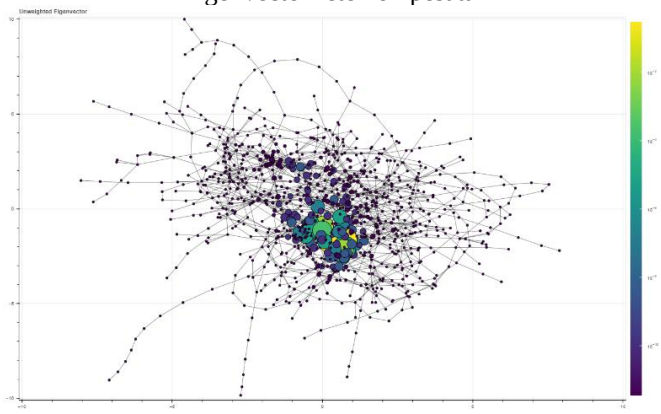


Per l'eigenvector i valori cambiano totalmente tra rete pesata e non. Nella rete pesata 204 e 923 si distaccano molto dagli altri come si può vedere nel grafico, causato probabilmente per la loro connessione con nodi con alta centralità

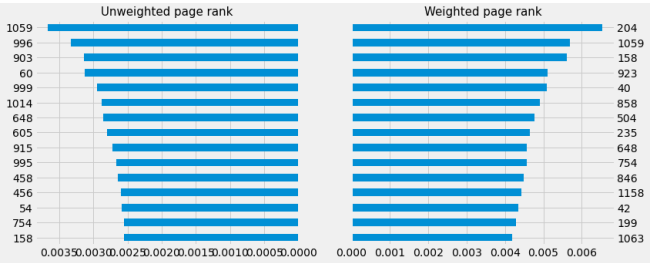
Eigenvector rete pesata



Eigenvector rete non pesata

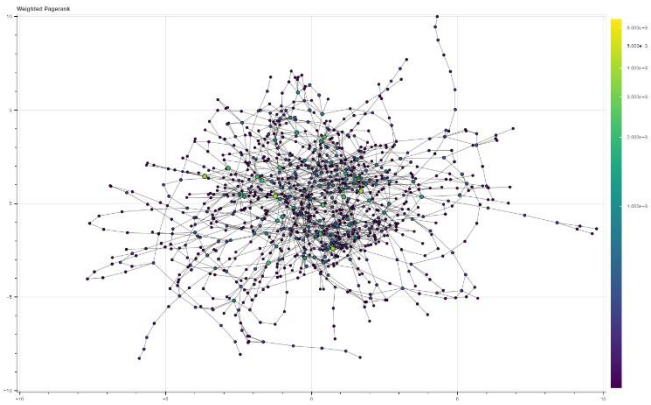


PAGE RANK

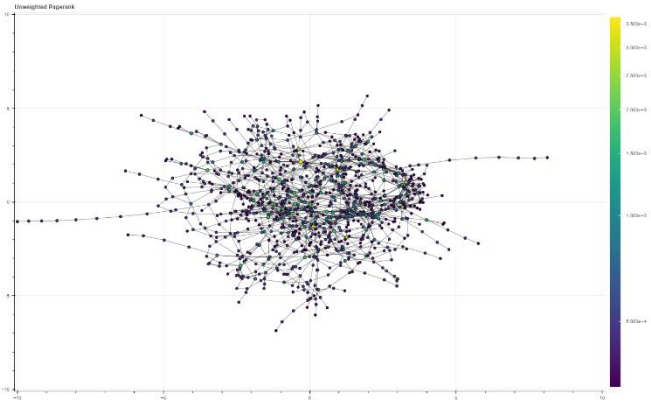


Come in altre metriche si riconfermano gli stessi nodi con valore più alto cioè 204, 996, 60

Page rank rete pesata



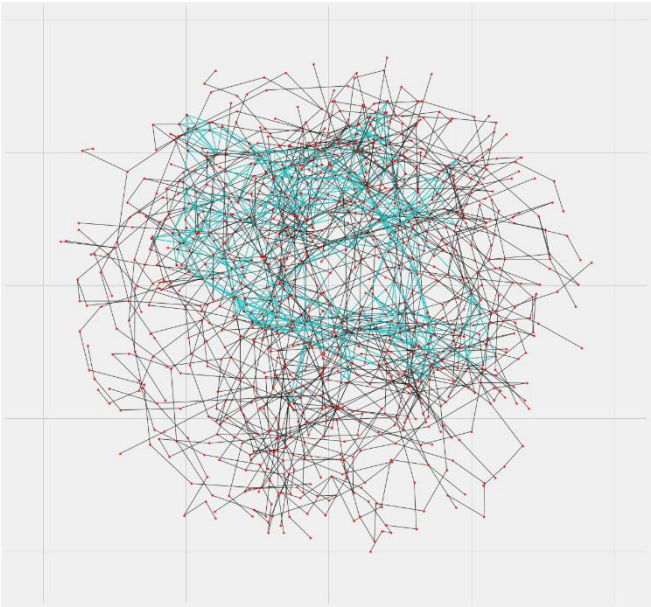
Page rank rete non pesata



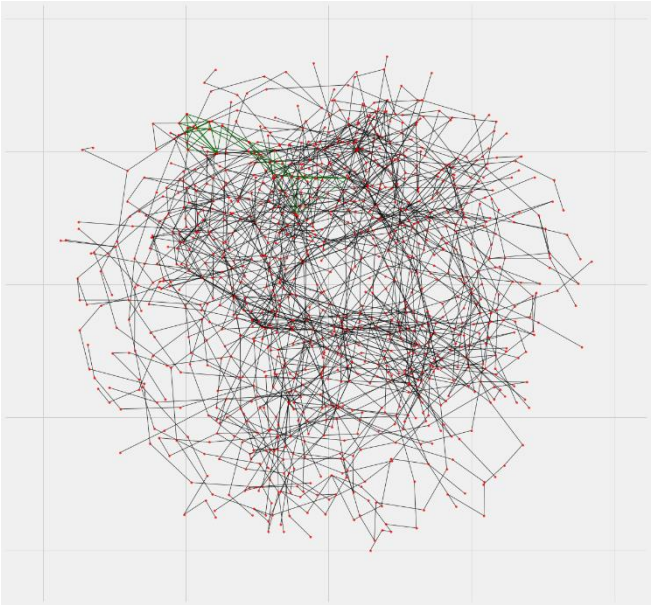
D. Medium Scale

Abbiamo cercato di capire se all'interno della rete fossero presenti delle communities attraverso il metodo K-core e facendolo iterare per tutti i valori possibili compresi tra 0 e 100. Per questa rete i valori più alti sono 3 e 4.

K-core con K=3



K-core con K=4



Anche qui la distinzione è ben evidente anche se in entrambi i casi i gruppi non sono molto numerosi.

WEAK TIE

I legami deboli sono molto importanti all'interno di una rete poiché collegano diverse comunità.

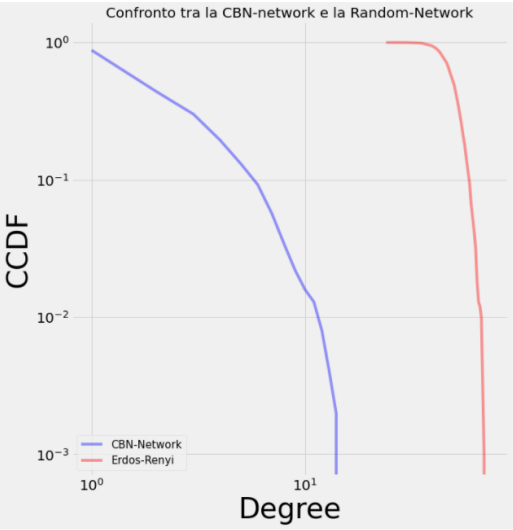
In questa network si possono trovare alcuni con span infinito, la rimozione di questi potrebbe comportare un distaccamento della rete.

Edge	Strength	Span
(1, 479)	1	9
(1, 601)	1	9
(479, 1203)	1	9
(601, 567)	1	9
(2, 385)	1	6
(2, 1085)	1	6
(385, 547)	1	4
(385, 1006)	1	4
(1085, 214)	1	3
(1085, 1088)	1	3
(3, 469)	1	10
(116, 235)	1	10
(469, 455)	1	10
(4, 313)	1	inf
(313, 12)	1	inf
(5, 108)	1	3
(5, 581)	1	4
(14, 132)	1	3
(108, 131)	1	3
(108, 132)	1	3
(278, 895)	1	3
(278, 896)	1	3
(278, 936)	1	4

E. Networks models

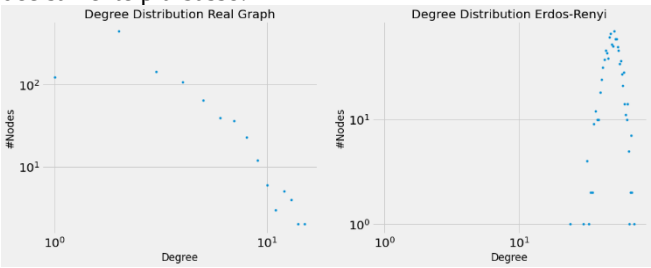
In ultima analisi abbiamo sviluppato 3 diversi modelli per simulare la nostra rete partendo da regole ben definite. Per ognuno di questi si è proceduto ad un confronto grafico della CCDF (distribuzione cumulativa, probabilità che un nodo abbia un certo grado o che sia maggiore di quell grado), ad un confronto su misure a larga scala e a un confronto della degree distribution.

RANDOM NETWORK ERDOS-RENYI



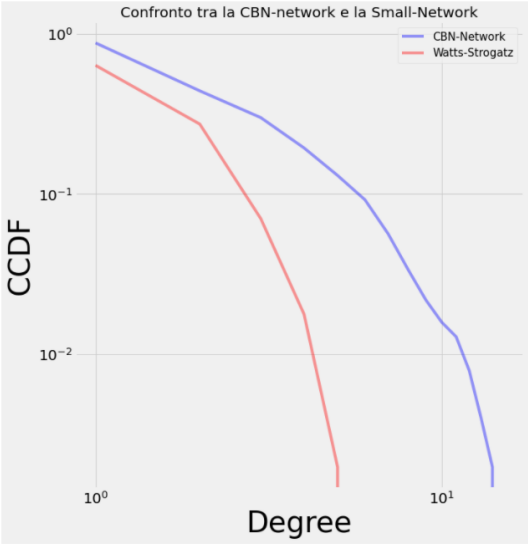

```
Average degree Random Network: 50.22596630327057
Average degree CBN Network: 3.191278493557978
Il global clustering della network CBN è:  0.105
Il global clustering della random network è:  0.049
L'APL della network CBN è:  23.605
L'APL della random network è:  2.028
```

Come si può notare dal grafico la CCDF del modello si discosta molto della nostra rete. Questo modello simula una rete con average degree molto più alto della rete reale e APL decisamente più basso.



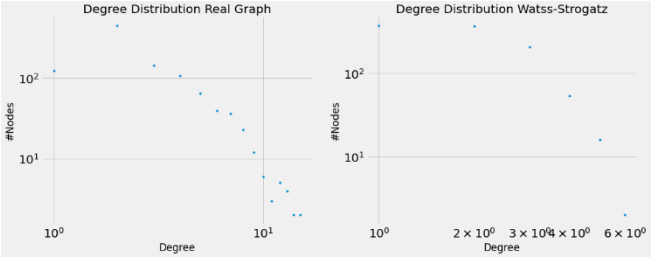
Anche la degree distribution è totalmente diversa. Possiamo concludere che il modello di Erdos-Renyi non è adatto a simulare la nostra rete.

WATTS-STROGATZ



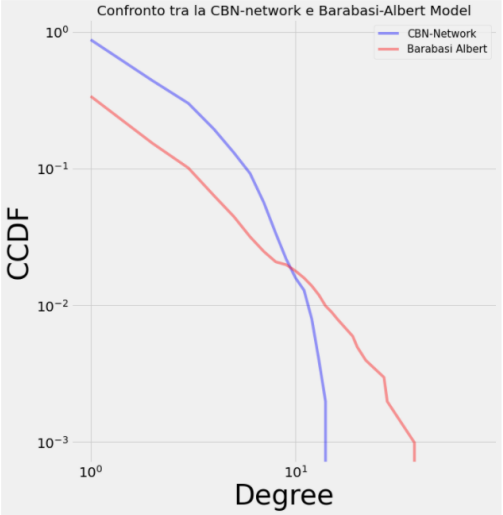

```
Average degree Small word: 2.0
Average degree CBN Network: 3.191278493557978
Il global clustering della network CBN è:  0.105
Il global clustering della Small word network è:  0
```

In questo caso sia la CCDF che le misure a larga scala si avvicinano molto di più alla nostra rete.



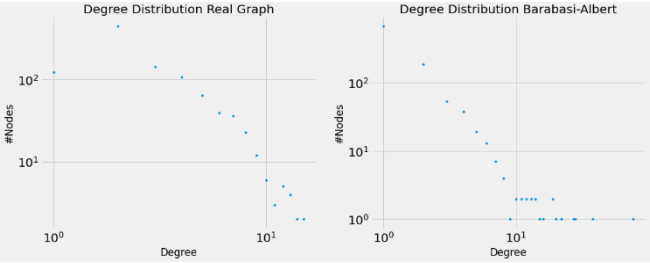
La degree distribution non si avvicini alla rappresentazione della nostra rete e anche in questo caso possiamo concludere affermando che questo modello non è adatto.

BARABASI ALBERT



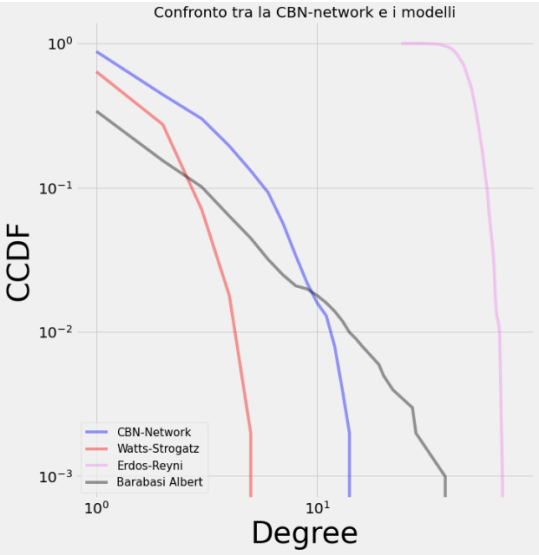

```
Average degree Barabasi-Albert: 1.998017839444995
Average degree CBN Network: 3.191278493557978
Il global clustering della network CBN è:  0.105
Il global clustering del modello di Barabasi-Albert è:  0
L'APL della network CBN è:  23.605
L'APL del modello di Barabasi- Albert è:  6.987
```

A differenza dei modelli precedenti la CCDF si avvicina molto di più mentre per alcune large scale come, l'average degree e APL il valore è più basso.



Per quanto riguarda la degree distribution segue abbastanza l'andamento della nostra rete.

Infine riportiamo un grafico di confornto fra tutte le CCDF.



In questo caso nessun modello si adatta particolarmente bene alla nostra network. Se si dovesse effettuare una scelta tra i tre ricaderebbe o su Barabasi o su Watts-Strogatz.

F. Conclusioni

Questa rete a differenza della precedente è caratterizzata da un numero minore di link e un maggior diametro e average path length.
Per quanto riguarda i singoli nodi non vi è più un singolo nodo importante ma più di uno, questo probabilmente dovuto dal fatto che la città di Chennai essendo al contempo industriale e artistica potrebbe possedere più di un punto di interesse.

V. DELHI

Delhi è una città situata nel Nord dell'India. Con una popolazione di 28.504.738 abitanti, Delhi è la città più popolosa dell'India, seguita da Mumbai, e la terza città più popolosa al mondo. Amministrativamente è un Territorio dell'Unione, chiamato ufficialmente Territorio Nazionale della Capitale di Delhi (National Capital Region - NCR). Situata sulle rive del fiume Yamuna, nel nord dell'India, è una delle più antiche città continuativamente abitate del mondo.

Delhi fu la capitale di diversi imperi indiani in epoca medievale, e una delle più grandi città lungo le antiche rotte commerciali tra l'India e le regioni nord-occidentali della Pianura Indo-Gangetica. Di quel periodo rimangono innumerevoli monumenti e siti archeologici. L'imperatore Shah Jahan costruì la città (ora nota come "Vecchia Delhi") che servì come capitale dell'Impero Moghul dal 1649 al 1857.

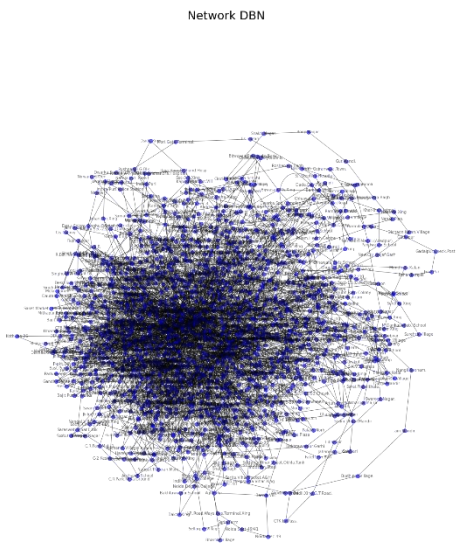
A. Analisi della rete

Il punto iniziale è stato il caricamento della Network e il calcolo dei nodi e link totali con successivo sviluppo di una matrice di adiacenza per una miglior visualizzazione dei dati.

```
NODI:
 1554

LINKS:
 3617
```

Grafico della Network
Una volta individuati i nodi e i link abbiamo graficato a schermo la Network attraverso la libreria networkx di Python



B. *Large Scale*

Ci permettono di comprendere la struttura generale della rete. Le metriche sono state calcolate considerando la rete pesata e non pesata.

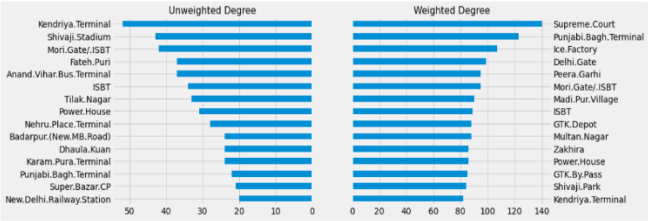
	Network non pesata	Network pesata
Average Path Length	5.5360	6.8810
Diametro	29.0000	29.0000
Density	0.0030	0.0030
Average Local Clustering	0.1833	0.0084
Global Clustering	0.1220	0.1220

Anche all'interno di questa rete il livello di density è abbastanza basso indicandoci che la rete è poco raggruppata, il diametro è il più grande tra le 3 reti analizzate e non vi è una gran differenza tra l'APL della rete pesata e non.

C. *Small Scale*

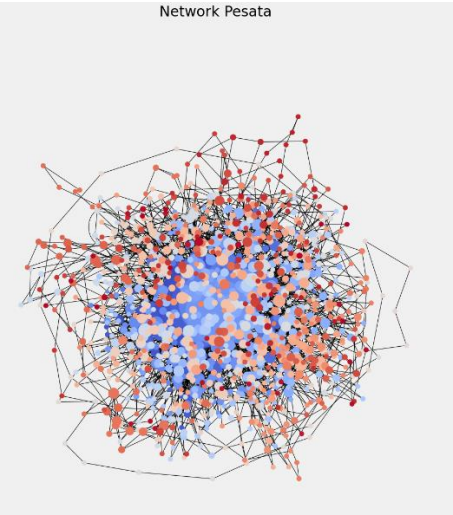
Ci consentono di valutare il ruolo dei nodi individuali all'interno della rete.
Per ogni metrica è stato eseguito un grafico rappresentante i 15 nodi con valore più alto relativo alla misura presa in considerazione.

DEGREE

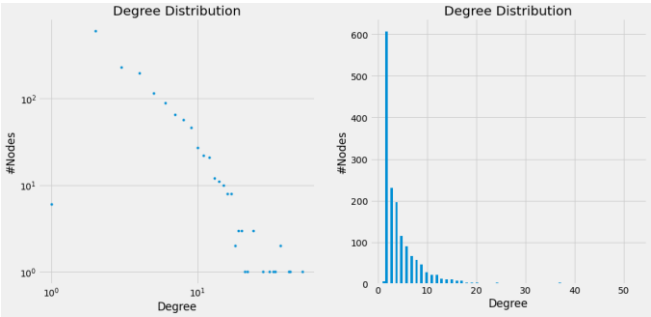


Nella rete non pesata i nodi con grado più alto sono il Kendriya.Terminal e il Shivaji.stadium che già come dicono i nomi presupponiamo siamo due luoghi molto visitati.
Nella rete pesata abbiamo il supreme.court e il Punjabi.bagh.terminal.

Grafico della Network basato sul peso dei nodi

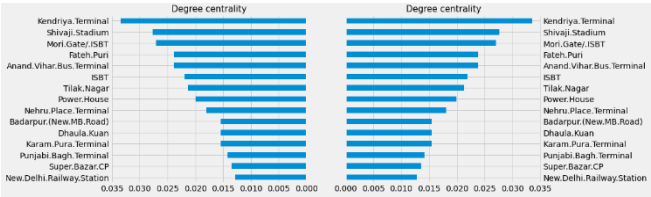


Istogramma della Degree distribution (fornisce la probabilità che un nodo selezionato casualmente nella rete ha grado K)



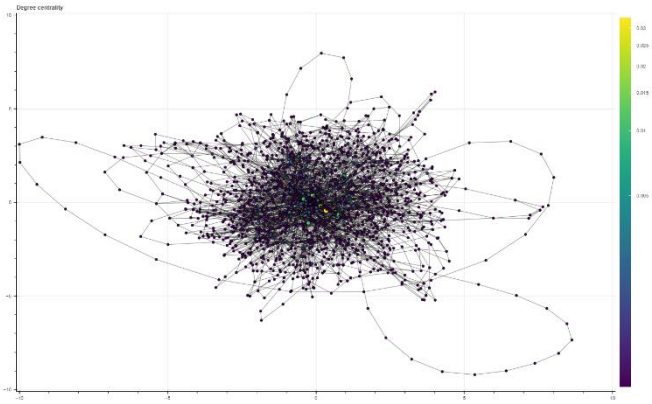
Da questo osserviamo che sono presenti molti nodi con grado basso soprattutto quelli con grado 1 che superano i 600 quindi più della metà della network

DEGREE CENTRALITY

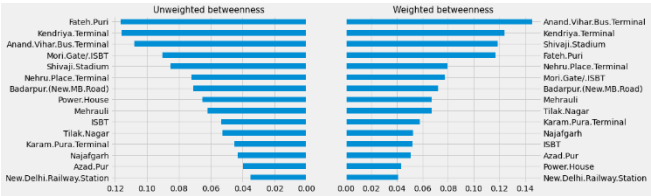


Per la degree centrality i nodi Kendriya.Terminal e Shivaji.stadium sono quelli con valore più elevato senza differenze tra rete pesata e non.

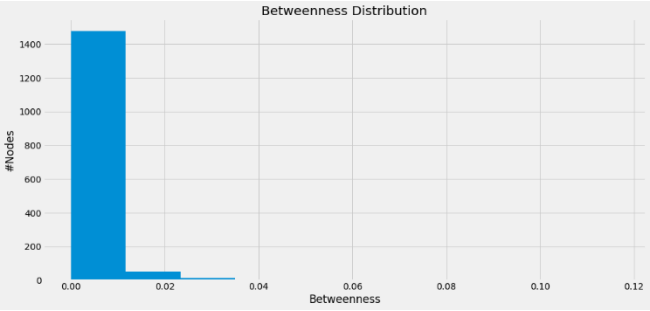
Da qui si inizia a vedere che questi due nodi possano risultare i più importanti nella rete



BETWEENNESS

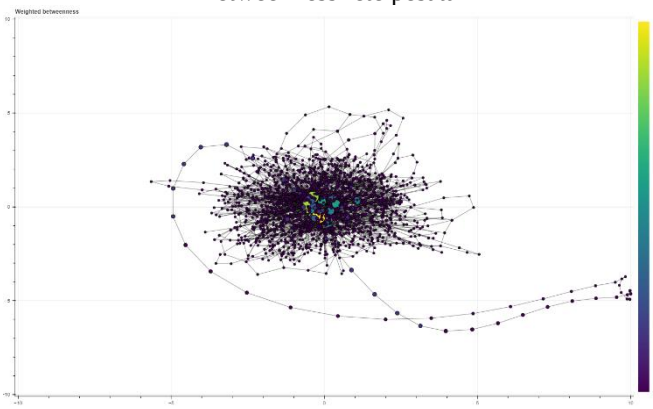


A differenza delle metriche precedenti il nodo con valore più elevato è Fateh.Puri per la rete non pesata e Anand.Vihar.Bus.Terminal per quella pesata, questo ci permette di affermare l'importanza di questo nodo che permette un collegamento tra nodi separati trovandosi spesso al centro degli shortest-path tra due nodi, poiché rappresente un terminale di autobus dove partono molte rotte per molti luoghi .

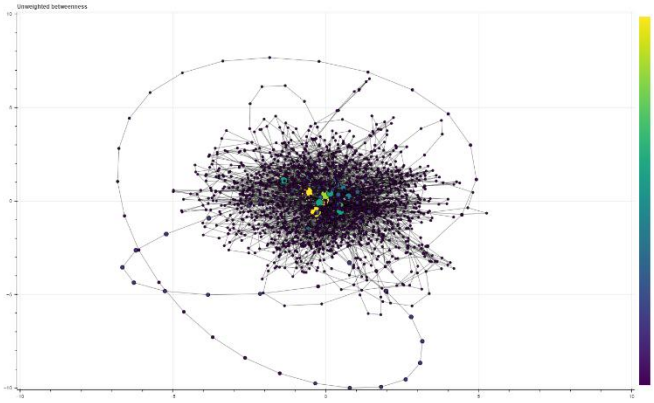


Il grafico conferma che nell'intera rete la maggioranza dei nodi ha un basso valore di betweenness.

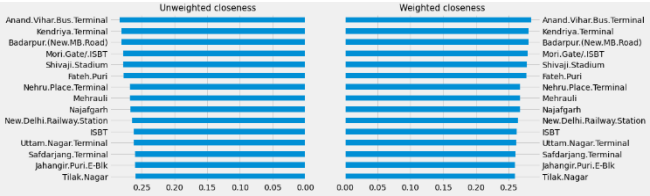
Betweenness rete pesata



Betweenness rete non pesata

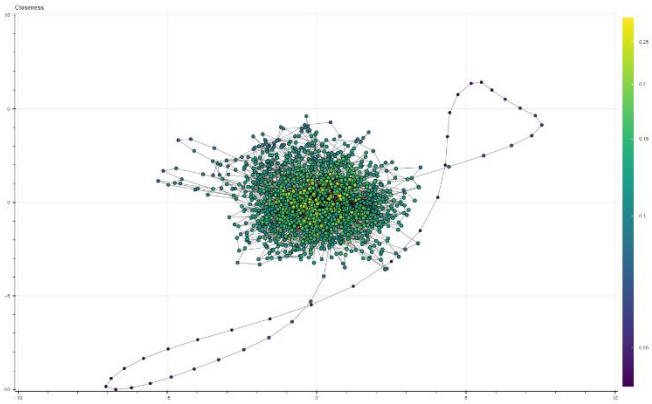


CLOSENESS

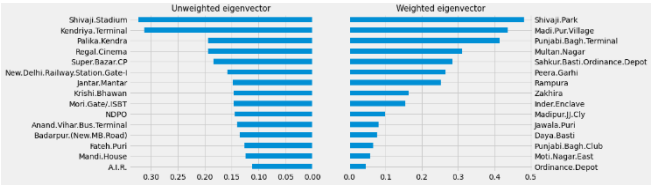


La closeness ci permette di stabilire la centralità di un nodo ovvero quanti nodi si possono raggiungere tramite gli shortest-path del nodo di partenza e non differisce tra rete pesata e non. Il valore di closeness dei 15 nodi con punteggio più alto è abbastanza omogeneo, in questo caso i primi sono il Anand.Vihar.Bus.Terminal e il Kendriya.Terminal entrambi terminali di autobus da cui partono la maggior parte dei bus.

Closeness

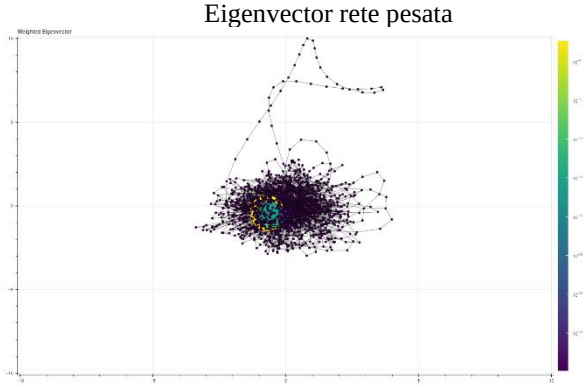


EIGENVECTOR

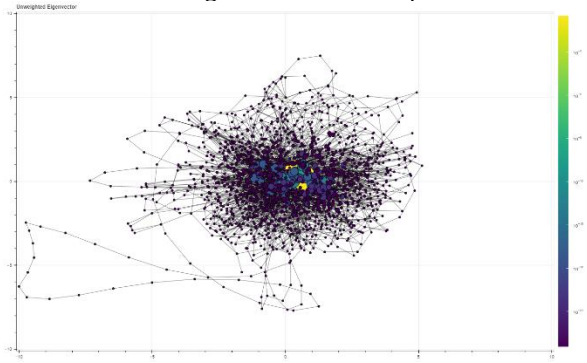


Per l'eigenvector i valori cambiano totalmente tra rete pesata e non. Nella rete pesata shivaji.park e madi.pur.village si distaccano molto dagli altri come si può vedere nel grafico, causato probabilmente per la loro connessione con nodi con alta centralità, dato che uno di loro rappresenta uno stadio e l'altro una meta turistica marittima.

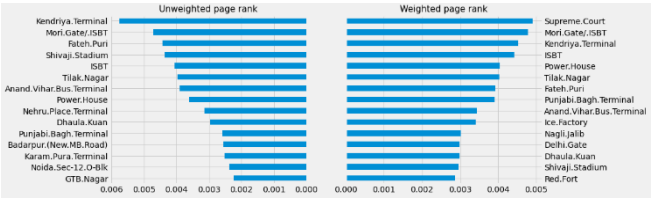
Eigenvector rete pesata



Eigenvector rete non pesata

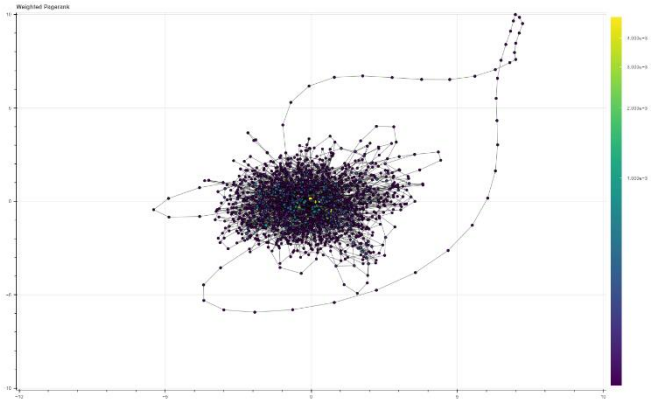


PAGE RANK

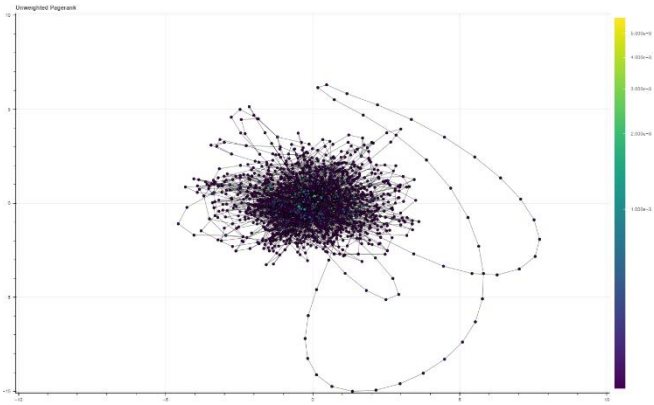


Per il page rank abbiamo al primo posto la supreme.court che ha un elevato numero di link in ingresso e i nodi che la linkano hanno un alto valore di centralità.

Page rank rete pesata



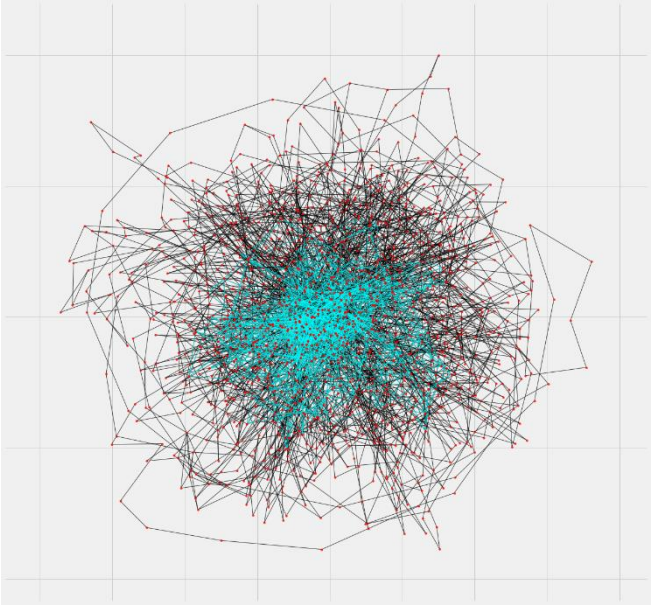
Page rank rete non pesata



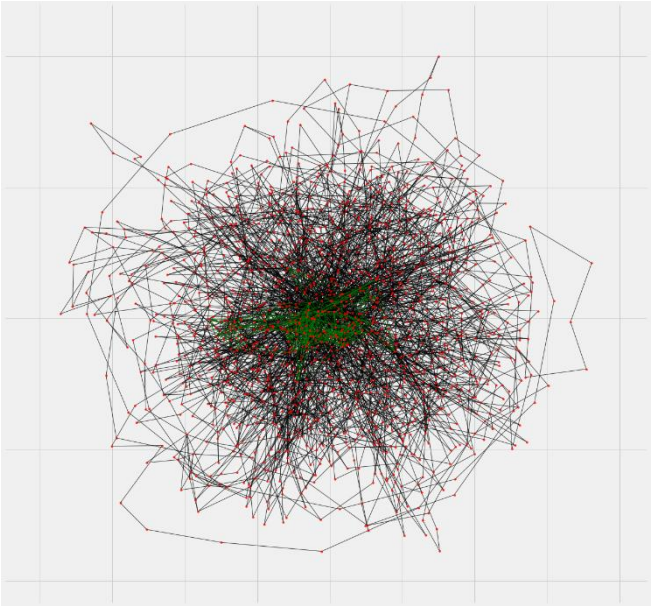
D. Medium Scale

Abbiamo cercato di capire se all'interno della rete fossero presenti delle communities attraverso il metodo K-core e facendolo iterare per tutti i valori possibili compresi tra 0 e 100. Per questa rete i valori più alti sono 4 e 5.

K-core con K=4



K-core con K=5



Questa volta sia con core 4 che 5 i gruppi sono molto numerosi anche se comunque con un core più basso si ha un maggior numero di nodi aggregati.

WEAK TIE

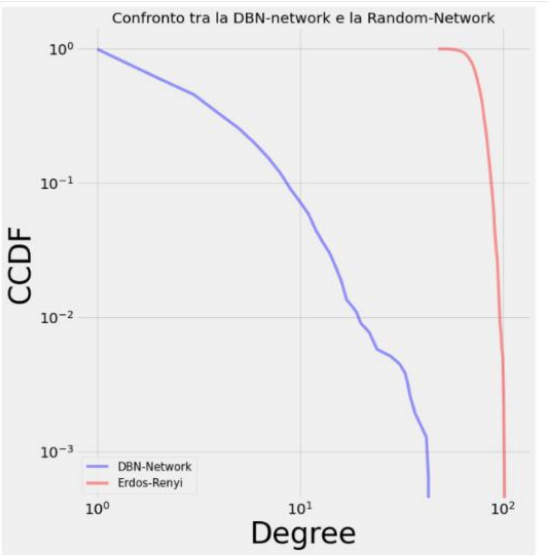
I legami deboli sono molto importanti all'interno di una rete poiché collegano diverse comunità.

```
Edge        Strength      Span
('Kendriya.Terminal', 'Gulabi.Bagh.Dda.Flats')  1       5
('Kendriya.Terminal', 'Kalkaji.DDA.Flats')      1       3
('Kendriya.Terminal', 'Bajit.Pur')      1       3
('Kendriya.Terminal', 'Nehru.Vihar')    1       3
('Kendriya.Terminal', 'Shalimar.Bagh.BH-Blk')   1       3
('Kendriya.Terminal', 'Jhingola.Village')       1       7
('Kendriya.Terminal', 'Holumbi.Kalan')  1       3
('Kendriya.Terminal', 'Manav.Chowk')    1       4
('Kendriya.Terminal', 'New.Kondali.Mkt')        1       4
('Kendriya.Terminal', 'Jagat.Pur.Temple')       1       6
('Kendriya.Terminal', 'Dilshad.Govt..School')   1       4
('Kendriya.Terminal', 'Trilok.Puri.27-Block')   1       3
('Kendriya.Terminal', 'Mayur.Vihar.Phase-II')   1       3
('Kendriya.Terminal', 'Dwarka.Metro.Station.Sec-14')    1       6
('Kendriya.Terminal', 'Khyala.Restt.Cly')       1       4
('Kendriya.Terminal', 'Adhyapak.Nagar') 1       3
('Kendriya.Terminal', 'Admn..Block')    1       3
('Kendriya.Terminal', 'Rohini.Sec-24.Pkt-18')   1       5
('Kendriya.Terminal', 'Neelwal.Village')        1       4
('Kendriya.Terminal', 'Bakar.Wala.JJ.Cly')      1       3
('Kendriya.Terminal', 'Kair.Village')   1       4
('Kendriya.Terminal', 'Ghuman.Hera.Village')    1       6
('Dr.RML.Hospital', 'Shadi.Pur.Depot')  1       3
```

E. Networks models

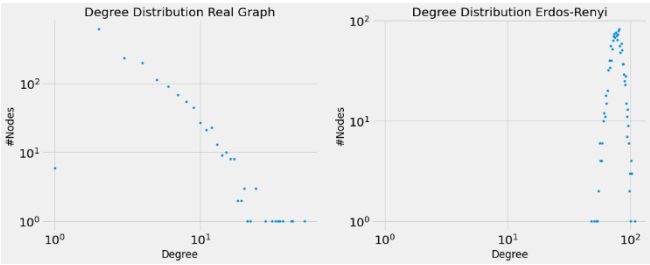
In ultima analisi abbiamo sviluppato 3 diversi modelli per simulare la nostra rete partendo da regole ben definite. Per ognuno di questi si è proceduto ad un confronto grafico della CCDF (distribuzione cumulativa, probabilità che un nodo abbia un certo grado o che sia maggiore di quell grado), ad un confronto su misure a larga scala e a un confronto della degree distribution.

RANDOM NETWORK ERDOS-RENYI



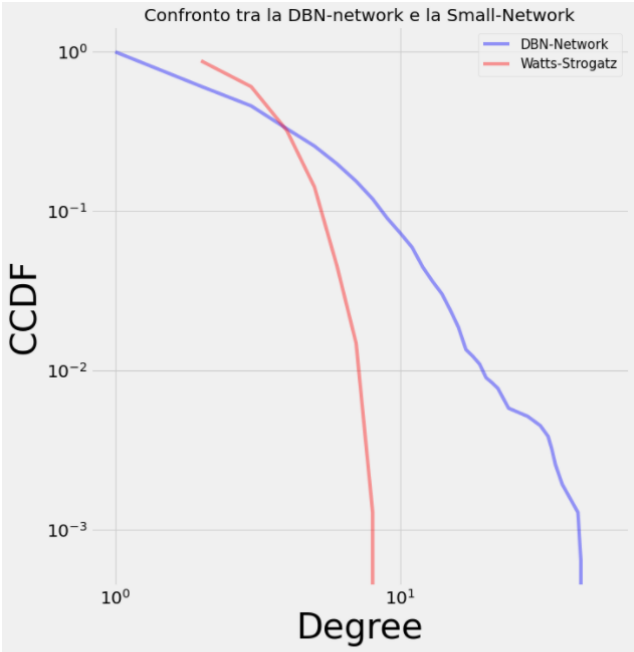

```
Average degree Random Network: 77.43629343629344
Average degree DBN Network: 4.633204633204633
Il global clustering della network DBN è:  0.122
Il global clustering della random network è:  0.05
L'APL della network DBN è:  5.536
L'APL della random network è:  1.97
```

Come si può notare dal grafico la CCDF si discosta molto dal similuzione della nostra rete. Questo modello simula una rete con average degree molto più alto della rete reale e APL decisamente più basso.



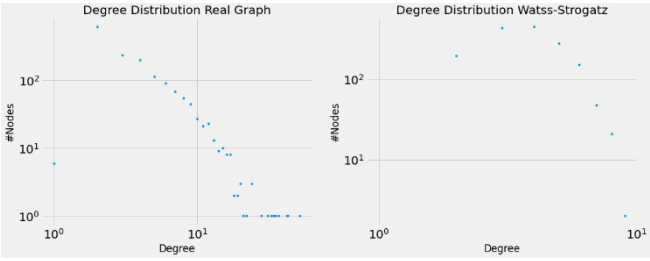
Anche la degree distribution è totalmente diversa. Possiamo concludere che il modello di Erdos-Renyi non è adatto a simulare la nostra rete.

WATTS-STROGATZ



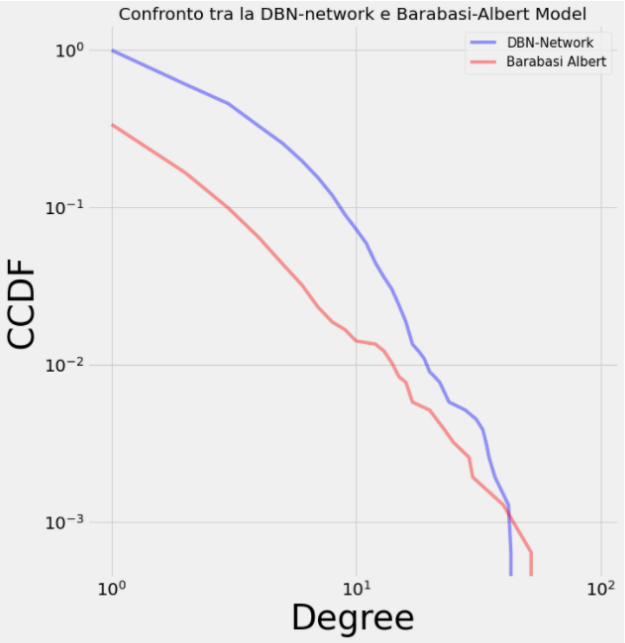

```
Average degree Small word: 4.0
Average degree DBN Network: 4.633204633204633
Il global clustering della network DBN è:  0.122
Il global clustering della Small word network è:  0.002
L'APL della network DBN è:  5.536
L'APL della small word network è:  5.704
```

In questo caso sia la CCDF che le misure a larga scala si avvicinano molto di più alla nostra rete, mentre il global clustering è decisamente più piccolo



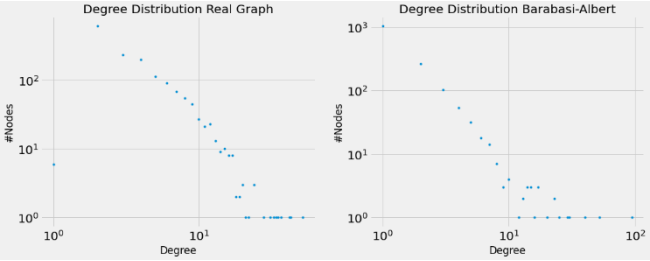
La degree distribution non si avvicina alla rappresentazione della nostra rete e anche in questo caso possiamo concludere affermando che questo modello non è adatto.

BARABASI ALBERT



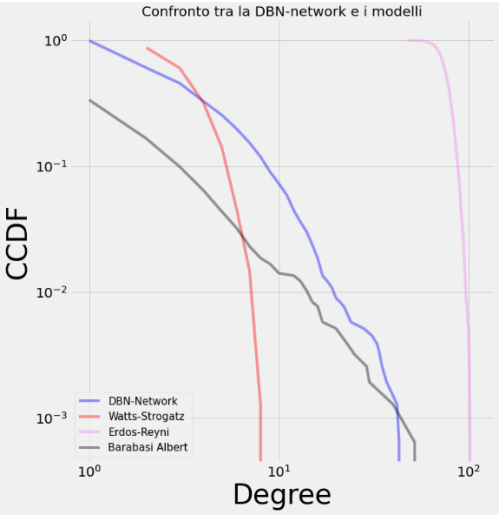

```
Average degree Barabasi-Albert: 1.9987129987129988
Average degree DBN Network: 4.633204633204633
Il global clustering della network DBN è:  0.122
Il global clustering del modello di Barabasi-Albert è:  0
L'APL della network DBN è:  5.536
L'APL del modello di Barabasi- Albert è:  7.45
```

A differenza dei modelli precedenti la CCDF si avvicina molto di più mentre per alcune large scale come, l'average degree il valore è più basso, al contrario per l'APL il valore è più alto



Per quanto riguarda la degree distribution segue abbastanza l'andamento della nostra rete.

Infine riportiamo un grafico di confornto fra tutte le CCDF.



Come nella prima città analizzata il modello di Barabasi basandoci sulla degree distribution è il migliore per la simulazione della nostra rete.

F. Conclusioni

In questa rete sono presenti diversi hubs ovvero i terminal degli autobus presenti in gran numero, che permettono molti collegamenti con molte tratte di autobus collegando cosi la città caratterizzata da un forte turismo. Abbiamo notato anche come lo Shivaji.stadium sia una meta molto gettonata seguito da alcune mete turistiche come madi.pur.village e Shivaji.park.

VI. Hyderabad

Hyderābād è una suddivisione dell'India, classificata come municipal corporation, capoluogo dei distretti di Hyderābād e Rangareddy, nello stato federato del Telangana. In base al numero di abitanti la città rientra nella classe I (da 100.000 persone in su).

Occupando 650 km quadrati lungo le rive del fiume Musi, ha una popolazione di circa 6,9 milioni e una popolazione metropolitana di circa 9,7 milioni, che la rende la quarta città più popolosa e la sesta agglomerazione urbana più popolosa in India. A un'altitudine media di 505 metri, gran parte di Hyderābād è situata su terreni collinari intorno a laghi artificiali, tra cui Hussain Sagar, che precede la fondazione della città, a nord del centro città.

A. Analisi della rete

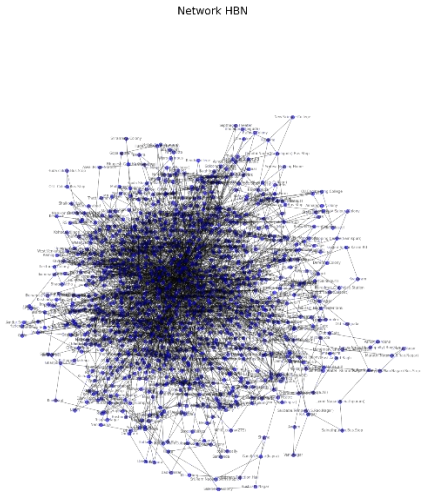
Il punto iniziale è stato il caricamento della Network e il calcolo dei nodi e link totali con successivo sviluppo di una matrice di adiacenza per una miglior visualizzazione dei dati.

```
NODI:
 1085

LINKS:
 2951
```

Grafico della Network

Una volta individuati i nodi e i link abbiamo graficato a schermo la Network attraverso la libreria networkx di Python



B. *Large Scale*

Ci permettono di comprendere la struttura generale della rete. Le metriche sono state calcolate considerando la rete pesata e non pesata.

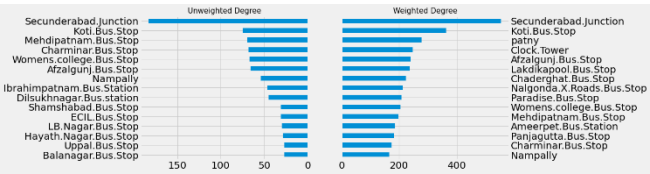
	Network non pesata	Network pesata
Average Path Length	3.8770	6.6790
Diametro	11.0000	11.0000
Density	0.0050	0.0050
Average Local Clustering	0.2597	0.0095
Global Clustering	0.1020	0.1020

Anche all'interno di questa rete il livello di density è abbastanza basso ma più alto delle altre reti, il diametro è il più piccolo tra le 4 reti analizzate.

C. *Small Scale*

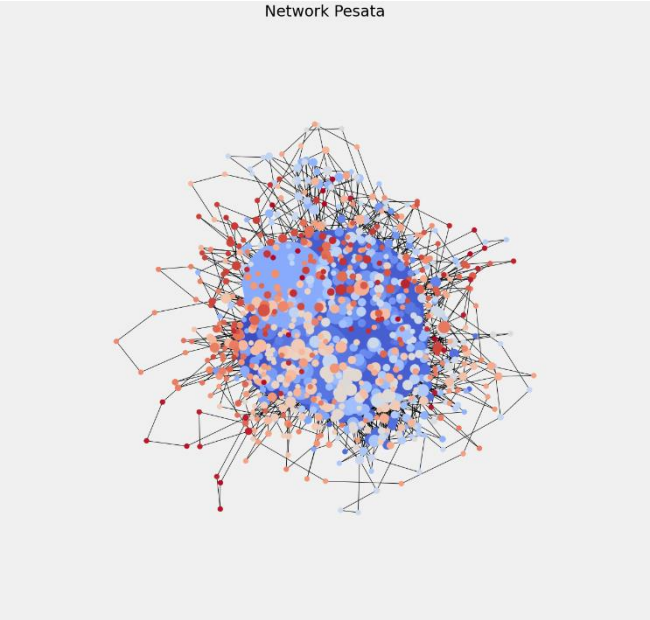
Ci consentono di valutare il ruolo dei nodi individuali all'interno della rete.
Per ogni metrica è stato eseguito un grafico rappresentante i 15 nodi con valore più alto relativo alla misura presa in considerazione.

DEGREE

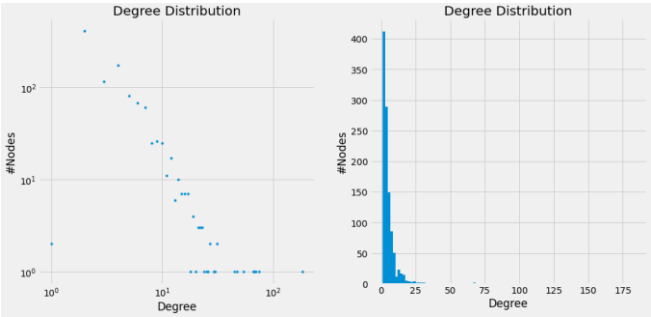


Nella rete non pesata e pesata i nodi con grado più alto sono il Secunderabad.juction e il koti.bus.stop rappresentanti stazioni.

Grafico della Network basato sul peso dei nodi

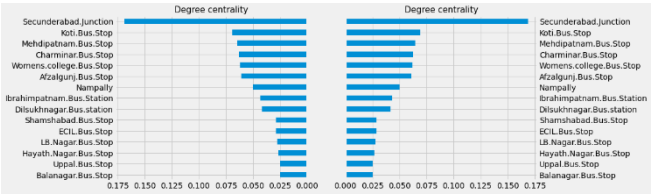


Istogramma della Degree distribution (fornisce la probabilità che un nodo selezionato casualmente nella rete ha grado K)



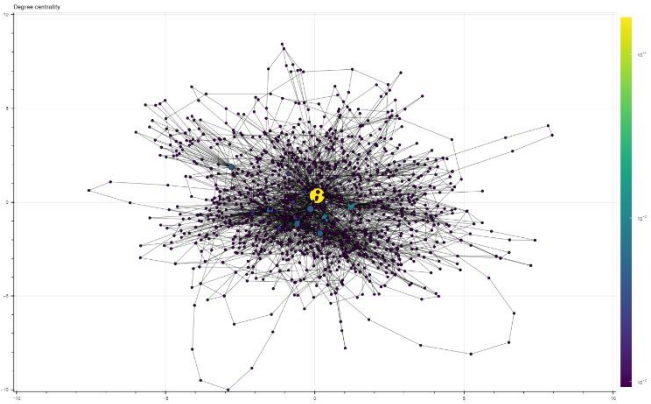
Da questo osserviamo che sono presenti molti nodi con grado basso

DEGREE CENTRALITY

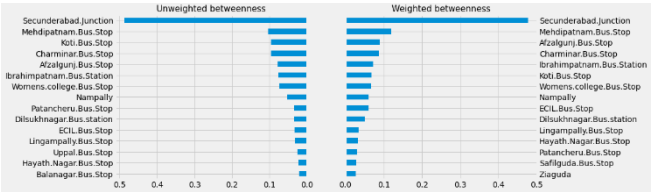


Per la degree centrality i nodi Secunderabad.juction e il koti.bus.stopsono quelli con valore più elavato senza differenze tra rete pesata e non.

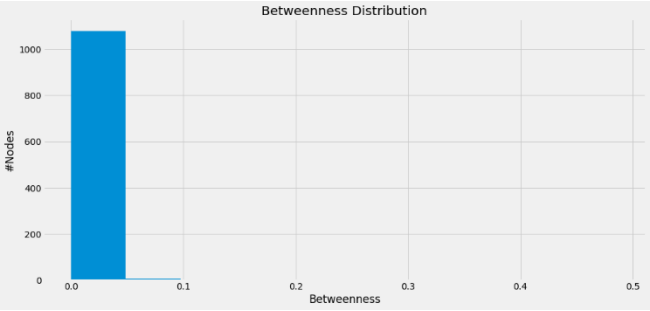
Da qui si inizia a vedere che questi due nodi possano risultare i più importanti nella rete, poiche uno dei due è la stazione principale della città.



BETWEENNESS

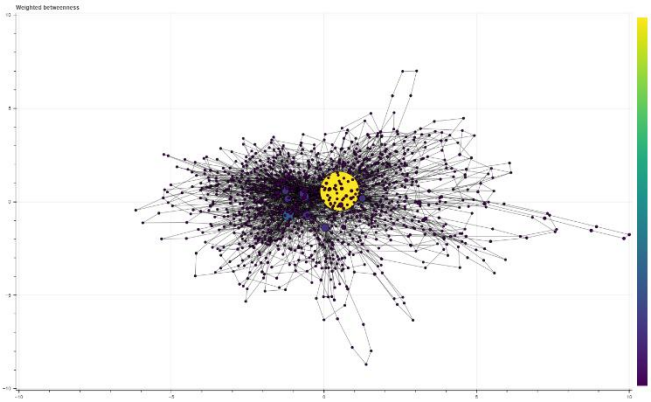


Come le metriche precedenti il nodo con valore più elevato è Secunderabad.juction per la rete non pesata, questo ci permette di affermare l'importanza di questo nodo che permette un collegamento tra nodi separati trovandosi spesso al centro degli shortest-path tra due nodi, poiché rappresenta la stazione principale di Secunderabad quartiere più importante della città.

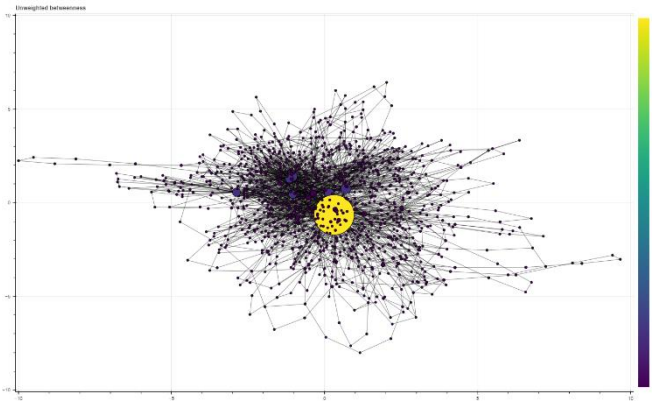


Il grafico conferma che nell'intera rete la maggioranza dei nodi ha un basso valore di betweenness.

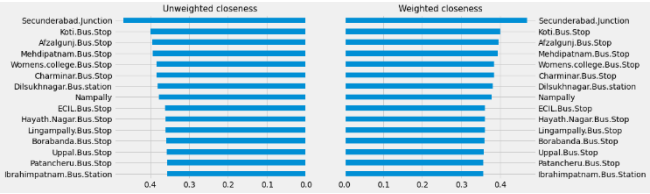
Betweenness rete pesata



Betweenness rete non pesata

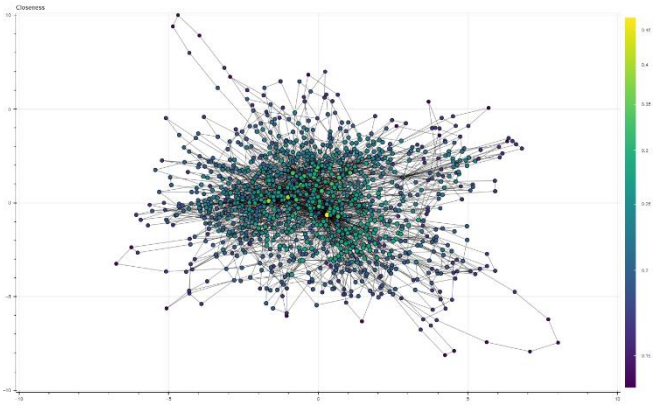


CLOSENESS

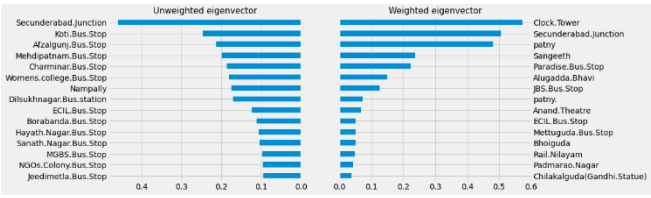


La closeness ci permette di stabilire la centralità di un nodo ovvero quanti nodi si possono raggiungere tramite gli shortest-path del nodo di partenza e non differisce tra rete pesata e non. Il valore di closeness dei 15 nodi con punteggio più alto è abbastanza omogeneo, in questo caso i primi sono il Secunderabad.juction e il koti.bus.stop terminali di autobus da cui partono la maggior parte dei bus.

Closeness

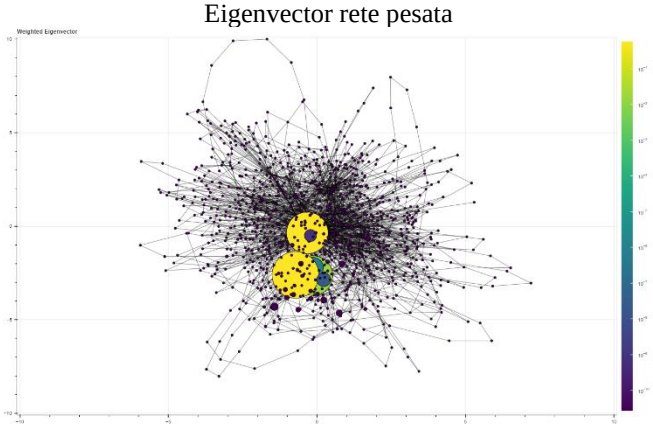


EIGENVECTOR

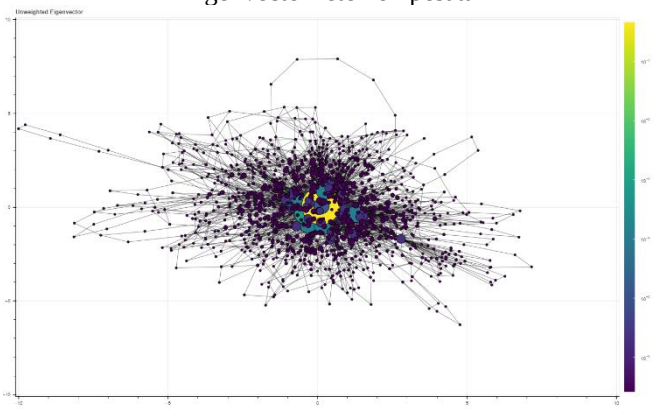


Per l'eigenvector i valori cambiano totalmente tra rete pesata e non confermando nella rete non pesata Secunderabad.juction e in quella pesata clock.tower essendo un monumento molto visitato presente nel quartiere di Secunderabad.

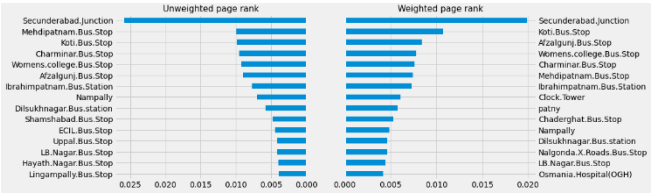
Eigenvector rete pesata



Eigenvector rete non pesata

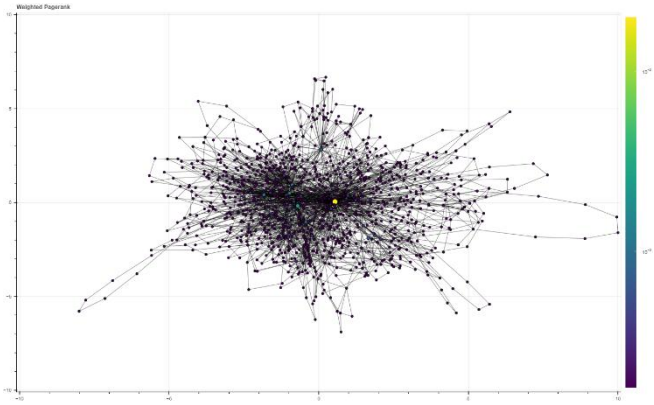


PAGE RANK

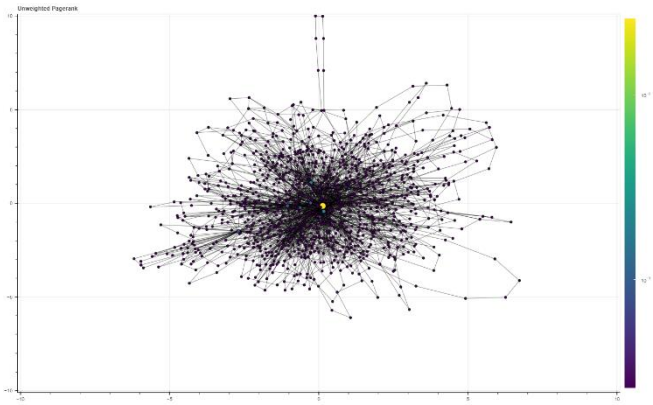


Per il page rank abbiamo al primo posto sempre Secunderabad.juction che ha un elevato numero di link in ingresso(essendo la città più visitata e popolosa) e i nodi che la linkano hanno un alto valore di centralità.

Page rank rete pesata



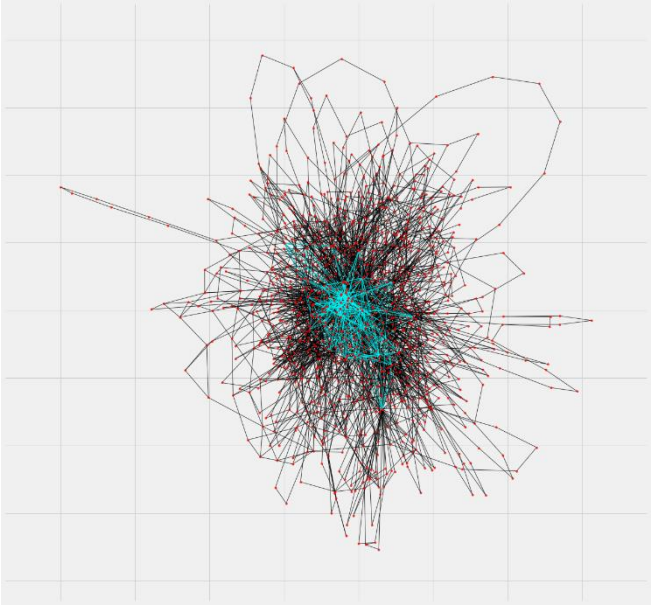
Page rank rete non pesata



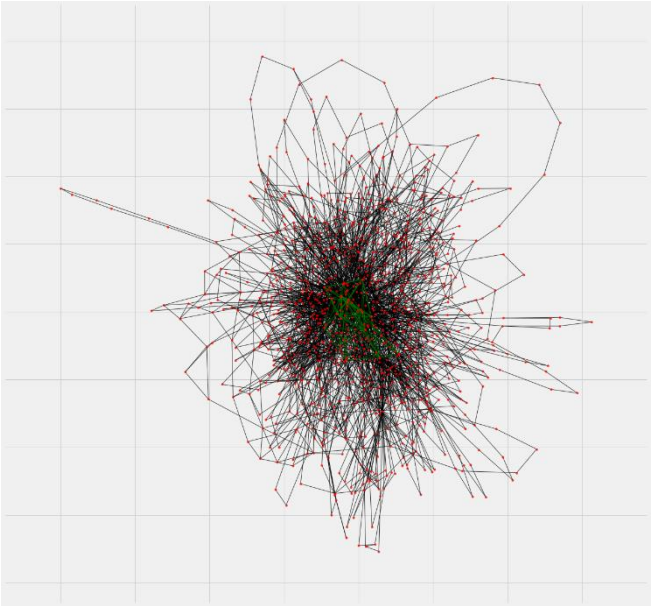
D. Medium Scale

Abbiamo cercato di capire se all'interno della rete fossero presenti delle communities attraverso il metodo K-core e facendolo iterare per tutti i valori possibili compresi tra 0 e 100. Per questa rete i valori più alti sono 6 e 7.

K-core con K=6



K-core con K=7



Sia con core di grado 6 che di grado 7 i gruppi non sono molto numerosi e tendono a raggrupparsi centralmente.

WEAK TIE

I legami deboli sono molto importanti all'interno di una rete poiché collegano diverse comunità.

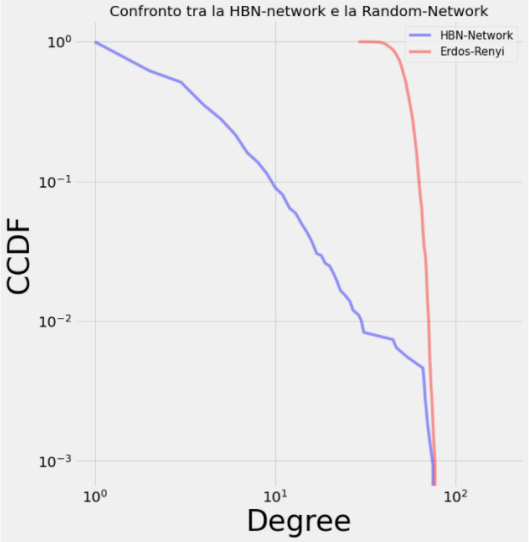
Si può individuare uno span infinito nell'edge che va da charminar.bus.stop (fermata autobus) a doodh.bowli, se quel ponte venisse eliminato si disconnetterebbe la rete.

```
('Afzalgunj.Bus.Stop', 'Gokhle.Nagar')  1       4
('Afzalgunj.Bus.Stop', 'Kokapeta.Lake') 1       3
('Afzalgunj.Bus.Stop', 'Mannyaguda')    1       3
('Afzalgunj.Bus.Stop', 'Dasrlapally')   1       3
('Afzalgunj.Bus.Stop', 'J.P.Durgah')    1       4
('CBS.Bus.Stop', 'Suraram')     1       3
('Charminar.Bus.Stop', 'A.G.Colony')    1       3
('Charminar.Bus.Stop', 'Babika.Chasma(Charminar)')      1       3
('Charminar.Bus.Stop', 'Doodh.Bowli')   1       inf
('Charminar.Bus.Stop', 'Bahadurpally')  1       3
('Charminar.Bus.Stop', 'HAL.Colony')    1       3
('Charminar.Bus.Stop', 'Santhosh.Nagar.Bus.Stand')      1       3
('Charminar.Bus.Stop', 'Lalitha.Bagh')  1       3
('Charminar.Bus.Stop', 'Teegala.Kunta') 1       3
('Charminar.Bus.Stop', 'Rajanna.Bai')   1       3
('Charminar.Bus.Stop', 'LB.Nagar.Bus.Sto')      1       3
('Charminar.Bus.Stop', 'Hafiz.Baba.Nagar')      1       4
('Charminar.Bus.Stop', 'Baba.Nagar')    1       3
('Charminar.Bus.Stop', 'Jalapalli')     1       3
('Charminar.Bus.Stop', 'Kalwakurthy.Bus.Stop')  1       3
('LB.Nagar.Bus.Stop', 'FCI.Colony')     1       3
('LB.Nagar.Bus.Stop', 'Laxmiguda')      1       3
('LB.Nagar.Bus.Stop', 'Thorrur')        1       4
```

E. Networks models

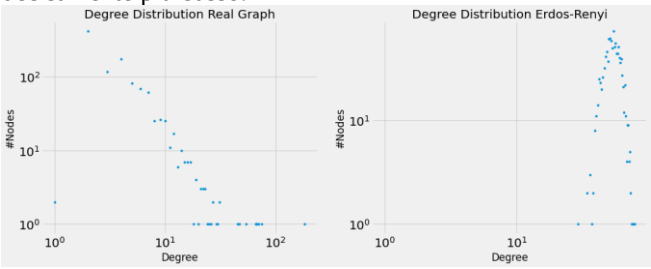
In ultima analisi abbiamo sviluppato 3 diversi modelli per simulare la nostra rete partendo da regole ben definite. Per ognuno di questi si è proceduto ad un confronto grafico della CCDF (distribuzione cumulativa, probabilità che un nodo abbia un certo grado o che sia maggiore di quell grado), ad un confronto su misure a larga scala e a un confronto della degree distribution.

RANDOM NETWORK ERDOS-RENYI



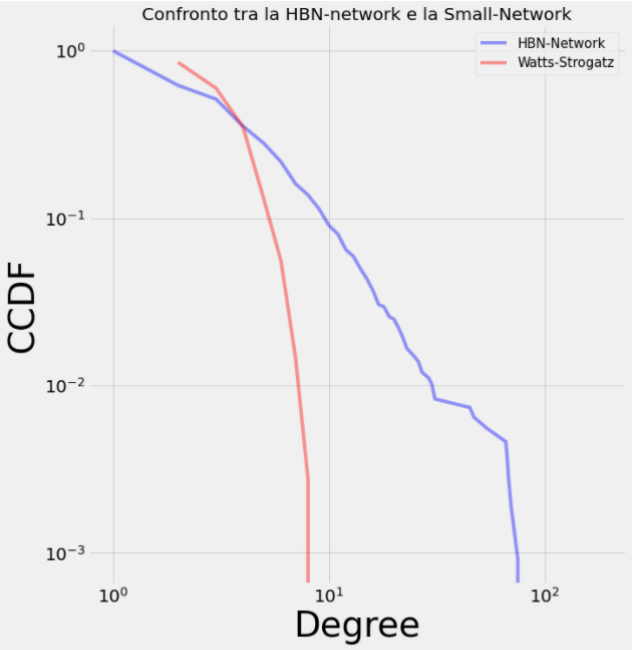

```
Average degree Random Network: 53.93364055299539
Average degree HBN Network: 5.4396313364055295
Il global clustering della network HBN è:  0.102
Il global clustering della random network è:  0.05
L'APL della network HBN è:  6.679
L'APL della random network è:  2.016
```

Come si può notare dal grafico la CCDF si discosta molto dal similuzione della nostra rete. Questo modello simula una rete con average degree molto più alto della rete reale e APL decisamente più basso.



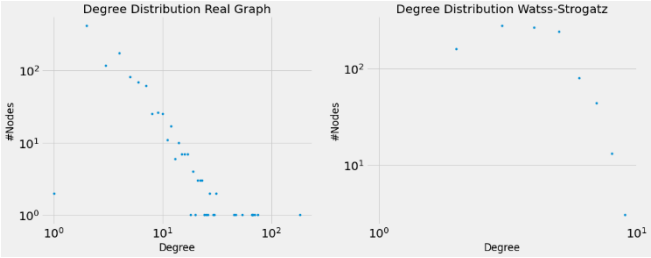
Anche la degree distribution è totalmente diversa. Possiamo concludere che il modello di Erdos-Renyi non è adatto a simulare la nostra rete.

WATTS-STROGATZ



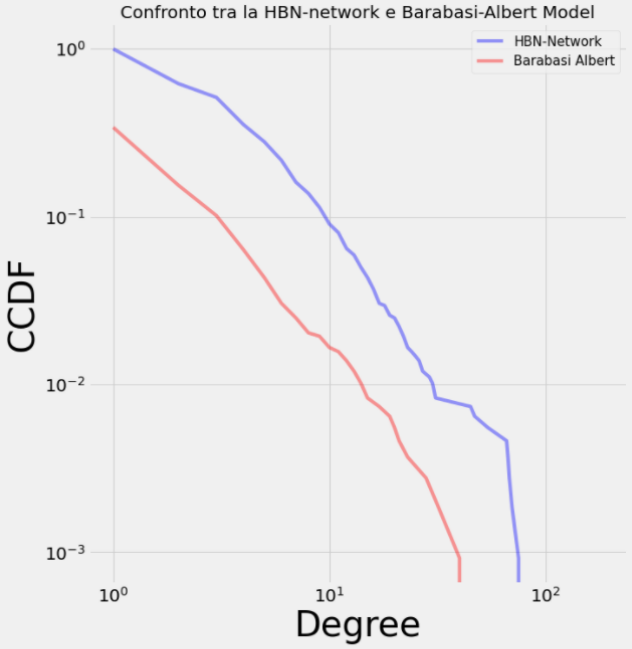

```
Average degree Small word: 4.0
Average degree HBN Network: 5.4396313364055295
Il global clustering della network HBN è:  0.102
Il global clustering della Small word network è:  0.004
L'APL della network HBN è:  6.679
L'APL della small word network è:  5.392
```

In questo caso la CCDF, l'average degree e l'APL si avvicinano molto di più alla nostra rete, mentre il global clustering è decisamente più piccolo



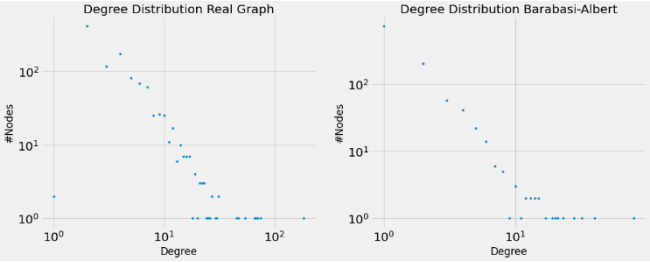
La degree distribution non si avvicina alla rappresentazione della nostra rete e anche in questo caso possiamo concludere affermando che questo modello non è adatto.

BARABASI ALBERT



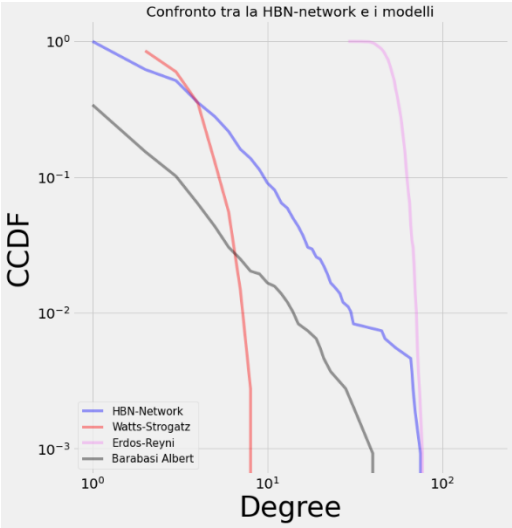

```
Average degree Barabasi-Albert: 1.9981566820276497
Average degree HBN Network: 5.4396313364055295
Il global clustering della network HBN è:  0.102
Il global clustering del modello di Barabasi-Albert è:  0
L'APL della network HBN è:  6.679
L'APL del modello di Barabasi- Albert è:  7.038
```

A differenza dei modelli precedenti la CCDF si avvicina molto di più mentre per alcune large scale come, l'average degree il valore è più basso, diversamente per l'APL il valore è simile.



Per quanto riguarda la degree distribution segue abbastanza l'andamento della nostra rete.

Infine riportiamo un grafico di confornto fra tutte le CCDF.



Il modello di Barabasi come in altri casi è il modello migliore per simulare la network.

F. Conclusioni

Essendo Hyderabad una città molto grande è composta da diverse stazioni e quartieri che corrispondo agli hubs nella network. La più importante è quella di Secunderabad essendo anche il quartiere più popoloso e largamente collegato con gli altri quartieri del distretto.

VII. KOLKATA

Calcutta è una città dell'India, capitale dello Stato federato del Bengala Occidentale. Sorge sulla sponda est del fiume Hooghly ed è il principale centro commerciale, culturale e formativo dell'India orientale. Il porto è il più antico dell'intero Stato e l'unico costruito su un fiume. In base al numero di abitanti, la città rientra nella classe I (più di 100 000 abitanti) ed è la terza in India dopo Mumbai e Delhi.

A. Analisi della rete

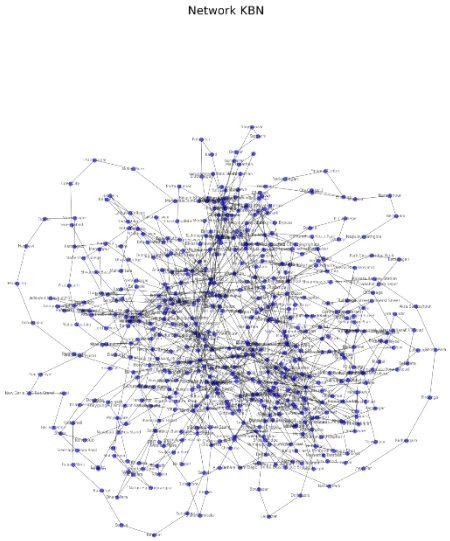
Il punto iniziale è stato il caricamento della Network e il calcolo dei nodi e link totali con successivo sviluppo di una matrice di adiacenza per una miglior visualizzazione dei dati.

```
NODI:
519

LINKS:
878
```

Grafico della Network

Una volta individuati i nodi e i link abbiamo graficato a schermo la Network attraverso la libreria networkx di Python



B. *Large Scale*

Ci permettono di comprendere la struttura generale della rete. Le metriche sono state calcolate considerando la rete pesata e non pesata.

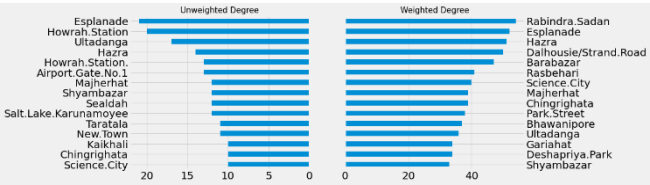
	Network non pesata	Network pesata
Average Path Length	5.7710	7.5670
Diametro	16.0000	16.0000
Density	0.0065	0.0065
Average Local Clustering	0.0793	0.0061
Global Clustering	0.0760	0.0760

Il diametro della rete è 16 indicandoci che lo shortest-path più grande non è troppo lungo, quindi possiamo pensare che le fermate più distanti si trovino a 16 edges, anche la density è bassa.

C. *Small Scale*

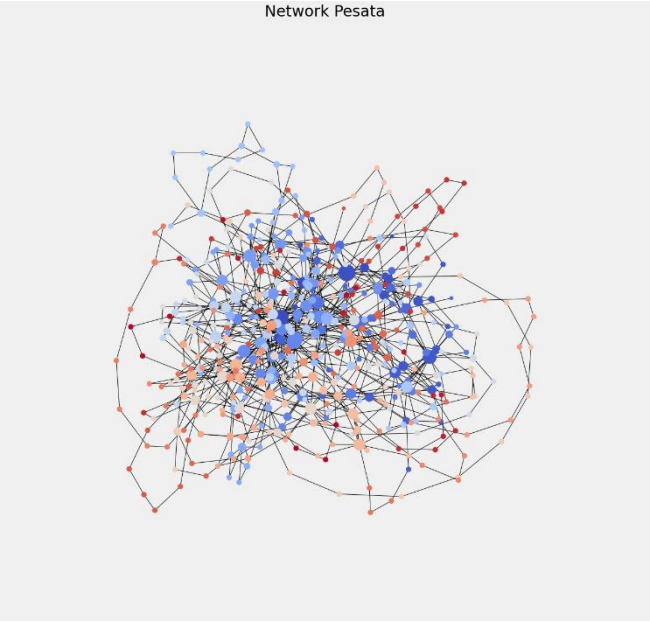
Ci consentono di valutare il ruolo dei nodi individuali all'interno della rete.
Per ogni metrica è stato eseguito un grafico rappresentante i 15 nodi con valore più alto relativo alla misura presa in considerazione.

DEGREE

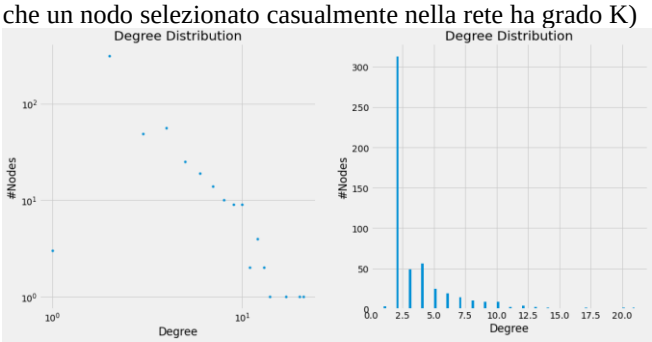


In questa rete otteniamo per la prima volta due diversi risultati dalla rete pesata e non. Per la non pesata i nodi con grado più alto sono l'Esplanade e Howrah.Station rappresentante una stazione.
Nella rete pesata abbiamo Rabindra.Sadan ed Esplanade.

Grafico della Network basato sul peso dei nodi

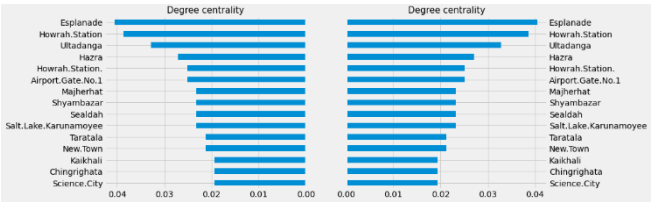


Istogramma della Degree distribution (fornisce la probabilità che un nodo selezionato casualmente nella rete ha grado K)



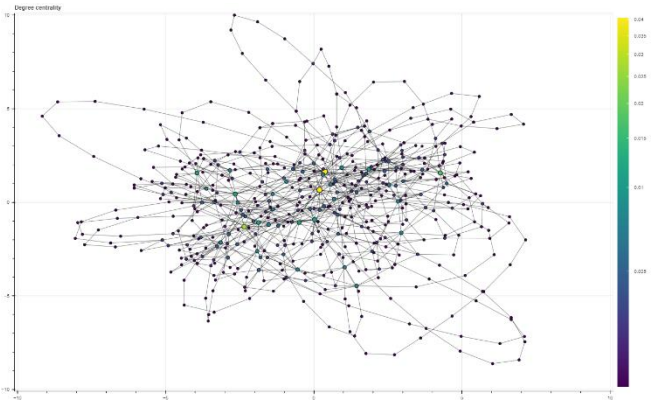
Da questo osserviamo che sono presenti molti nodi con grado basso

DEGREE CENTRALITY

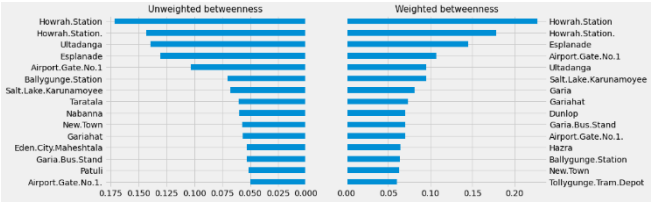


Per la degree centrality i nodi Esplanade e Howrah.Station sono quelli con valore più elavato senza differenze tra rete pesata e non.

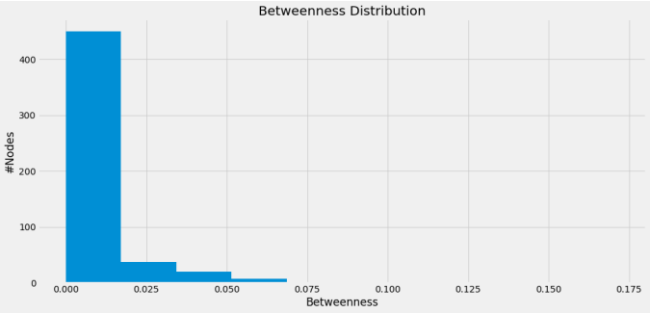
Da qui si inizia a vedere che questi due nodi possano risultare i più importanti nella rete, poiche Esplanade è il quartiere centrale storico di Calcutta.



BETWEENNESS

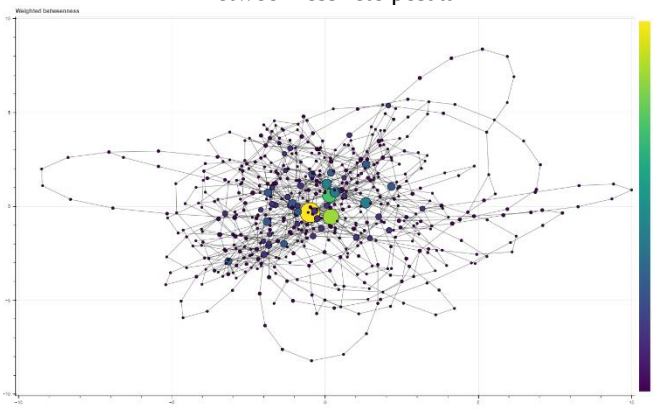


A differenze delle prime metriche il nodo con valore più elevato è Howrah.Station per la rete pesata e non, questo ci permette di affermare l'importanza di questo nodo che permette un collegamento tra nodi separati trovandosi spesso al centro degli shortest-path tra due nodi.

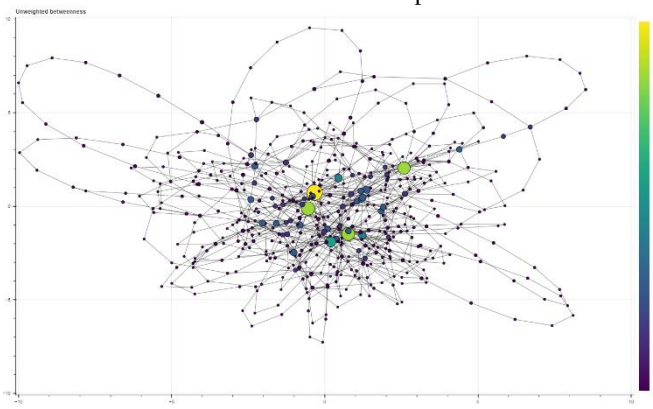


Il grafico conferma che nell'intera rete la maggioranza dei nodi ha un basso valore di betweenness.

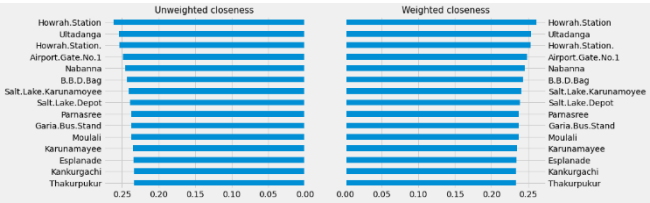
Betweenness rete pesata



Betweenness rete non pesata

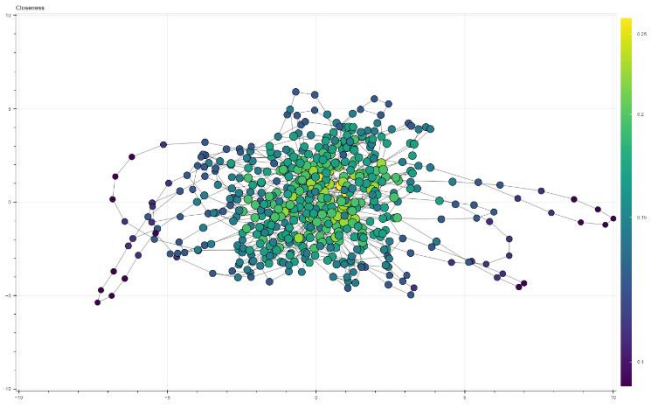


CLOSENESS

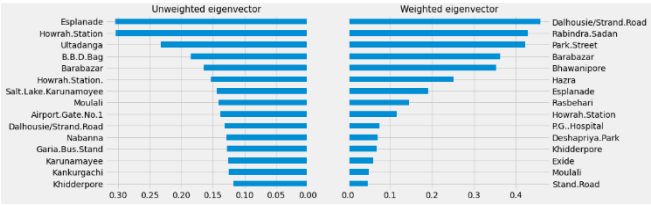


La closeness ci permette di stabilire la centralità di un nodo ovvero quanti nodi si possono raggiungere tramite gli shortest-path del nodo di partenza e non differisce tra rete pesata e non. Il valore di closeness dei 15 nodi con punteggio più alto è abbastanza omogeneo, in questo caso i primi sono Ultadanga e Howrah.Station due stazioni principali di Calcutta.

Closeness

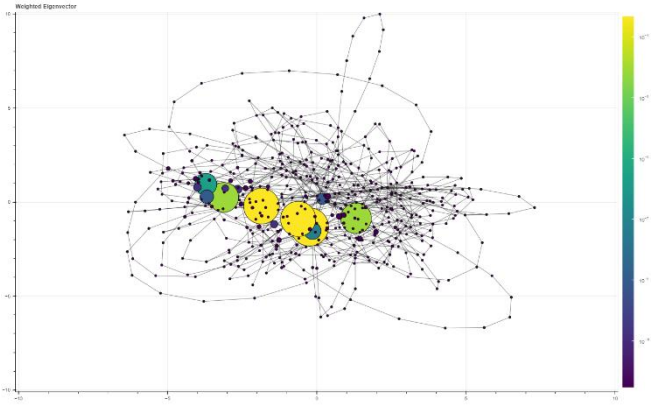


EIGENVECTOR

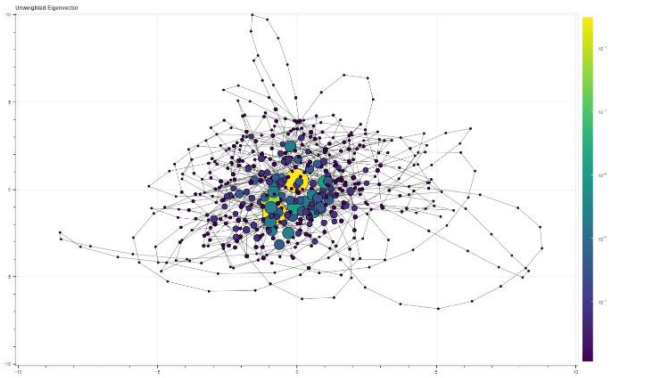


Per l'eigenvector i valori cambiano totalmente tra rete pesata e non, confermando nella rete non pesata Esplanade, mentre in quella pesata dalhousie/strand.road emerge visto che è una strada molto importante di calcutta

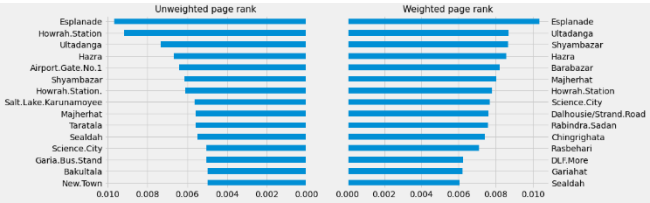
Eigenvector rete pesata



Eigenvector rete non pesata

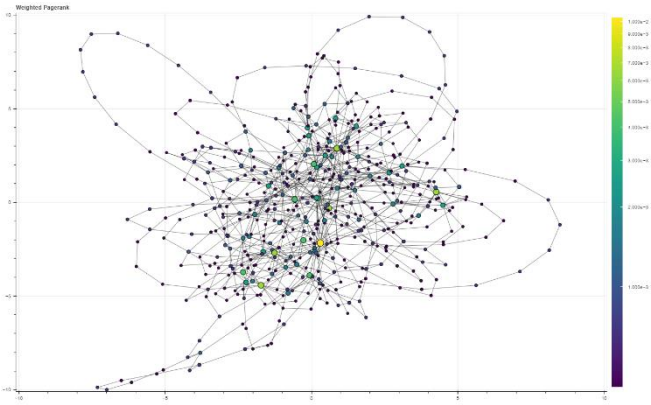


PAGE RANK

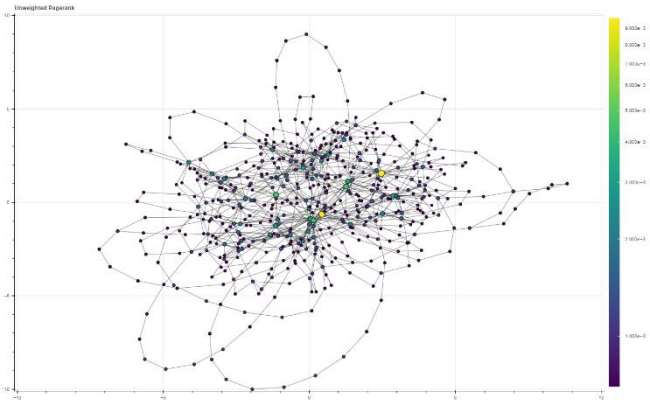


Per il page rank abbiamo al primo posto sempre esplanade che ha un elevato numero di link in ingresso(essendo il punto storico più visitato) e i nodi che la linkano hanno un alto valore di centralità.

Page rank rete pesata



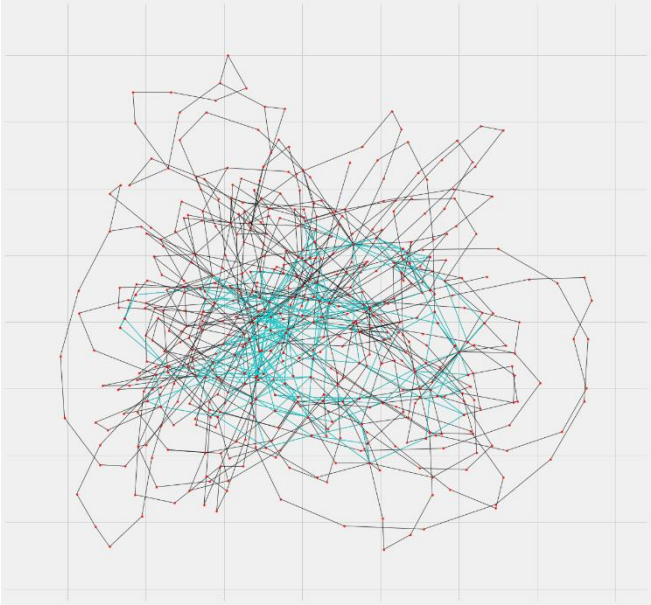
Page rank rete non pesata



D. Medium Scale

Abbiamo cercato di capire se all'interno della rete fossero presenti delle communities attraverso il metodo K-core e facendolo iterare per tutti i valori possibili compresi tra 0 e 100. Per questa rete il valore più alto è 3.

K-core con K=3



Per questa network abbiamo deciso di utilizzare un k-core con grado 3 poiché era l'unico sviluppabile, qui anche nodi più periferici tendono a raggrupparsi

WEAK TIE

I legami deboli sono molto importanti all'interno di una rete poiché collegano diverse comunità.

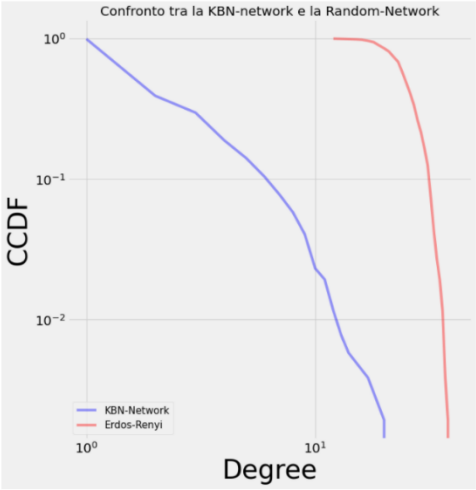
Si può individuare uno span infinito nell'edge che va da dalhousie/strand.road park.street, se quell' edge venisse eliminato si disconnetterebbe la rete.

```
Edge        Strength        Span
('Howrah.Station', 'Nabanna')   1       3
('Howrah.Station', 'Howrah.Maidan.')    1       5
('Howrah.Station', 'Taratala.') 1       3
('Howrah.Station', 'Moulali')   1       3
('Howrah.Station', 'Garia.Bus.Stand.')  1       3
('Howrah.Station', 'Sona.Mukhi.BAZAR')  1       5
('Howrah.Station', 'Dumdum.Station.')   1       5
('Howrah.Station', 'Howrah.Court.Foresore.Road')        1       3
('Howrah.Station', 'Saltlake.Sec.5')    1       3
('Howrah.Station', 'Thakurpukur')       1       4
('Howrah.Station', 'Raipur')    1       6
('Howrah.Station', 'Nilgunge')  1       6
('Howrah.Station', 'Baruipur')  1       4
('Howrah.Station', 'Sapoorji.Estate')   1       4
('Barabazar', 'Dalhousie/Strand.Road')  1       3
('Barabazar', 'Manicktala')     1       4
('Barabazar', 'Thakurpukur.')   1       4
('Barabazar', 'Mahatma.Gandhi.Road')    1       3
('Barabazar', 'M.G.Road')       1       3
('Barabazar', 'MG.Road')        1       5
('Dalhousie/Strand.Road', 'Park.Street.')       1       inf
('Manicktala', 'Ultadanga')     1       4
('Thakurpukur.', 'Kankurgachi.')        1       3
```

E. Networks models

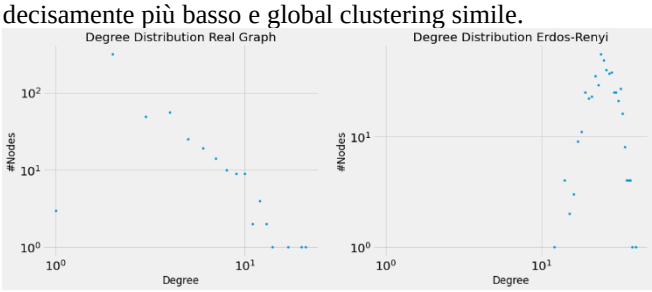
In ultima analisi abbiamo sviluppato 3 diversi modelli per simulare la nostra rete partendo da regole ben definite. Per ognuno di questi si è proceduto ad un confronto grafico della CCDF (distribuzione cumulativa, probabilità che un nodo abbia un certo grado o che sia maggiore di quell grado), ad un confronto su misure a larga scala e a un confronto della degree distribution.

RANDOM NETWORK ERDOS-RENYI



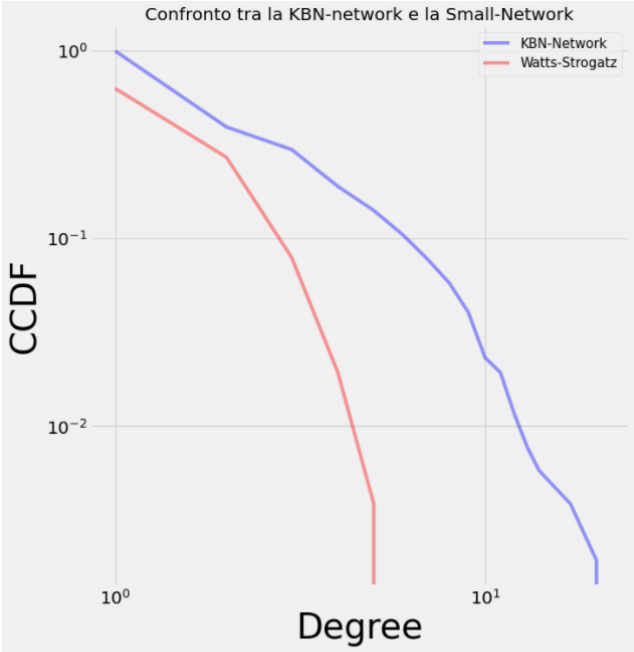

```
Average degree Random Network: 25.58766859344894
Average degree KBN Network: 3.3834296724470136
Il global clustering della network KBN è:  0.076
Il global clustering della random network è:  0.049
L'APL della network KBN è:  7.567
L'APL della random network è:  2.219
```

Come si può notare dal grafico la CCDF si discosta molto dal similuzione della nostra rete. Questo modello simula una rete con average degree molto più alto della rete reale e APL decisamente più basso e global clustering simile.



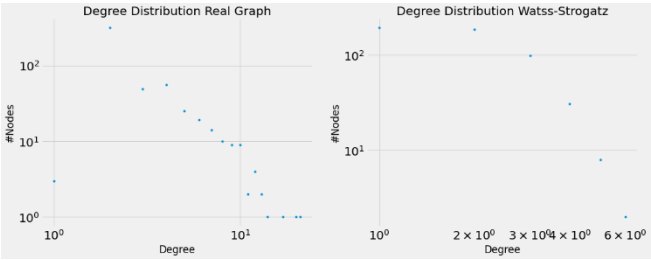
Anche la degree distribution è totalmente diversa. Possiamo concludere che il modello di Erdos-Renyi non è adatto a simulare la nostra rete.

WATTS-STROGATZ



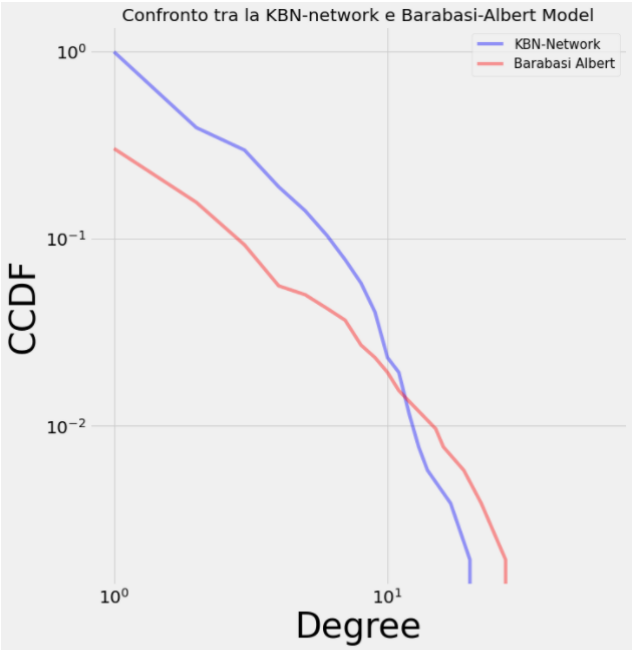

```
Average degree Small word: 2.0
Average degree KBN Network: 3.3834296724470136
Il global clustering della network KBN è:  0.076
Il global clustering della Small word network è:  0
```

In questo caso la CCDF, l'average degree si avvicinano molto di più alla nostra rete



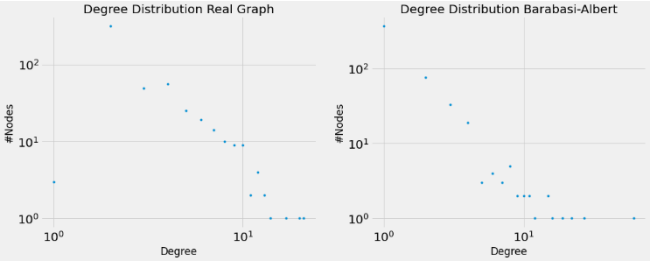
La degree distribution non si avvicina alla rappresentazione della nostra rete e anche in questo caso possiamo concludere affermando che questo modello non è adatto.

BARABASI ALBERT



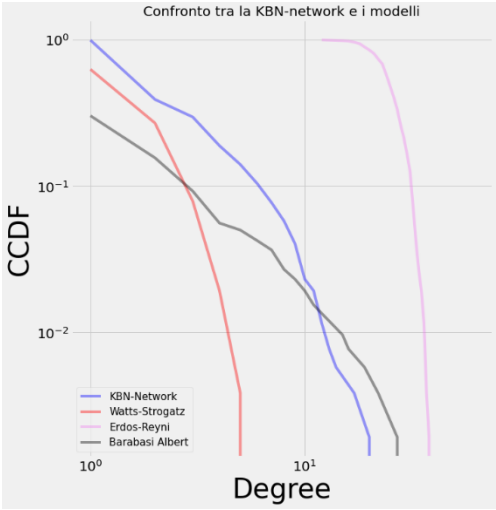

```
Average degree Barabasi-Albert: 1.9961464354527938
Average degree KBN Network: 3.3834296724470136
Il global clustering della network KBN è:  0.076
Il global clustering del modello di Barabasi-Albert è:  0
L'APL della network KBN è:  7.567
L'APL del modello di Barabasi- Albert è:  6.204
```

A differenza dei modelli precedenti la CCDF si avvicina molto di più mentre per alcune large scale come, l'average degree il valore è più basso, diversamente per l'APL il valore si avvicina.



Per quanto riguarda la degree distribution segue abbastanza l'andamento della nostra rete.

Infine riportiamo un grafico di confornto fra tutte le CCDF.



Come si sta confermando in quasi tutte le reti il modello di Barabasi è sempre quello che riesce a simulare la rete in maniera migliore.

F. Conclusioni

Anche se Calcutta è la terza città più grande dell'India e una delle più turistiche il numero di fermate degli autobus non è eccessivo e si può notare che anche qui gli hubs sono le stazioni tranne che per alcuni casi per esempio il centro storico di Esplanade e l'mportnte via di dalhousie/strand.road.

VIII. MUMBAI

Mumbai ufficialmente Bombay fino al 1995, è una città dell'India, capitale dello stato del Maharashtra. Con una popolazione stimata di 13 198 662 di abitanti, e con i suoi 31 214,47 ab/km² è la prima città per densità di popolazione al mondo. Insieme con la vicina periferia di Navi Mumbai e Thane, forma un agglomerato urbano di 21 milioni di abitanti, rendendola tredicesima area metropolitana più popolosa del pianeta. Mumbai si trova sulla costa occidentale dell'India e possiede un profondo porto naturale, che movimenta quasi la metà del traffico marittimo merci dell'India

Mumbai è la capitale commerciale e dell'intrattenimento dell'India. Genera il 5% del PIL del paese, il 25% della produzione industriale, il 40% del commercio marittimo, e il 70% delle transazioni di capitali dell'economia indiana.

A. Analisi della rete

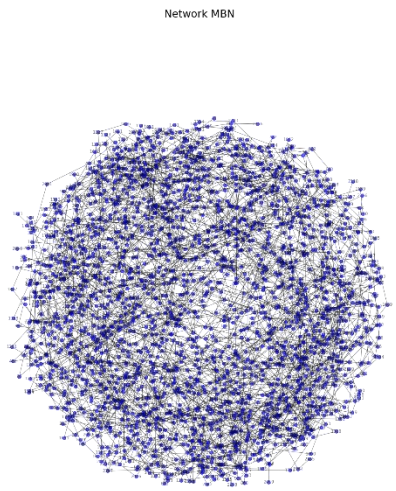
Il punto iniziale è stato il caricamento della Network e il calcolo dei nodi e link totali con successivo sviluppo di una matrice di adiacenza per una miglior visualizzazione dei dati.

```
NODI:
 2266

LINKS:
 3042
```

Grafico della Network

Una volta individuati i nodi e i link abbiamo graficato a schermo la Network attraverso la libreria networkx di Python



B. *Large Scale*

Ci permettono di comprendere la struttura generale della rete. Le metriche sono state calcolate considerando la rete pesata e non pesata.

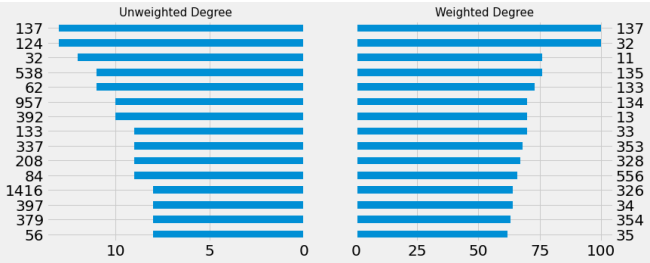
	Network non pesata	Network pesata
Average Path Length	25.6090	103.0780
Diametro	98.0000	98.0000
Density	0.0012	0.0012
Average Local Clustering	0.1490	0.0158
Global Clustering	0.1880	0.1880

In questo caso la differenza tra average path length della rete pesata e non pesata è alta, il diametro è il più grande di tutte le reti e la densità è molto bassa.

C. *Small Scale*

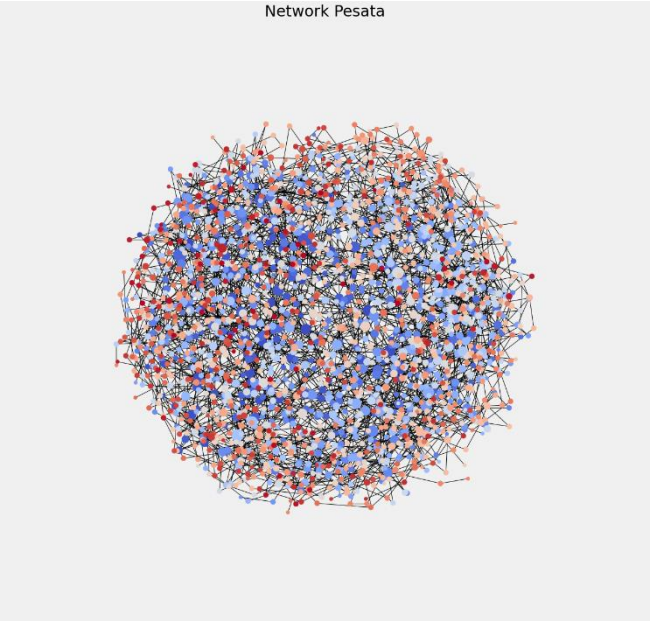
Ci consentono di valutare il ruolo dei nodi individuali all'interno della rete.
Per ogni metrica è stato eseguito un grafico rappresentante i 15 nodi con valore più alto relativo alla misura presa in considerazione.

DEGREE

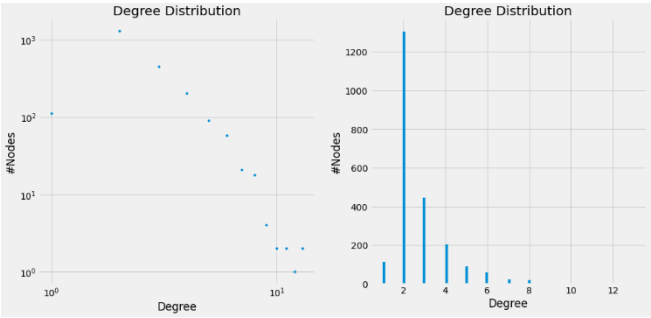


Nella rete non pesata e pesata i nodi con grado più alto sono il 137, 124 e il 32

Grafico della Network basato sul peso dei nodi

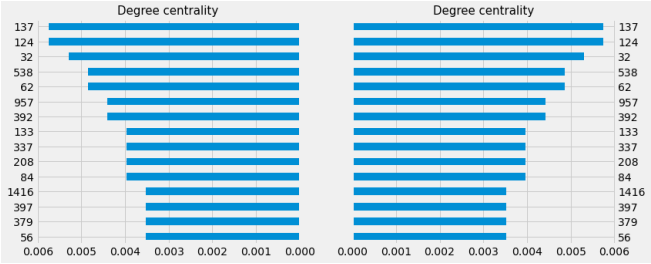


Istogramma della Degree distribution (fornisce la probabilità che un nodo selezionato casualmente nella rete ha grado K)



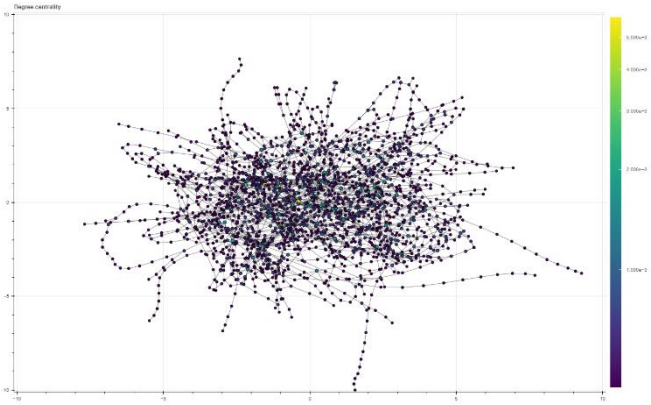
Da questo osserviamo che sono presenti molti nodi con grado 2

DEGREE CENTRALITY

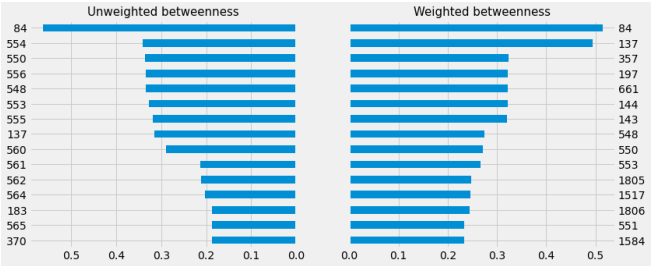


Per la degree centrality i nodi 137 e 124 sono quelli con valore più elavato senza differenze tra rete pesata e non.

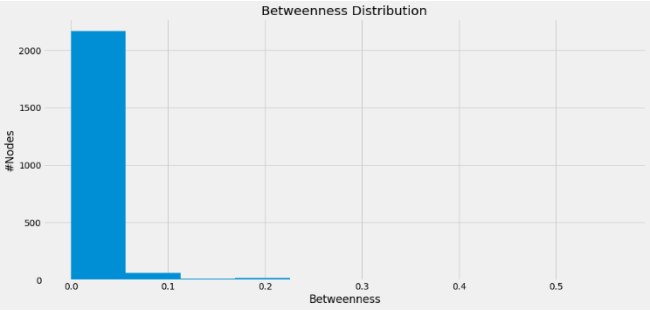
Anche se non abbiamo il nome di questi due nodi siamo in grado di poter dire che molto probabilmente si tratta di stazioni.



BETWEENNESS

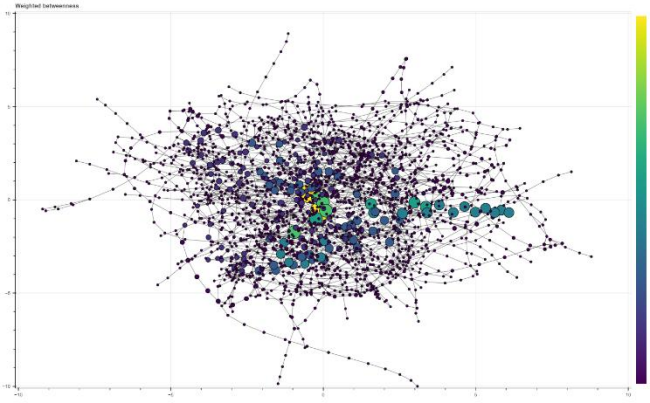


Il nodo con valore più alto è l'84 a differenza delle metriche precedenti, questo ci permette di affermare l'importanza di questo nodo che permette un collegamento tra nodi separati trovandosi spesso al centro degli shortest-path tra due nodi.

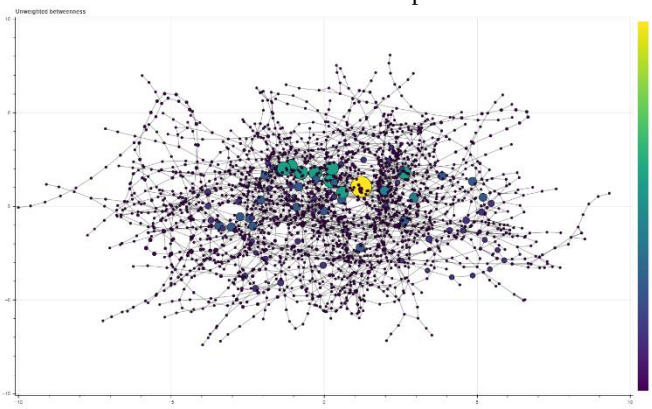


Il grafico conferma che nell'intera rete la maggioranza dei nodi ha un basso valore di betweenness.

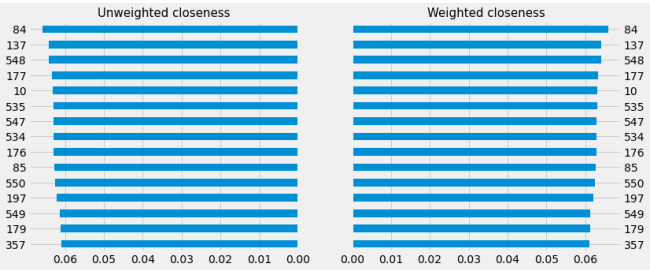
Betweenness rete pesata



Betweenness rete non pesata

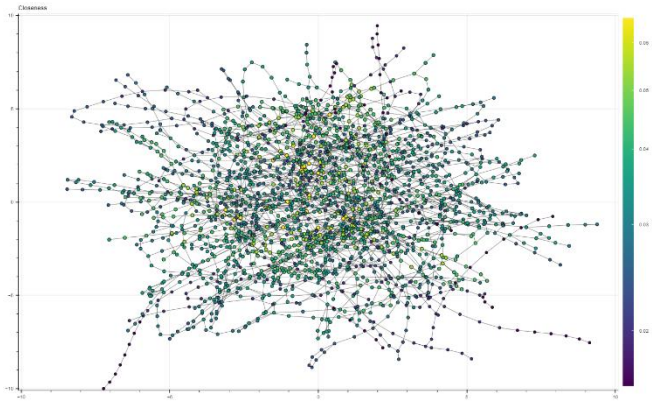


CLOSENESS

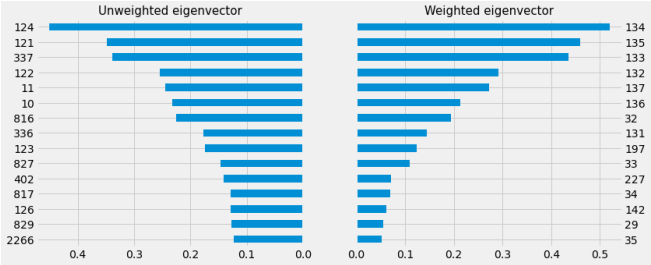


La closeness ci permette di stabilire la centralità di un nodo ovvero quanti nodi si possono raggiungere tramite gli shortest-path del nodo di partenza e non differisce tra rete pesata e non. Il valore di closeness dei 15 nodi con punteggio più alto è abbastanza omogeneo. Rimane come valore più alto l'84 come nella metrica precedente.

Closeness

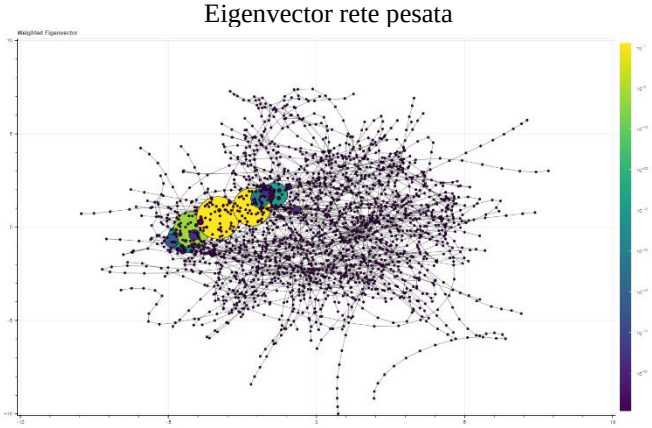


EIGENVECTOR

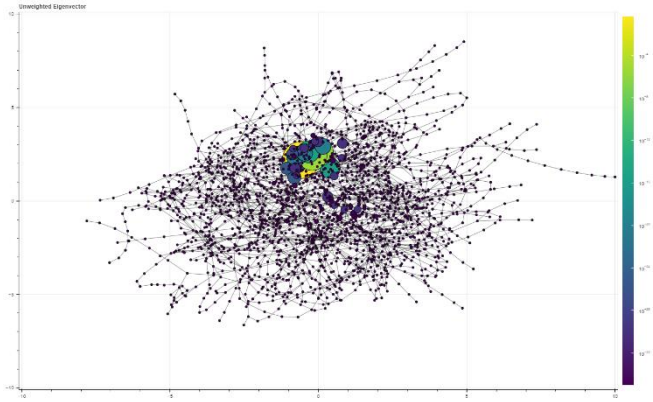


Per l'eigenvector i valori cambiano totalmente tra rete pesata e non pesata. In quella pesata 134,135,133 hanno un volore decisamente più elevato degli altri mentre nella non pesata è il 124.

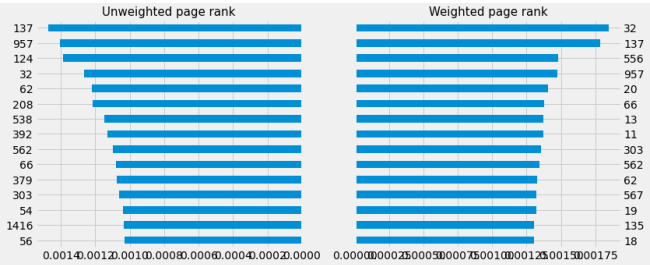
Eigenvector rete pesata



Eigenvector rete non pesata

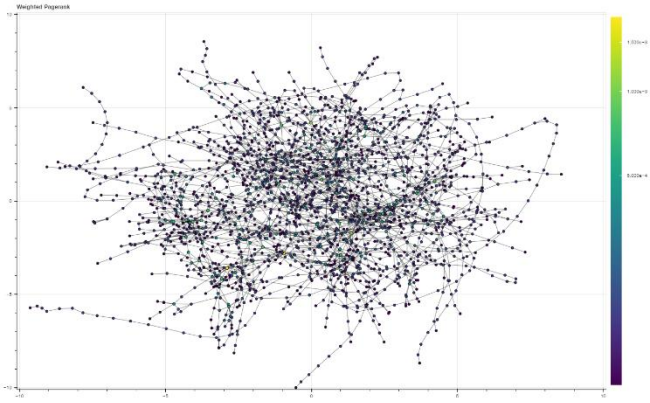


PAGE RANK

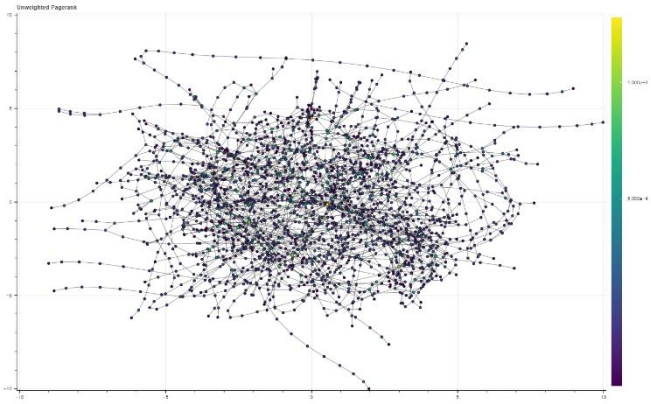


Per il page rank abbiamo al primo posto 32 per la rete pesata che ha un elevato numero di link in ingresso e i nodi che la linkano hanno un alto valore di centralità, e 137 per la non pesata.

Page rank rete pesata



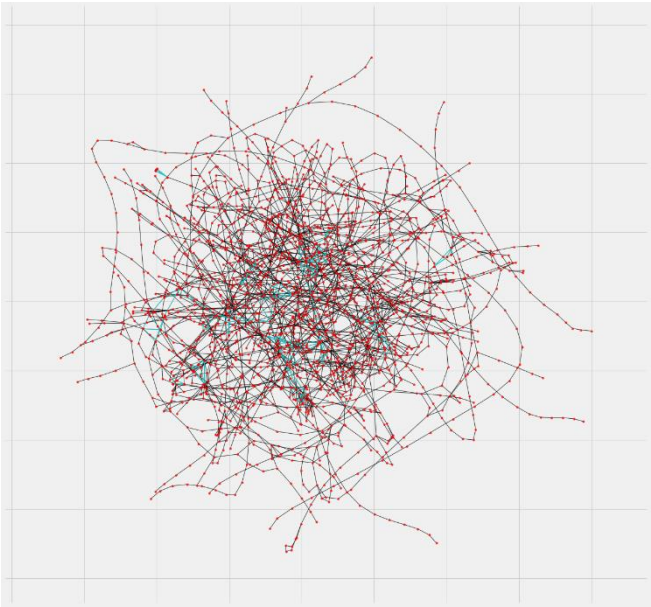
Page rank rete non pesata



D. Medium Scale

Abbiamo cercato di capire se all'interno della rete fossero presenti delle communities attraverso il metodo K-core e facendolo iterare per tutti i valori possibili compresi tra 0 e 100. Per questa rete il valore più alto è 3.

K-core con K=3



Come nello studio della rete precedente l'unico valore disponibile per il k-core è 3 e possiamo vedere dal grafico che il raggruppamento è abbastanza piccolo.

WEAK TIE

I legami deboli sono molto importanti all'interno di una rete poiché collegano diverse comunità.

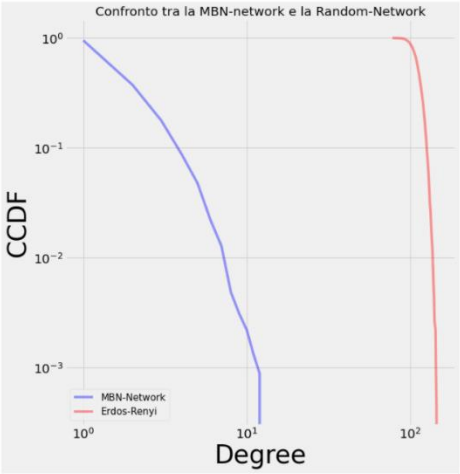
Si può individuare uno span infinito negli edge che vanno da 83 a 841 e da 10 a 825.

```
(1, 2)  1       3
(1, 4)  1       3
(2, 3)  1       3
(4, 3)  1       3
(4, 83) 1       11
(83, 82)        1       11
(83, 841)       1       inf
(7, 8)  1       11
(8, 9)  1       11
(9, 10) 1       11
(10, 11)        1       3
(10, 825)       1       inf
(46, 684)       1       7
(46, 47)        1       4
(46, 685)       1       11
(337, 804)      1       3
(337, 814)      1       3
(124, 802)      1       3
(126, 127)      1       7
(126, 210)      1       3
(13, 47)        1       4
(13, 89)        1       4
(13, 607)       1       5
```

E. Networks models

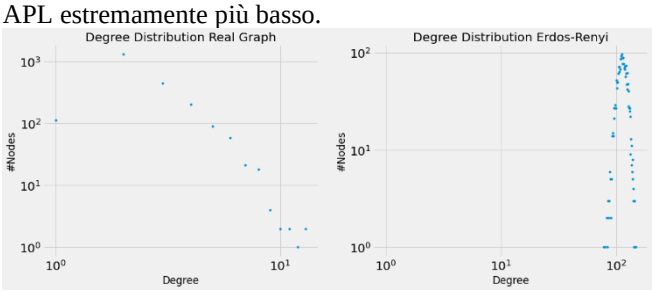
In ultima analisi abbiamo sviluppato 3 diversi modelli per simulare la nostra rete partendo da regole ben definite. Per ognuno di questi si è proceduto ad un confronto grafico della CCDF (distribuzione cumulativa, probabilità che un nodo abbia un certo grado o che sia maggiore di quell grado), ad un confronto su misure a larga scala e a un confronto della degree distribution.

RANDOM NETWORK ERDOS-RENYI



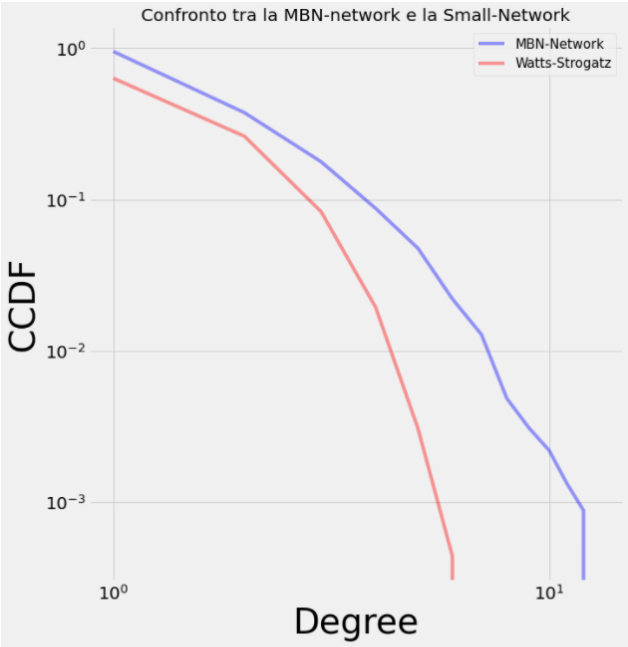

```
Average degree Random Network: 113.27007943512798
Average degree MBN Network: 2.6849073256840246
Il global clustering della network MBN è:  0.188
Il global clustering della random network è:  0.05
L'APL della network MBN è:  103.078
L'APL della random network è:  1.953
```

Come si può notare dal grafico la CCDF si discosta molto dal similuzione della nostra rete. Questo modello simula una rete con average degree eccessivamnete più alto della rete reale e APL estremamente più basso.



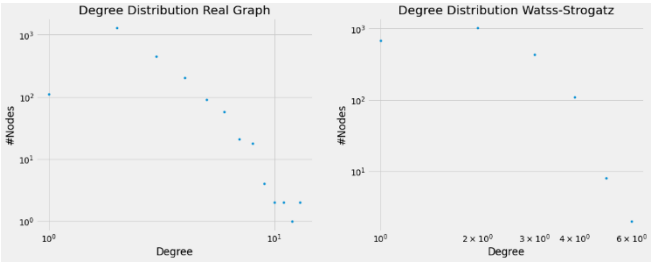
Anche la degree distribution è totalmente diversa. Possiamo concludere che il modello di Erdos-Renyi non è adatto a simulare la nostra rete.

WATTS-STROGATZ



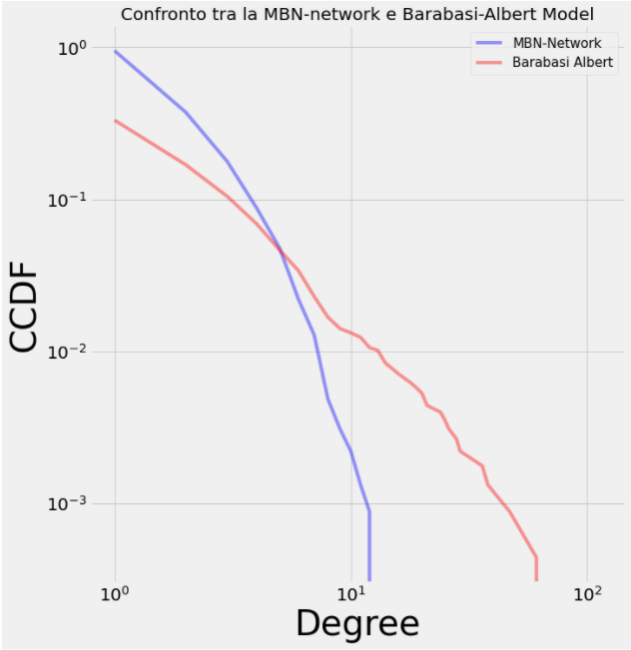

```
Average degree Small word: 2.0
Average degree MBN Network: 2.6849073256840246
Il global clustering della network MBN è:  0.188
Il global clustering della Small word network è:  0
```

In questo caso la CCDF si avvicina e l'average degree è simile alla nostra rete.



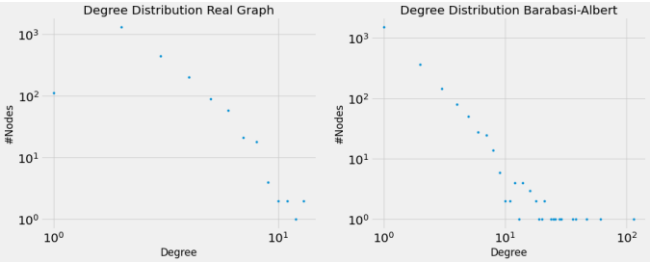
La degree distribution si avvicina alla rappresentazione della nostra rete e anche in questo caso possiamo concludere affermando che questo modello non è totalmente sbagliato per la nostra simulazione.

BARABASI ALBERT



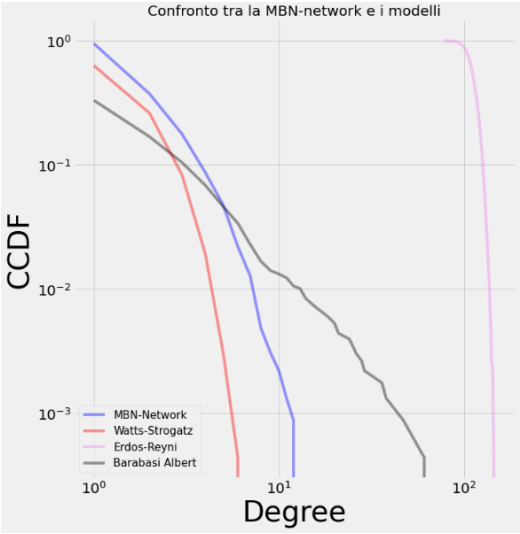

```
Average degree Barabasi-Albert: 1.999117387466902
Average degree MBN Network: 2.6849073256840246
Il global clustering della network MBN è:  0.188
Il global clustering del modello di Barabasi-Albert è:  0
L'APL della network MBN è:  103.078
L'APL del modello di Barabasi- Albert è:  7.815
```

A differenza dei modelli precedenti la CCDF si avvicina molto di più mentre per alcune large scale come, l'average degree il valore è più basso come l'APL.



Per quanto riguarda la degree distribution segue abbastanza l'andamento della nostra rete anche se alcuni nodi presentano un elevato degree rispetto alla rete reale.

Infine riportiamo un grafico di confornto fra tutte le CCDF.



In questo caso i modelli che più si avvicinano alla nostra rete sono Watts-Strogatz e Barabasi.

F. Conclusioni

Mumbai essendo una delle città più popolose e turistiche dell'India, come per le altre città probabilmente giocano un ruolo fondamentale le stazioni che fungono da hubs e che permettono la connessione dei vari nodi.

IX. CONFRONTO FINALE

Una volta terminato il lavoro di analisi per ogni singola rete abbiamo confrontato le misure a larga scala di quest'ultime.

- Average Path length: le network di Ahmedabad, Delhi, Hyderabad e Kolkata hanno un valore decisamente più basso rispetto a Chennai e Mumbai, il che ci indica che le reti autobus delle prime città hanno efficenza di trasporto migliore e più veloce delle seconde.
- Diametro: tutte le reti esclusa quella di Mumbai hanno un diametro compreso tra gli 11 e 27 mentre quest'ultima ha un diametro di 98. Questo ci permette di dire che la rete autobus di Mumbai è sicuramente molto più larga e lunga delle altre città, essendo la città più grande presa in considerazione.
- Densità: riscontrando una densità molto bassa in tutte le città possiamo affermare che la resilienza della network (capacità di ovviare ad errore o attacchi) è bassa, questo è probabilmente dovuto alla conformazione delle rotte degli autobus , cioè due rotte posso avere in comune una o due fermate ma non tutte.
- Clustering: per quanto riguarda questa metrica tutte le reti hanno un valore relativamente basso, indicando che i nodi non tendono a raggrupparsi in cricche.