

Napoli Road Network

Punti di interesse e quotazioni immobiliari

Arianna Nocente

a.nocente@studenti.unipi.it
Student ID: 476249

Jarir Salame Younis

salameyounis@studenti.unipi.it
Student ID: 583084

Marco Cozzolino

m.cozzolino5@studenti.unipi.it
Student ID: 582370

ABSTRACT

In questo studio è stata analizzata la conformazione della rete stradale di Napoli esplorandone grado medio, misure di centralità, resilienza. Sono stati applicati algoritmi di community discovery, con lo scopo di scoprire se essi sono in grado di individuare i quartieri della città ottenendo medi risultati e algoritmi di link prediction, i quali dovrebbero, in questo caso, prevedere la costruzione di nuove strade in base alla struttura della rete ma ottengono scarsi risultati. Ci si chiede se esiste una relazione tra struttura della rete, distribuzione territoriale dei punti di interesse e prezzi delle zone catastali. Mediante un'analisi statica emergono correlazioni poco significative: i prezzi più alti della città non sono dovuti alla presenza di una buona rete stradale o di aree dense di interesse. Tuttavia, esplorando l'andamento delle quotazioni immobiliari emerge che l'introduzione di nuovi punti di interesse causa un aumento dei prezzi, se pur non rendendo i prezzi più alti in assoluto.¹

KEYWORDS

Social Network Analysis , Analisi Reti Sociali, Rete Stradale, Mercato Immobiliare.

ACM Reference Format:

Arianna Nocente, Jarir Salame Younis, and Marco Cozzolino. 2019. Napoli Road Network: Punti di interesse e quotazioni immobiliari. In *Social Network Analysis '19*. , 9 pages.

¹Project Repositories

Data Collection: https://github.com/sna-unipi/data-collection-not_just_nerds

Analytical Tasks: https://github.com/sna-unipi/network-analysis-analytical-tasks-not_just_nerds

Report: https://github.com/sna-unipi/project-report-not_just_nerds

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

SNA, 2018/19, University of Pisa, Italy

© 2019 Association for Computing Machinery.

1 INTRODUZIONE

La rete stradale è stata definita come multigrafo diretto² in cui i link sono le strade e i nodi sono le intersezioni tra esse. L'intera rete è stata distinta per quartieri e zone catastali, alle quali sono stati aggiunti i prezzi di vendita degli immobili e il numero di abitanti. Il grafo è stato ulteriormente caratterizzato aggiungendo punti di interesse: parchi, castelli, scuole, università, ospedali.

2 DATA COLLECTION E DATA MANIPULATION

La raccolta dati è avvenuta mediante OSMnx (Python for street networks) pacchetto di Python che permette di recuperare, modellare, analizzare e visualizzare reti stradali e altri dati spaziali da OpenStreetMap.

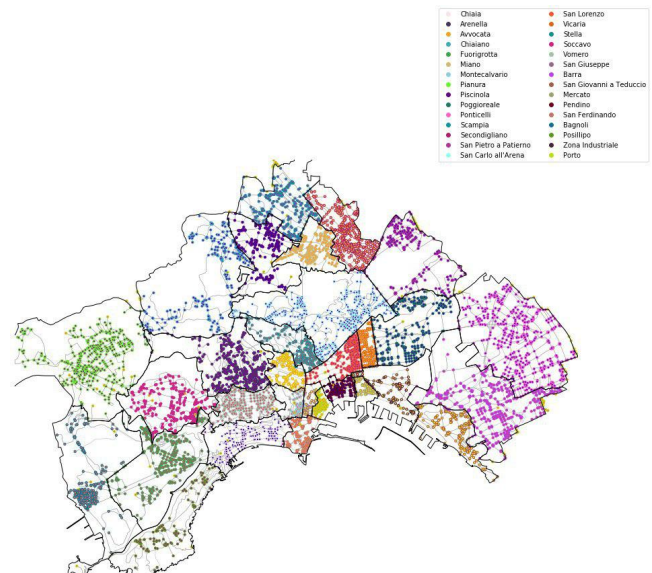


Figure 1: Rete stradale Napoli suddivisa in quartieri con nodi colorati

Quartieri

Mediante la sezione OpenData del sito del comune, sono stati identificati 30 quartieri sotto forma di area. Ad ogni nodo

²Multigrafo

grafo in cui per ogni coppia di nodi esistono più links

della rete è stato assegnato il quartiere di appartenenza ed è stato verificato che non ci fossero nodi in comune tra due quartieri differenti. Per una migliore visualizzazione i nodi sono stati rappresentati sulla mappa con un colore diverso per ciascun quartiere.

Numero di abitanti di ogni quartiere

E' stato considerato il numero di abitanti per ogni quartiere utilizzando come fonte la sezione OpenData del comune di Napoli (anno censimento: 2011).

Punti di interesse

- Ospedali pubblici e cliniche private (definite come punti)
È stato utilizzato un tool per la raccolta dati da pagine web, tale "F-Scraper", da cui è stato scaricato il csv con i nomi degli ospedali pubblici e privati di Napoli e le corrispondenti coordinate geografiche. Gli ospedali sono stati rappresentati sotto forma di punti sulla rete (punti rossi) e sono stati individuati i nodi della rete più vicini (punti verdi). La visualizzazione è resa più intuitiva utilizzando Folium, libreria di Python che genera una mappa interattiva per esplorare il network.

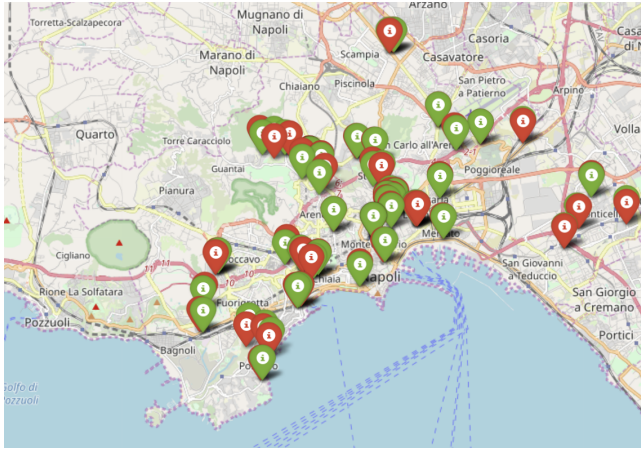


Figure 2: Identificazione ospedali e nodi della rete (rispettivamente rosso e verde)

- Ospedali pubblici (definiti come aree)
Dalla sezione OpenData del sito del comune di Napoli, che tiene conto dei database catastali, sono stati scaricati i dati relativi alle aree degli ospedali pubblici in modo da avere non più un nodo singolo, bensì identificando per ogni area il numero di nodi compresi in quest'ultima. Si ritiene non trascurabile la dimensione dell'area ospedaliera, pertanto per le strutture pubbliche è stato tenuto il dato sotto forma di area e di conseguenza di nodi compresi nell'area. Sono stati

eliminati i duplicati. Sono stati individuati i nodi ambigui, cioè i nodi che appartengono a più di un ospedale e si è risolta la disambiguazione andando ad individuarli nella mappa tramite la libreria Folium sia per osservare da vicino le zone ospedaliere corrispondenti sulla mappa, sia per vedere a quale zona appartenessero i vicini di tale nodo.

- Università e Istituti Superiori
Dalla sezione OpenData del sito del M.i.u.r. sono stati scaricati i dati di università e istituti superiori. In seguito sono stati rappresentati come nodi sulla rete intercettando per ogni punto il nodo della rete più vicino.
- Parchi
Dalla sezione OpenData del sito del comune di Napoli sono stati scaricati i dati riferiti ai parchi pubblici e sono stati rappresentati come nodi sulla rete intercettando per ogni punto il nodo della rete più vicino.
- Castelli
Da OpenData Campania è stato scaricato l'elenco dei castelli e mediante le coordinate sono stati individuati i nodi della rete più vicini.

Zone OMI

Sono state estratte da GEOPOI le aree delle zone OMI ed è stata applicata la stessa procedura dei quartieri.

Numero di abitanti di ogni zona OMI

È stato stimato il numero di abitanti delle zone OMI utilizzando i dati sul numero di abitanti dei quartieri assumendo che la popolazione sia equamente distribuita all'interno dei quartieri utilizzando la seguente formula:

$$N_{zq} = \frac{A_{zq}}{A_q} N_q \quad (1)$$

$$N_z = \sum_q N_{zq} \quad (2)$$

Dove A_{zq} rappresenta l'area di intersezione tra la zona OMI z e il quartiere q , N_{zq} la stima del numero di abitanti di tale area, N_q il numero di abitanti nel quartiere q e N_z la stima del numero di abitanti della zona OMI z .

Quotazioni immobiliari OMI

Per rispondere al quesito riguardante la correlazione tra prezzi delle case, struttura della rete stradale e numero di punti di interesse sono stati utilizzati i dati forniti dall'Agenzia dell'Entrate - Banca dati delle quotazioni immobiliare (OMI), relativi alle quotazioni immobiliari nelle zone catastali di

Napoli divisi per semestre dal 2002 al 2015.

3 NETWORK ANALYSIS

È stata fatta un'analisi della rete di Napoli e si sono confrontati i risultati con le reti random di Erdos-Renyi e Barabasi-Albert.

Erdos-Renyi : probabilità = $\frac{|E|}{|N|(|N|-1)}$
Poiché le reti reale e random hanno lo stesso numero di nodi $|N|$ e lo stesso numero di archi $|E|$, il valore della densità è lo stesso per entrambe.

Invece Barabasi-Albert, parametro $m = \frac{|E|}{|N|}$ quindi la densità può cambiare.

Dalle analisi è emerso che nonostante sembri condividere alcune caratteristiche con le reti random, in particolare con quella di Erdos, la struttura della rete sembra essere differente da entrambe.

Table 1: Statistiche del network originale

Number of nodes:	7790
Number of links:	15820
Pairs of nodes with 2 links (multigraphs):	103
Pairs of nodes with 3 links (multigraphs):	1

Degree Distribution

La degree distribution $p(k)$ rappresenta la probabilità che un nodo scelto casualmente abbia grado k .

$$p(k) = \frac{N_k}{N}$$

dove N_k è il numero di nodi con grado k e N è il numero totale di nodi. Nel caso di una rete random, la degree distribution è ben approssimata dalla distribuzione di Poisson

$$p(k) = e^{-\langle k \rangle} \frac{\langle k \rangle^k}{k!}$$

mentre per una rete scale-free è approssimata dalla power-law $p_k k^{-\gamma}$ dove γ rappresenta il degree exponent della distribuzione.

Table 2: Degree Centrality

	Napoli	Erdos-Renyi	Barabasi
Min	1.000	0.000	2.000
Max	10.000	9.000	213.000
Mean	4.061	2.011	3.998
Stdev	1.549	1.421	6.062

La distribuzione del grado nella rete reale può essere ricondotta alla distribuzione del grado Erdos-Renyi

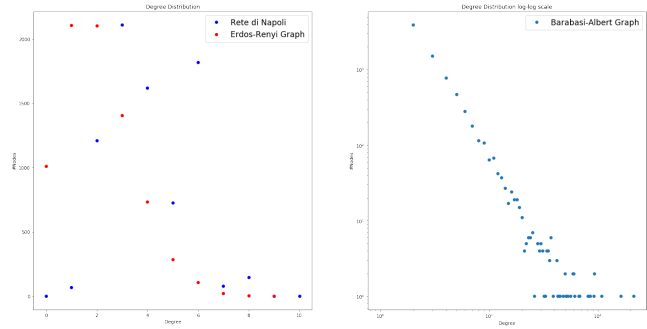


Figure 3: Degree distribution della rete Erdos-Renyi e della rete Barabasi

Connected Components Analysis

L'intero grafo per definizione di rete stradale dovrebbe essere un'unica grande strongly connected component. Infatti un'auto dovrebbe essere in grado di andare da un punto all'altro della città seguendo i corretti sensi di marcia. Tuttavia possiamo individuare nel network una sola weakly connected component e 14 strongly connected component, di cui la più grande è composta da 7596 nodi, la seconda da 24 nodi (centro storico), la terza da 9 nodi, la quarta da 8 nodi. Emerge che ci sono 194 nodi che formano altre componenti connesse e si avanzano alcune ipotesi: sono nodi ai bordi del network, in posizione periferica, vicoli o zone pedonali. Per quanto riguarda i nodi individuati in posizione periferica, se venissero allargati un po' i confini della mappa entro i quali includere la rete stradale, essi si ritroverebbero parte della giant component. Per quanto riguarda gli altri, analizzando la rete pedonale sempre tramite l'utilizzo della libreria OSMnx si scopre che esiste un'unica componente connessa. Pertanto si deduce che questi nodi, per lo più presenti in centro storico, siano in corrispondenza delle zone pedonali.

Path Analysis

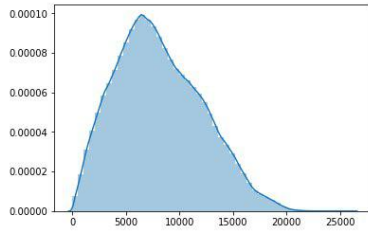
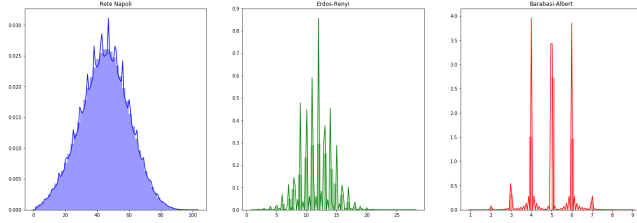
Si definisce path una sequenza di nodi in cui ogni nodo è adiacente al successivo. In particolare lo shortest path è il path minimo tra due nodi.

E' stata calcolata la distribuzione degli shortest path della rete espressa in metri in quanto ritenuta più significativa per il tipo di rete in analisi e si nota che segue una distribuzione normale.

In seguito sono stati calcolati gli shortest path in termini di link al fine di poterli confrontare con la distribuzione degli shortest path della rete Erdos-Renyi e della rete Barabasi. Confrontando i valori degli shortest paths è possibile notare che nella rete di Napoli non è presente la proprietà small

Table 3: Shortest Path

	Napoli	Erdos-Renyi	Barabasi
Min	1.0000	1.000	1.000
Max	102.000	28.000	9.000
Mean	44.647	11.692	4.922
Stdev	15.050	2.651	0.924
$\frac{\log(n)}{\log(\langle k \rangle)}$	6.393	12.818	6.464

**Figure 4: Distribuzione degli shortest path della rete espressa in metri****Figure 5: Shortest path della rete Erdos-Renyi e della rete Barabasi**

world³ che caratterizza le reti random e molte delle reti reali.

Clustering Coefficient

E' stata calcolata l'Average clustering coefficient, che rappresenta il grado medio di clustering della rete:

$$\langle C \rangle = \frac{1}{N} \sum_{i=1}^N C_i \quad (3)$$

dove

$$C_i = \frac{2L_i}{k_i(k_i - 1)} \quad (4)$$

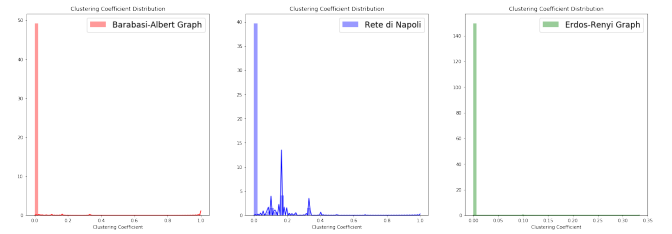
C_i rappresenta il local clustering coefficient e L_i è il numero di archi tra i k_i vicini del nodo i . Si tratta di una misura della densità locale della rete: più i vicini del nodo i sono interconnessi tra loro, più alto è C_i , il cui valore è compreso

³La distanza media tra due nodi cresce logaritmicamente rispetto alle dimensioni del network

tra 0 e 1.

Table 4: Clustering Coefficient

	Napoli	Erdos-Renyi	Barabasi
Min	0.0000	0.0000	0.0000
Max	1.000	0.3333	1.0000
Mean	0.0377	0.0001	0.0068
Stdev	0.088	0.0047	0.0751

**Figure 6: Rappresentazione della clustering coefficient della rete di Napoli confrontata con quella dei grafi di Erdos-Renyi e Barabasi-Albert**

Osservando la tabella si può notare come la variabilità del clustering coefficient sia simile nella rete reale e in quella di Barabasi, anche se dal grafico si nota che nel primo caso i valori sono di gran lunga maggiori e meglio distribuiti.

Density analysis

La Graph Density rappresenta invece il numero di archi presenti nel grafo rispetto a quelli totali possibili.

$$D(G) = \frac{2|E|}{[|N|(|N| - 1)]} \quad \text{undirected} \quad (5)$$

$$D(G) = \frac{|E|}{[|N|(|N| - 1)]} \quad \text{directed} \quad (6)$$

I valori ottenuti, sono:

- Rete reale - density: $2.607 \cdot 10^{-4}$
- Rete random - density: $2.587 \cdot 10^{-4}$
- Rete scale-free - density: $5.134 \cdot 10^{-4}$

La densità del grafo di Erdos-Renyi è leggermente minore rispetto a quella della rete di Napoli a causa della diversa topologia dei due network; questa diversità è infatti dovuta alla presenza di coppie di nodi collegati da più di un link. Per quanto riguarda la scale-free, notiamo che è circa il doppio rispetto alla rete reale.

Centrality measures

Attraverso le seguenti misure si vuol determinare quanto sia importante un nodo all'interno della rete.

Degree Centrality

$$k_i = \sum_{j=1}^N A_{ij} \quad (7)$$

La degree centrality k_i rappresenta il numero dei nodi a cui il nodo i è connesso; il grado del nodo i è ottenuto dagli elementi della matrice di adiacenza.

Per consultare i risultati si rimanda alla sezione degree distribution.

Closeness Centrality

$$C(i) = \frac{N-1}{\sum_{j=1}^{n-1} d_{ij}} \quad (8)$$

La closeness centrality $C(i)$ è il reciproco della somma delle distanze da i a tutti gli altri $N-1$ nodi della rete.

Table 5: Closeness Centrality

	Napoli	Erdos-Renyi	Barabasi
Min	0.000	0.000	0.152
Max	0.036	0.094	0.322
Mean	0.022	0.058	0.205
Stdev	0.004	0.029	0.021

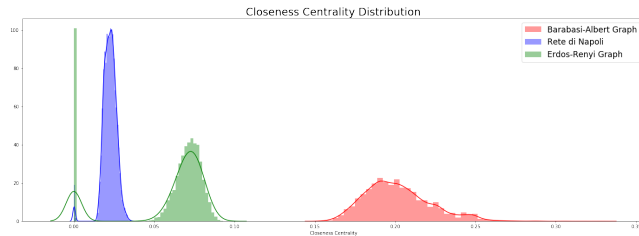


Figure 7: Rappresentazione della closeness della rete di Napoli confrontata con quella dei grafi di Erdos-Renyi e Barabasi-Albert

La distribuzione della closeness centrality della rete reale è simile a quella relativa alla rete di Erdos-Renyi con un andamento normale e un piccolo estremo alla curva in 0. Tuttavia ci sono delle differenze nei valori; oltre alla diversità di altezza dei due picchi in zero è possibile notare che la rete reale, oltre ad avere una media inferiore ha una varianza molto bassa, il che significa che la closeness tra i vari nodi è quasi costante. Per quanto riguarda Barabasi-Albert media e varianza sono maggiori rispetto a Erdos-Renyi e quindi molto lontani dalla rete reale.

Betweenness Centrality

$BC(i)$ =shortest paths through i

La betweenness centrality $BC(i)$ calcola il numero di cammini minimi che attraversano il nodo i .

Table 6: Betweenness Centrality

	Napoli	Erdos-Renyi	Barabasi
Min	0.0000	0.0000	0.0000
Max	0.2063	0.0165	0.1503
Mean	0.0054	0.0008	0.0005
Stdev	0.0142	0.0014	0.0035

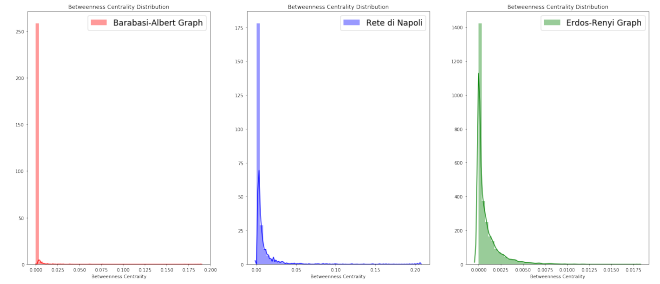


Figure 8: Rappresentazione della betweenness della rete di Napoli confrontata con quella dei grafi di Erdos-Renyi e Barabasi-Albert

Guardando le distribuzioni sembra che la rete reale si comporti in maniera simile a quella di Erdos, tuttavia se si considerano i valori ottenuti nel caso della rete reale si hanno valori di betweenness molto maggiori e con una maggiore variabilità.

Selfloops

Sono stati individuati 23 selfloops e, esplorando la mappa interattiva è emerso che rappresentano la situazione in cui la strada ritorna su se stessa, il che è lecito considerata la topologia del network.

4 ANALYTICAL TASKS

Community discovery

Si applicano differenti algoritmi di community discovery utilizzando la libreria di Python CDlib con lo scopo di verificare se essi sono in grado di individuare i quartieri. Per questa ragione si è deciso di escludere gli algoritmi che producono community di tipo overlap. Sono stati esclusi quelli applicabili solo ai grafi undirected. In prima istanza sono stati considerati gli algoritmi che non richiedono parametri, in

seconda istanza algoritmi i cui parametri sono stati scelti in modo da ottenere un numero di community ragionevole rispetto alle dimensioni della rete e al numero dei quartieri:

- Infomap: 1567 communities di piccoli nodi (16 nodi circa)
- Walktrap: 153 communities
- Significance communities: 1363 communities
I risultati dei tre algoritmi precedenti non evidenziano un numero di community rilevante.
- Spinglass: 25 communities
Spinglass fa affidamento su un'analogia tra un modello meccanico statistico molto popolare chiamato Potts spin glass e la struttura delle communities. Applica la tecnica di ottimizzazione della "ricottura" simulata su questo modello per ottimizzare la modularità, di seguito la definizione:

$$Q(C) = \frac{1}{2m} \sum_{(i,j)} \left(A_{ij} - \frac{k_i k_j}{2m} \right) \delta(c_i, c_j) \quad (9)$$

Q è la modularità, C_i sono le communities, $\frac{k_i k_j}{2m}$ rappresenta la densità attesa, A è la matrice di adiacenza, $\delta(c_i, c_j)$ è il delta della funzione di Kronecker che è 1 se i nodi i e j appartengono alla stessa community e 0 altrimenti.

- Eigenvector: 50 communities
È un algoritmo di tipo Divide Et Impera in cui ogni split viene fatto massimizzando la modularità rispetto al network originale.
- Leiden: 54 communities
L'algoritmo di Leiden è basato sulla modularità, costituito da tre fasi: (1) spostamento locale dei nodi, (2) perfezionamento della partizione (3), aggregazione della rete basata sulla partizione perfezionata utilizzando la partizione non raffinata per creare una partizione iniziale per l'aggregato rete.
- Cpm: 30 communities
Abbastanza bilanciate nel numero di nodi, utilizzando come parametro 0.00018. L'algoritmo verte sulla modularità.

CPM è un modello Leiden in cui la funzione di qualità da ottimizzare è:

$$Q = \sum_{ij} (A_{ij} - \gamma) \delta(\sigma_i, \sigma_j) \quad (10)$$

dove A è la matrice di adiacenza, σ_i denota la community del nodo i , δ è la delta di Kronecker. Determinare un particolare γ corrisponde a scegliere di trovare comunità di una densità particolare.

- Rber pots:
agisce similmente al CPM con il parametro 0.32 trovando 30 communities abbastanza bilanciate per numero di nodi.

$$Q = \sum_{ij} (A_{ij} - \gamma p) \delta(\sigma_i, \sigma_j) \quad (11)$$

dove A è la matrice di adiacenza, σ_i denota la community del nodo i , p è la densità del grafo, δ è la delta di Kronecker. Determinare un particolare γ corrisponde a scegliere di trovare comunità di una densità particolare.

Le communities sono state rappresentate graficamente in una mappa contenente i confini territoriali dei quartieri utilizzando un colore diverso per ogni community in modo da visualizzare come si distribuiscono rispetto ai quartieri.

Per ogni algoritmo si studia la corrispondenza:

- per ogni quartiere qual è la community di maggioranza, in particolare quanti nodi di essa sono in quel quartiere;
- per ogni community qual è il quartiere di maggioranza, in particolare quanti nodi di esso sono in quella community.

È stato deciso di usare come misura di valutazione una misura di accuracy definita nel seguente modo:

$$accuracy = \frac{\sum Q_i + \sum C_j}{2n} \quad (12)$$

dove Q_i è il dato di cui sopra relativo al quartiere i e C_j il dato di cui sopra relativo alla community j .

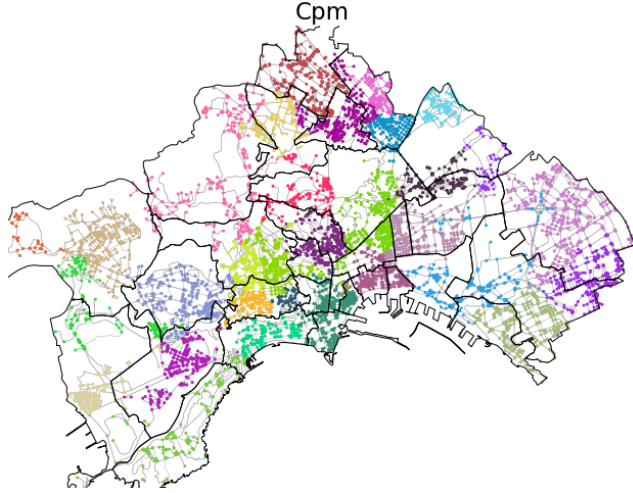
Questi algoritmi non sono deterministici pertanto, eseguendoli più volte, si ottengono diversi risultati. Si effettua quindi un ciclo per poter scegliere tra tutti i risultati quello che mostra la migliore accuracy.

CPM risulta l'algoritmo di community discovery con migliore accuracy in quanto prevede la formazione di 30 communities con accuracy=67%.

Dall'immagine si può osservare come l'algoritmo sia riuscito ad individuare molto bene alcuni quartieri ma non sempre riesce ad essere preciso. Se ne deduce che è possibile che alcuni quartieri non siano ben separati in termini di network.

Table 7: Accuracy dei vari algoritmi

Accuracy of spinglass algorithm:	0.447
Accuracy of eigenvector algorithm:	0.496
Accuracy of leiden algorithm:	0.639
Accuracy of cpm algorithm:	0.675
Accuracy of rber pots algorithm:	0.674

**Figure 9: Funzionamento algoritmo Katz****Link Prediction**

La link prediction è stata effettuata con la libreria linkpred seguendo tale procedimento:

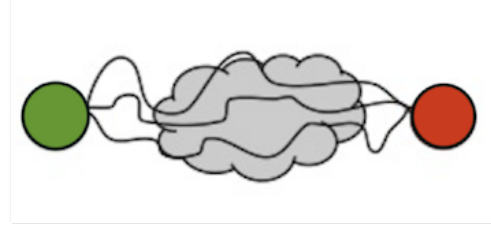
- dal network si rimuovono in maniera casuale 1000 link;
- sul network ridotto appena generato si applicano gli algoritmi unsupervised di link prediction;
- si considerano i 1000 link con score maggiore;
- si confrontano i 1000 link appena predetti con i 1000 link rimossi preliminarmente al fine di valutare la bontà della predizione;
- si confrontano i link rimossi con i link random.

Seguendo questo criterio sono stati applicati i seguenti algoritmi:

- Katz: path based.

$$katz\ score(u, v) = \sum_{i=1}^{\infty} \beta^i |P^i(u, v)| \quad (13)$$

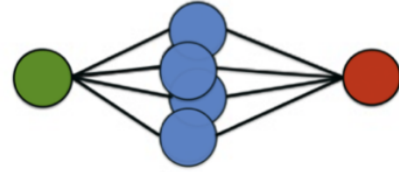
effettua la somma ponderata su tutti i percorsi tra due nodi.

**Figure 10: Funzionamento algoritmo Katz**

- Common Neighbors: neighborhood.

$$common\ neighbors\ score(u, v) = |\Gamma(u) \cap \Gamma(v)| \quad (14)$$

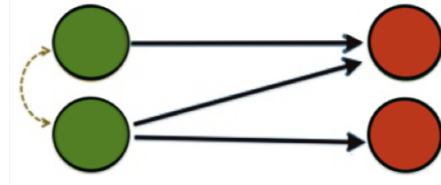
più è alto il numero di vicini che due nodi condividono, più è probabile che entreranno in connessione.

**Figure 11: Funzionamento algoritmo Common Neighbors**

- Jaccard: neighborhood.

$$jaccard\ score(u, v) = \frac{|\Gamma(u) \cap \Gamma(v)|}{|\Gamma(u) \cup \Gamma(v)|} \quad (15)$$

più è simile la cerchia di vicini di due nodi, più probabilmente entreranno in connessione.

**Figure 12: Funzionamento algoritmo Jaccard**

- Sim Rank: Similarity

L'algoritmo è andato in memory error pertanto non è stato possibile esplorare i risultati.

La procedura è stata ripetuta 50 volte ed è stata calcolata la media dei risultati per ciascun algoritmo. Nessun algoritmo ha prodotto buoni risultati.

Sembra che essi performino come la predizione random.

Si riportano di seguito le medie delle percentuali di link correttamente prodotti:

- Katz: 0.012%
- Common Neighbors: 0.000%
- Jaccard: 0.000%
- Random: 0.004%

E' possibile concludere che tali tecniche di link prediction, in quanto unsupervised, non riescono a prevedere la formazione di nuovi link.

Network Resilience

L'analisi della resilienza della rete è avvenuta inizialmente rimuovendo i link in ordine crescente e poi in ordine decrescente di Edge Betweenness. Dopo ogni rimozione è stata ricalcolata la dimensione della Giant component. Lo stesso procedimento è stato fatto rimuovendo i link in maniera casuale.

Edge Betweenness

Rappresenta il numero di shortest path che attraversano il link.



Figure 13: Giant Component Betweenness, sull'asse x gli edges rimossi, sull'asse y la frazione della Giant Component

Rimuovendo gli edges in ordine decrescente di betweenness si osserva che dopo aver eliminato fino a 500 edges la dimensione della Giant component rimane quasi invariata. Da 1000 edges in poi la dimensione della Giant component diminuisce fortemente. La rimozione mediante edge betweenness in ordine decrescente simula attacchi mirati alla rete.

Random

Si rimuovono i link in maniera casuale simulando il random failure, nel nostro caso specifico si potrebbe associare la rimozione di un link alla chiusura di una strada per lavori in corso.

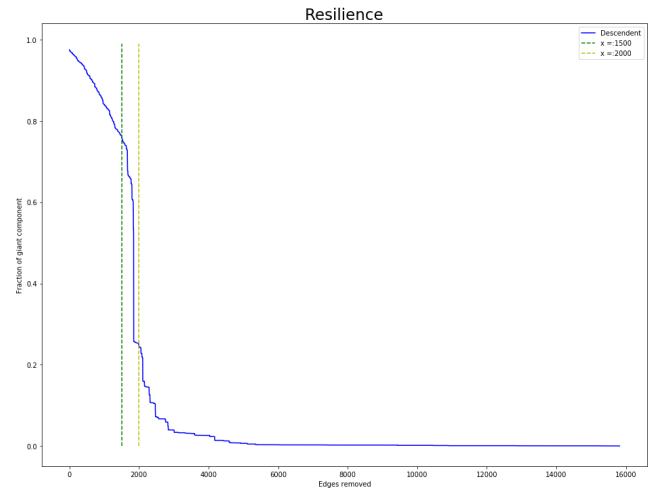


Figure 14: Giant Component Random, sull'asse x gli edges rimossi, sull'asse y la frazione della Giant Component

Rimuovendo gli edges in maniera casuale si osserva che dopo aver eliminato fino a 1500 edges la dimensione della Giant component rimane quasi invariata. Da 2000 edges in poi la dimensione della Giant component diminuisce fortemente.

5 OPEN PROBLEM

Esiste una relazione tra punti di interesse, quotazioni immobiliari e struttura della rete?

Per rispondere a questa domanda è stata fatta dapprima un'analisi statica e poi un'analisi dinamica ponendo l'attenzione alle zone catastali in cui sono stati costruiti due nuovi punti di interesse.

Analisi statica

E' stata analizzata la struttura della rete stradale di ogni zona catastale: grado medio, lunghezza media delle strade, numero di punti di interesse, betweenness centrality, closeness centrality e clustering coefficient. Si vuole indagare se i prezzi di vendita delle case siano correlati alla struttura della rete stradale e ai punti di interesse: scuole, università, castelli, ospedali e parchi. Si è ritenuto necessario rapportare i dati relativi ai punti di interesse e il numero di link al numero di abitanti di ciascuna zona catastale per tenere conto della popolazione. È per questo che sono state introdotte le nuove variabili:

$$ratio_edges = \frac{\text{numero di link}}{\text{numero di abitanti}}$$

$$ratio_nodi_interesse = \frac{\text{numero di nodi di interesse}}{\text{numero di abitanti}}$$

Qui di seguito è riportata la matrice di correlazione tra le grandezze prese in analisi:

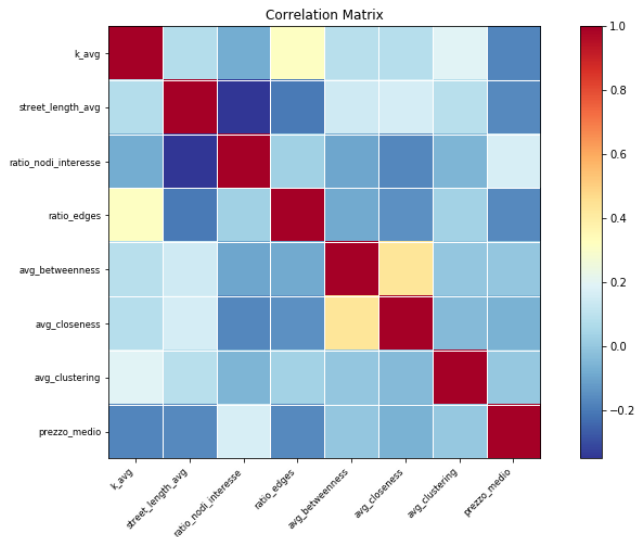


Figure 15: Tabella di correlazione Kendall

Osservando l'ultima riga della correlation matrix emerge che tutte le misure non sono correlate ai prezzi di vendita (dato relativo al 2015). Sono stati utilizzati gli indici di correlazione di Kendall e Pearson e per entrambi si ottengono valori molto vicini allo zero.

Da ciò emerge che i prezzi di vendita più alti non si hanno né nelle zone ben servite dalla rete stradale, né in quelle ricche di punti di interesse.

Si conclude deducendo che evidentemente ci sono altre variabili più rilevanti che determinano il prezzo degli immobili: pregio storico, turistico. Si ritiene quindi necessaria un'ulteriore analisi, in questo caso dinamica.

Analisi dinamica

È stata effettuata un'analisi dinamica delle quotazioni immobiliari distinte per zone catastali relative al periodo che va dal 2002 al 2013. I dati sono stati rappresentati sotto forma di time series.

Nella maggior parte delle time series si ha un trend positivo con media 19.24. Facendo un passo indietro e studiando l'evoluzione della città nel tempo si scopre che a Napoli sono stati costruiti due grandi poli di interesse negli ultimi quindici anni. Il primo è l'Ospedale del Mare, redatto secondo i riferimenti del progetto di un modello di ospedale firmato da Renzo Piano sviluppato nel 2001 per il Ministero della Salute.

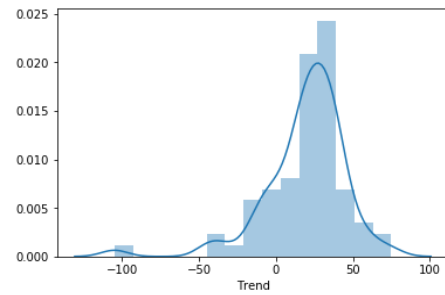


Figure 16: Distribuzione del trend

I lavori sono iniziati nel 2006 e nel 2016 l'ospedale è stato inaugurato. Il secondo grande punto di interesse di recente costruzione è il polo universitario di Napoli Est, progetto di riqualificazione, ex stabilimento Cirio. I lavori di costruzione dell'Università sono iniziati nel 2003, intensificati 10 anni dopo e conclusi ufficialmente con l'inaugurazione del 2015. Analizzando il trend relativo alle zone catastali E08 (zona in cui si trova l'Ospedale del mare) e alla zona catastale E17 (zona della nuova università), emergono i trend:

Zona E08: 35.84

Zone E17: 41.24

I quali risultano circa il doppio rispetto alla media (19.24).

Ciò conferma che una riqualificazione urbana porta ad un forte incremento dei prezzi, se pur non in termini di prezzi più alti in assoluto, causati invece da altri fattori come emerso dall'analisi statica e avvalorato dal seguente grafico.

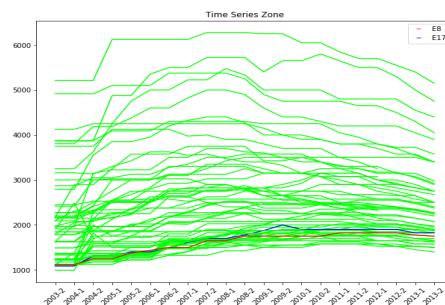


Figure 17: Time series, sull'asse x i semestri, sull'asse y le quotazioni immobiliari

La riqualificazione e quindi l'aggiunta di importanti punti di interesse, causa l'aumento dei prezzi di vendita degli immobili rispetto al trend medio di crescita dei prezzi ma il prezzo rimane comunque inferiore ai prezzi più alti in assoluto. Le quotazioni immobiliari risultano invece indipendenti dalle caratteristiche strutturali del network.