Heidelberg University
Institute of Computer Science
Database Systems Research Group

Lecture: Complex Network Analysis

Prof. Dr. Michael Gertz

Assignment 8

Degree Correlations and Assortativity

`https://github.com/nilskre/CNA_assignments`

Team Member: Patrick Günther, 3660886,
Applied Computer Science
rh269@stud.uni-heidelberg.de

Team Member: Felix Hausberger, 3661293,
Applied Computer Science
eb260@stud.uni-heidelberg.de

Team Member: Nils Krehl, 3664130,
Applied Computer Science
pu268@stud.uni-heidelberg.de

Problem 6-3 Degree Correlation Coefficient

December 13, 2021

1 Lecture: Complex Network Analysis

Prof. Dr. Michael Gertz

Winter Semester 2021/22

1.1 Assignment 6 - Degree Correlations and Assortativity

Students: Felix Hausberger, Nils Krehl, Patrick Günther

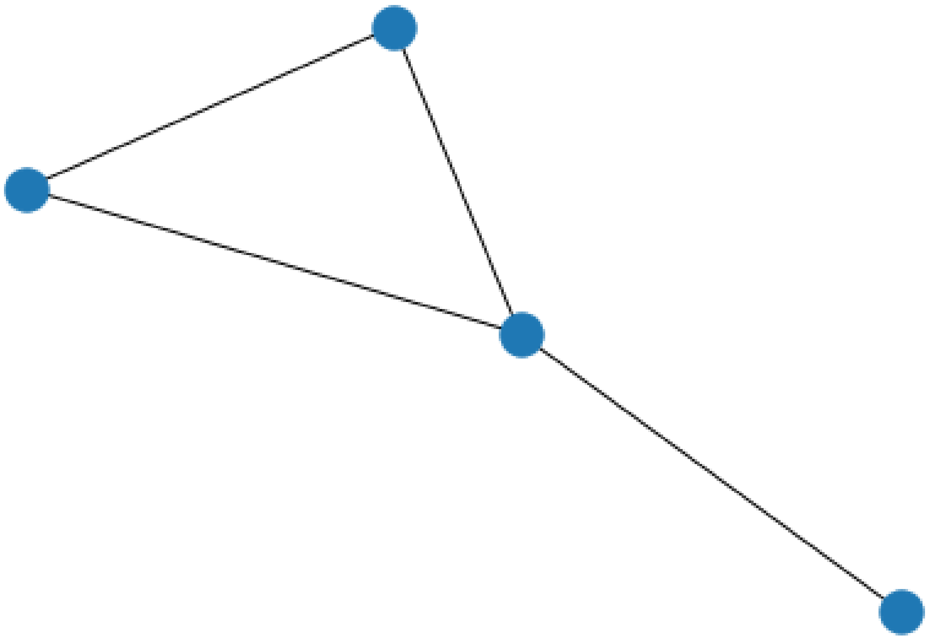
2 1. Compute the degree correlation matrix

[1]:
```
import numpy as np
import networkx as nx
```

[2]:
```
A = np.matrix([[0, 1, 0, 1],
               [1, 0, 1, 1],
               [0, 1, 0, 0],
               [1, 1, 0, 0]])
```

[3]:
```
G = nx.convert_matrix.from_numpy_matrix(A)
```

[4]:
```
nx.draw(G)
```



[5]:
```
G.degree
```

[5]:
```
DegreeView({0: 2, 1: 3, 2: 1, 3: 2})
```

[6]:
```
max_degree = max(deg for n, deg in G.degree)
mapping = {x: x for x in range(max_degree+1)}
deg_corr_mat = nx.degree_mixing_matrix(G, mapping=mapping)
```

[7]:
```
deg_corr_mat
```

[7]:
```
array([[0.   , 0.   , 0.   , 0.   ],
       [0.   , 0.   , 0.   , 0.125],
       [0.   , 0.   , 0.25 , 0.25 ],
       [0.   , 0.125, 0.25 , 0.   ]])
```

(the first column/ row is for degree 0... that could be cut out, since a node with degree 0 never connects to any other node)

3 2. Compute the probabilities q_k of having a degree k node at the end of a random link

[8]:
```
avg_degree = sum(deg for n, deg in G.degree)/len(G.degree)

q_k = {}
for deg in range(max_degree + 1):
    p_k = [deg for n, deg in G.degree].count(deg)/len(G.degree)
    q_k[deg] = (deg * p_k)/avg_degree
```

[9]:
```
q_k
```

[9]:
```
{0: 0.0, 1: 0.125, 2: 0.5, 3: 0.375}
```

4 3. Compute the degree correlation coefficient r

[10]:
```
sigma_squared = sum([(k**2) * q_k[k] for k in q_k]) - sum([k * q_k[k] for k in␣
↪q_k])**2

r = []

for j, row in enumerate(deg_corr_mat):
    for k, e_jk in enumerate(row):
        qk = q_k[k]
        qj = q_k[j]
        r.append((j*k*(e_jk-qj*qk))/sigma_squared)

r = sum(r)
```

[11]:
```
print(f"The degree correlation coefficient of the network is {r}.")
```

```
The degree correlation coefficient of the network is -0.7142857142857144.
```

[12]:
```
# to check our computation, we also use the inbuild function of networkx
nx.algorithms.assortativity.degree_assortativity_coefficient(G)
```

[12]:
```
-0.7142857142857143
```

[ ]:
```

```

1. Compute the modularity of the graph.

   - $L = 15$
   - $n = 10$
   - $n_c = 3$
   - $L_A = 5; L_B = 2; L_C = 3$
   - $k_A = 3 + 3 + 3 + 4 = 13; k_B = 2 + 4 + 3 = 9; k_C = 3 + 2 + 3 = 8$

   Plugging this values into the given equation, we get $M(G, C) = 0.318$.

2. Give proof that for every partitioning of every simple graph, it always holds that $M \leq 1$.

   For each component, its links within itself can be at most $L_c = L$. In this case, the component would hold all the links of the graph. For this component $\frac{L_c}{L}$ would be 1. Since there is the term $\frac{k_c}{2L}^2$ subtracted from that, the summand for this component in the equation cannot be greater than 1. All other components remaining in the graph cannot have any links, which means that these are just unconnnected single nodes. This means that their summand in the equation (ignorring the division by zero) would be 0. The whole equation thus cannot be greater than 1.

   For other cases where there are multiple components containing links, the total number of L_c for all components is still limited by L. Thus it has to hold that $\sum_{c=1}^{n_c}(\frac{L_c}{L}) \leq 1$. Since the other term in each summand $(\frac{k_c}{2L}^2)$ can only decrease the overall sum, $M \leq 1$ has to hold.

   For the trivial case of all nodes being disconnected, $M = 0$ (ignoring the division by 0).