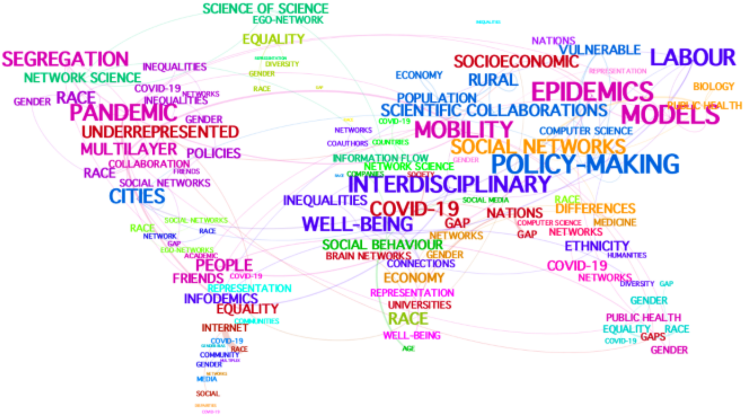
Comparison of network complexity measures

Yipei Zhao

Aston University

October 31, 2021

Network Science



https://appliednetsci.springeropen.com/networked-inequality--studies-on-diversity-and-marginalization

Complexity measures

- Different subgraph measures
  - $C_{1e,st}$; counting the number of different subgraphs with different number of spanning trees after deleting one edge.
  - $C_{1e,spec}$; counting the number of different subgraphs with different spectrums after deleting one edge.
  - $C_{2e,spec}$; counting the number of different subgraphs with different spectrums after deleting two edges.
- Product measures
  - MA_g; product of redundancy and mutual information.
  - MA_{RI}; product of redundancy and mutual information using a different normalisation method than MA_g.
  - Cr; largest eigenvalue of adjacency matrix.
  - Ce; efficiency of the graph.
- Entropy measure
  - OdC; calculating the entropy of node-node link correlation matrix.

MA_{RI}

A product measure that is based on the idea of MA_g.

- Redundancy of a graph: $R = \frac{1}{m} \sum_{i,j>i} ln(d_i d_j)$
- Mutual information of a graph: $I = \frac{1}{m} \sum_{i,j>i} ln(\frac{2m}{d_i d_j})$
- An alternative way to state the mutual information: $I = ln(2m) - R$
- Highest redundancy: $R_{clique} = 2ln(n-1)$
- Lowest redundancy: $R_{path} = 2(\frac{n-2}{n-1})ln(2)$
- Highest mutual information: $I_{path} = ln(n-1) - (\frac{n-3}{n-1})ln2$
- Lowest mutual information: $I_{clique} = ln(\frac{n}{n-1})$

We can define the complexity to be $C = (R - R_{path})(I - I_{clique})$.

$$MA_g = 16(\frac{R-R_{path}}{R_{clique}-R_{path}})(1 - \frac{R-R_{path}}{R_{clique}-R_{path}})(\frac{I-I_{clique}}{I_{path}-I_{clique}})(1 - \frac{I-I_{clique}}{I_{path}-I_{clique}})$$

MA_{RI} continue

To compare different complexity measures, they need to be normalised: $0 < C < 1$.
The complexity measure can be rewritten as:
$C = (R - R_{path})(ln(2m) - R - I_{clique})$.

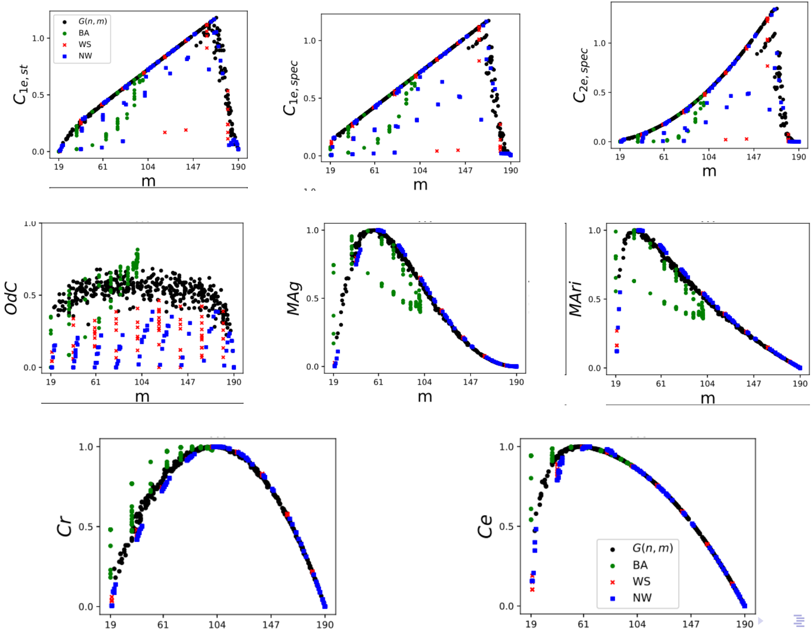
$$C = -R^2 + (ln(2m) - I_{clique} + R_{path})R + (-R_{path} ln(2m) + R_{path} I_{clique})$$

$$R_{max} = \frac{ln(2m) - I_{clique} + R_{path}}{2}$$

$$C_{max} = \frac{(ln(2m) - I_{clique} - R_{path})^2}{4}$$

$$MA_{RI} = \frac{4(R - R_{path})(I - I_{clique})}{(ln(2m) - I_{clique} - R_{path})^2}$$

Result



G(n, m)
BA
WS
NW
C1e, st
C1e, spec
C2e, spec
OdC
MAg
MAri
Cr
Ce
m
0.0
0.5
1.0
19
61
104
147
190

Result continue

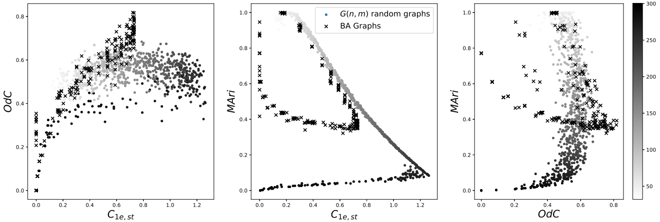


Figure: Correlation between complexity measures, all graphs have 25 nodes and random number of edges. The darker the data point, the graph has more number of nodes.

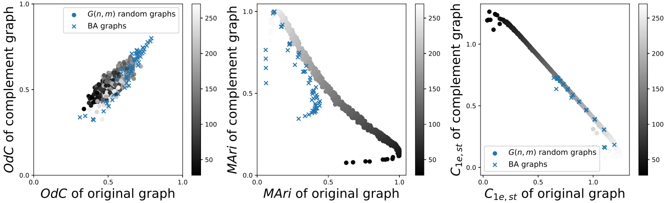
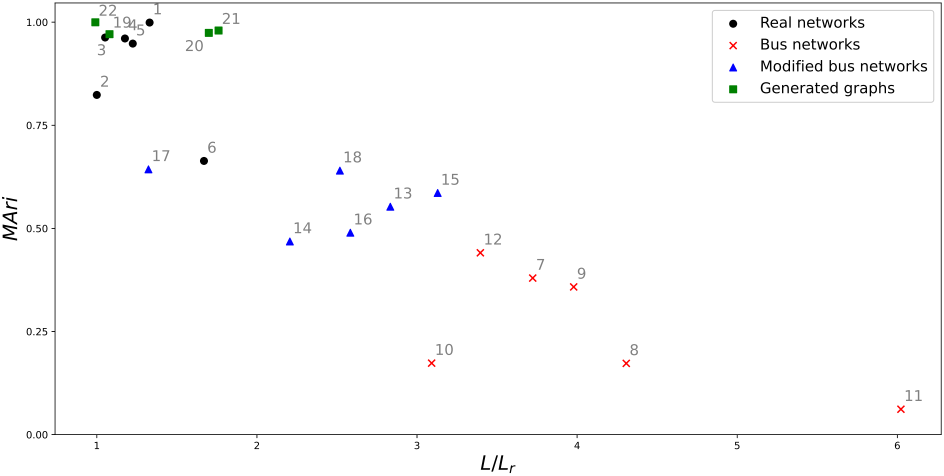


Figure: Complexities of the original graphs and complement graphs with $n = 20$.

Result continue



MA_{RI} complexity of real networks, bus networks, modified bus networks and graphs generated by graph models.

Conclusion

- Compared different complexity measures
- Introduced MA_{RI}
- Compared complexity measures on different types of graph
- Investigated the uniqueness of transportation networks
- How to invent an optimal complexity measure?
- Do transportation networks require different complexity measures compare to other real networks?
- Should different type of networks use different measure?

Question time

Any question for me?