Evaluating Structural Connectivity in Multiple Networks

Michele Coscia

Harvard University

NetSci - June 3rd, 2013
Structural Connectivity

Degree-related measures are very useful in complex networks.

Does it change significantly in multiple networks?

Do we add a new degree of freedom for new degree-related measures?
My Degree: 9
My Neighbors: 4

... Neighbors in at least 3 networks: 1

... Neighbors in the "Friendship" network: 2

...
Question #1

How many neighbors will I lose if we remove one network?
If “Financial” goes down, we lose one neighbor.

If “Work” goes down, we lose three edges but no neighbors.

“Financial” looks very important, even if it had only that one edge.
Formally:

$$\text{Neighbors}_{\text{XOR}}(v, D) = \left| \{ u \in V \mid \exists d \in D : (u, v, d) \in E \land \nexists d' \notin D : (u, v, d') \in E \} \right|$$

Counting the edges appearing in that network

Discarding the ones also appearing in other networks
Question #2

What is the most relevant network for a node (given different relevance criteria)?
Criterion #1

The ratio of neighbors connected in that network

We call it Dimension Relevance:

\[
\text{DimRelevance}(v, D) = \frac{\text{Neighbors}(v, D)}{\text{Neighbors}(v, L)}
\]

\[
\begin{align*}
\text{DimRelevance}(\text{Me, Financial}) &= 3 / 4 \\
\text{DimRelevance}(\text{Me, Work}) &= 3 / 4 \\
\text{DimRelevance}(\text{Me, Friendship}) &= 2 / 4 \\
\text{DimRelevance}(\text{Me, Sport}) &= 1 / 4
\end{align*}
\]
Criterion #2

Weight the # of neighbors on the number of other networks in which I can connect to them

\[ \text{DimRelevance}_W(v, D) = \frac{\sum_{u \in \text{NeighborSet}(v, D)} \frac{n_{uvd}}{n_{uv}}}{\text{Neighbors}(v, L)} \]

- DimRelevance(Me, Financial) = 1.75 / 4
- DimRelevanceW(Me, Work) = 1.25 / 4
- DimRelevanceW(Me, Friendship) = 0.75 / 4
- DimRelevanceW(Me, Sport) = 0.25 / 4
Criterion #3

The ratio of neighbors lost if the network disappears

We call it Dimension Relevance XOR:

\[ \text{DimRelevance}_{XOR}(v, D) = \frac{\text{Neighbors}_{XOR}(v, D)}{\text{Neighbors}(v, L)} \]

- DimRelevance(Me, Financial) = 1 / 4
- DimRelevanceW(Me, Work) = 0 / 4
- DimRelevanceW(Me, Friendship) = 0 / 4
- DimRelevanceW(Me, Sport) = 0 / 4
A Nice Property

$$\text{DimRelevance} \geq \text{DimRelevanceW} \geq \text{DimRelevanceXOR}$$
Why do we care?

Query contextualization
- Term-Term network from AOL query logs
  Networks: Rank of the URL clicked by the user

A-Social User Behavior
- User-User network from Flickr
  Networks: Friendship, Comment, Favorite, Tag

Science Jumpers
- Co-authorship network from DBLP
  Networks: Year of the collaboration
Generally bad results

In Flickr:
Users appearing only in the “Friendship” networks

In DBLP:
Authors changing their neighborhood from one year to the other
Capturing network interplay

- DR Distributions in the original network
- MultidimRandom (Only the # of Nodes/Edges are preserved)
- Multidim Preferential Attachment (rich-get-richer with same # of Nodes/Edges)
- Multidim Jaccard (correlations between dimensions preserved)
Question #3

Are there networks that show the very same patterns of connections?
Network Correlation

The Jaccard coefficient over the nodes with at least one edge in the network; or over the edges themselves

\[
\rho_{node}(d_1, d_2) = \frac{|V_{d_1} \cap V_{d_2}|}{|V_{d_1} \cup V_{d_2}|}
\]

\[
\rho_{edge}(d_1, d_2) = \frac{|E_{d_1} \cap E_{d_2}|}{|E_{d_1} \cup E_{d_2}|}
\]

The Jaccard coefficient over the nodes with at least one edge in the network; or over the edges themselves
Application: Network Eras

Nodes: Authors; Edges: Coauthorship; Each Network: a year
For each consecutive years we calculate the network correlation
Application: Network Eras

In the same era we expect a regular evolution.
The difference between the expectation and reality is how we detect new eras.
Conclusion

Network connectivity is much different in multiple networks

We have an additional degree of freedom

Advanced analysis can be performed even with measures as simple as the degree
Bibliography


M. Schich and M. Coscia: Untangling the Complex Overlap of Subject Themes in Classical Archaeology. MLG @ KDD, (2011)
Thank you

Questions?