Structural measures for multiplex networks

LASAGNE
Cutting through complexity

Federico Battiston, Vincenzo Nicosia and Vito Latora

School of Mathematical Sciences, Queen Mary University of London

Mathematics of Networks 2014 - Imperial College London, Sept. 10, 2014

SEVENTH FRAMEWORK PROGRAMME

Outline

- General formalism for multiplex networks
- Measures to characterise the multiplexity of a system
  1. Basic node and link properties
  2. Local properties (clustering)
  3. Global properties (transitivity, reachability, centrality)
- Validation of all measures on a genuine multi-layer dataset of Indonesian terrorists

General formalism for multiplex networks

A multiplex is a system whose basic units are connected through a variety of different relationships. Links of different kind are embedded in different layers.

- Node index $i = 1, \ldots, N$
- Layer index $\alpha = 1, \ldots, M$

General formalism for multiplex networks

A multiplex is a system whose basic units are connected through a variety of different relationships. Links of different kind are embedded in different layers.

- Node index $i = 1, \ldots, N$
- Layer index $\alpha = 1, \ldots, M$

For each layer α:

- adjacency matrix $A^{[\alpha]} = \{a_{ij}^{[\alpha]}\}$
- node degree $k_i^{[\alpha]} = \sum_j a_{ij}^{[\alpha]}$
- $\sum_i k_i^{[\alpha]} = 2K^{[\alpha]}$ $K^{[\alpha]}$ is the size of layer α

General formalism for multiplex networks

A multiplex is a system whose basic units are connected through a variety of different relationships. Links of different kind are embedded in different layers.

- Node index $i = 1, \ldots, N$
- Layer index $\alpha = 1, \ldots, M$

For each layer α:

- adjacency matrix $A^{[\alpha]} = \{a_{ij}^{[\alpha]}\}$
- node degree $k_i^{[\alpha]} = \sum_j a_{ij}^{[\alpha]}$
- $\sum_i k_i^{[\alpha]} = 2K^{[\alpha]}$ $K^{[\alpha]}$ is the size of layer α

For the multiplex:

- vector of adjacency matrices $\mathbf{A} = \{A^{[1]}, ..., A^{[M]}\}$.
- vector of degrees $\mathbf{k}_i = (k_i^{[1]}, ..., k_i^{[M]})$.

Vectorial variables are necessary to store all the richness of multiplexes.

General formalism: aggregated matrices

Aggregated topological network:

- adjacency matrix $\mathcal{A} = \{a_{ij}\}$:

$$a_{ij} = \begin{cases} 1 & \text{if } \exists\, \alpha : a_{ij}^{[\alpha]} = 1 \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

- node degree $k_i = \sum_j a_{ij}$
- $\sum_i k_i = 2K$

General formalism: aggregated matrices

Aggregated topological network:

- adjacency matrix $\mathcal{A} = \{a_{ij}\}$:

$$a_{ij} = \begin{cases} 1 & \text{if } \exists\, \alpha : a_{ij}^{[\alpha]} = 1 \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

- node degree $k_i = \sum_j a_{ij}$
- $\sum_i k_i = 2K$

Aggregated overlapping network:

- adjacency matrix $\mathcal{O} = \{o_{ij}\}$:

$$o_{ij} = \sum_{\alpha} a_{ij}^{[\alpha]} \quad \text{edge overlap} \tag{2}$$

- node degree $o_i = \sum_j o_{ij} = \sum_\alpha k_i^{[\alpha]}, \quad o_i \geq k_i$
- $\sum_i o_i = 2O$

General formalism: aggregated matrices

Aggregated topological network:

- adjacency matrix $\mathcal{A} = \{a_{ij}\}$:

$$a_{ij} = \begin{cases} 1 & \text{if } \exists\, \alpha : a_{ij}^{[\alpha]} = 1 \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

- node degree $k_i = \sum_j a_{ij}$
- $\sum_i k_i = 2K$

Aggregated overlapping network:

- adjacency matrix $\mathcal{O} = \{o_{ij}\}$:

$$o_{ij} = \sum_\alpha a_{ij}^{[\alpha]} \quad \text{edge overlap} \tag{2}$$

- node degree $o_i = \sum_j o_{ij} = \sum_\alpha k_i^{[\alpha]}, \quad o_i \geq k_i$
- $\sum_i o_i = 2O$

Scalar variables describing system's multiplexity cannot disregard the layer index $[\alpha]$.
Generalisation to the case of weighted layers: $a_{ij}^{[\alpha]} \to w_{ij}^{[\alpha]} \quad k_i^{[\alpha]} \to s_i^{[\alpha]} \quad o_{ij} \to o_{ij}^{\mathrm{w}}$

The multi-layer network of Indonesian terrorists

- 78 nodes
- 911 edges representing 4 social relationships:
  1. Trust (weighted edges)
  2. Operations (weighted edges)
  3. Communications
  4. Businness (only few information)

The multi-layer network of Indonesian terrorists

LAYER	CODE	N_{act}	K	S	O	O^{w}
MULTIPLEX	M	78	623	/	911	1014
Trust	T	70	259	293	/	/
Operations	O	68	437	506	/	/
Communications	C	74	200	200	/	/
Business	B	13	15	15	/	/

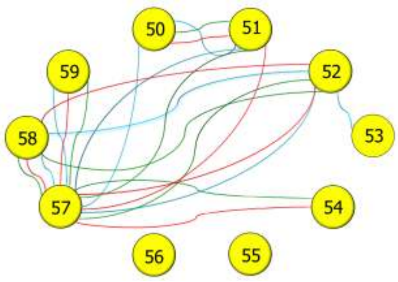
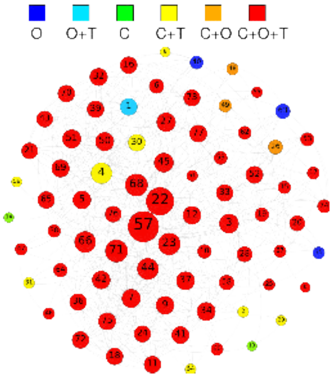


Figure : Coloured-edge representation of a subset of 10 nodes for the multiplex network of Indonesian terrorist: green edges represent trust, red edges communications and blue edges common operations.

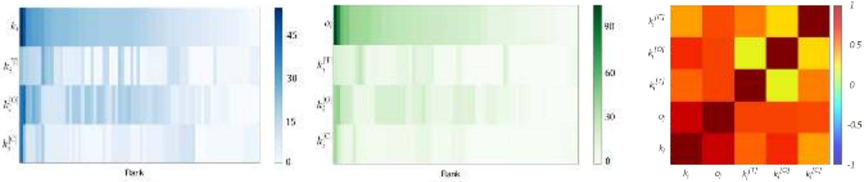
The multi-layer network of Indonesian terrorists

LAYER	CODE	N_{act}	K	S	O	O^{w}
MULTIPLEX	M	78	623	/	911	1014
Trust	T	70	259	293	/	/
Operations	O	68	437	506	/	/
Communications	C	74	200	200	/	/
Businness	B	13	15	15	/	/



Basic node properties

A layer-by-layer exploration of node properties: the case of the degree distribution.



(a) (b) (c)

Different layers show different patterns.

Basic node properties: cartography of a multiplex

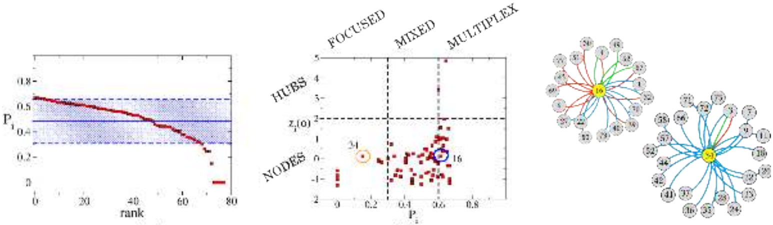
Participation coefficient: $P_i = 1 - \sum_{\alpha=1}^{M} \left(\frac{k_i^{[\alpha]}}{o_i} \right)^2$

1. Focused nodes $0 \leq P_i \leq 0.3$
2. Mixed-pattern nodes $0.3 < P_i \leq 0.6$
3. Truly multiplex nodes $P_i > 0.6$

Z-score of the overlapping degree: $z_i(o) = \frac{o_i - <o>}{\sigma_o}$

1. Simple nodes $-2 \leq z_i(o) \leq 2$
2. Hubs $z_i(o) > 2$

Basic node properties: cartography of a multiplex



(a) (b) (c)

Edge overlap and social reinforcement

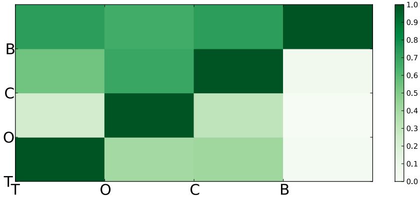
o_{ij}	Percentage of edges (%)
1	46
2	27
3	23
4	4

Edge overlap and social reinforcement

o_{ij}	Percentage of edges (%)
1	46
2	27
3	23
4	4

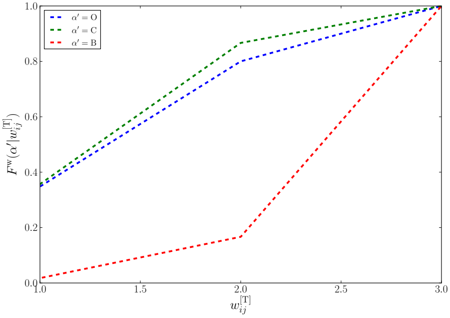
Probability conditional overlap:

$$F(a_{ij}^{[\alpha']}|a_{ij}^{[\alpha]}) = \frac{\sum_{ij} a_{ij}^{[\alpha']} a_{ij}^{[\alpha]}}{\sum_{ij} a_{ij}^{[\alpha]}} \quad (3)$$



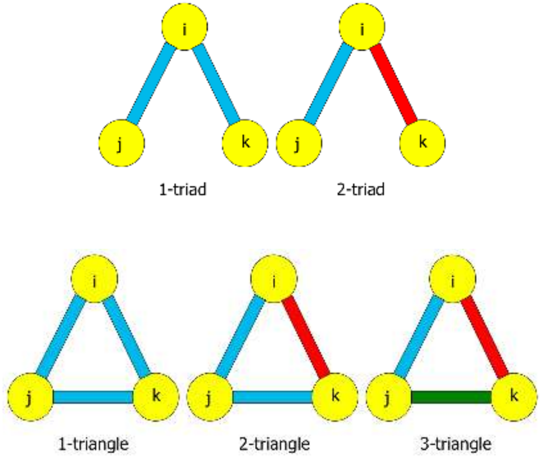
Edge overlap and social reinforcement

$$F(a_{ij}^{[\alpha']}|a_{ij}^{[\alpha]}) \rightarrow F^{\mathrm{w}}(a_{ij}^{[\alpha']}|w_{ij}^{[\alpha]})$$

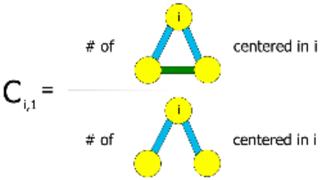


The existence of strong connections in the Trust layer, which represents the strongest relationships between two people, actually fosters the creation of links in other layers.

Triads and triangles

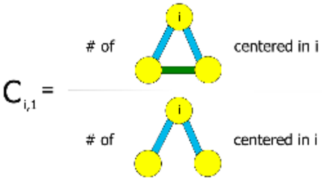


Clustering

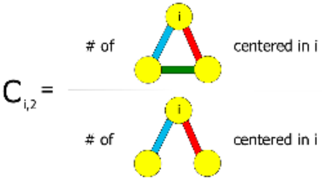


$$C_{i,1} = \frac{\sum_{\alpha} \sum_{\alpha' \neq \alpha} \sum_{j \neq i, m \neq i} (a_{ij}^{[\alpha]} a_{jm}^{[\alpha']} a_{mi}^{[\alpha]})}{\sum_{\alpha} \sum_{j \neq i, m \neq i} (a_{ij}^{[\alpha]} a_{mi}^{[\alpha]})} \tag{4}$$

Clustering

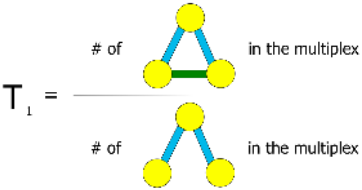


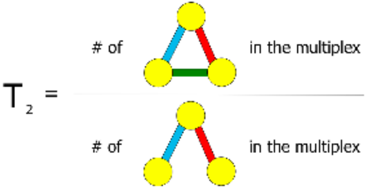
$$C_{i,1} = \frac{\sum_{\alpha} \sum_{\alpha' \neq \alpha} \sum_{j \neq i, m \neq i} (a_{ij}^{[\alpha]} a_{jm}^{[\alpha']} a_{mi}^{[\alpha]})}{\sum_{\alpha} \sum_{j \neq i, m \neq i} (a_{ij}^{[\alpha]} a_{mi}^{[\alpha]})} \quad (4)$$



$$C_{i,2} = \frac{\sum_{\alpha} \sum_{\alpha' \neq \alpha} \sum_{\alpha'' \neq \alpha, \alpha'} \sum_{j \neq i, m \neq i} (a_{ij}^{[\alpha]} a_{jm}^{[\alpha'']} a_{mi}^{[\alpha']})}{\sum_{\alpha} \sum_{\alpha' \neq \alpha} \sum_{j \neq i, m \neq i} (a_{ij}^{[\alpha]} a_{mi}^{[\alpha']})} \quad (5)$$

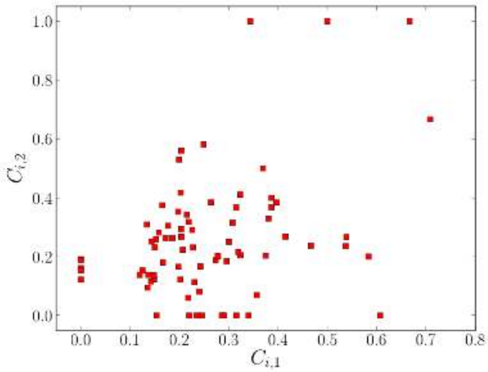
Transitivity

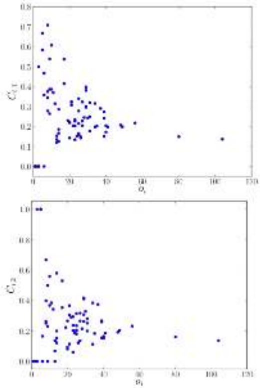




Clustering

$C_{i,1}$ and $C_{i,2}$ show different patterns of multi-clustering and are not correlated with o_i.





Transitivity and clustering

We call configuration model (CM) the set of multiplexes obtained from the original system by randomising edges and keeping fixed the sequence of degree vectors $\mathbf{k}_1, \mathbf{k}_2, \ldots, \mathbf{k}_N$, i.e. keeping fixed the degree sequence at each layer α.

Variable	Real data	Randomised data
C_1	0.26	0.17
C_2	0.26	0.18
T_1	0.21	0.15
T_2	0.21	0.16

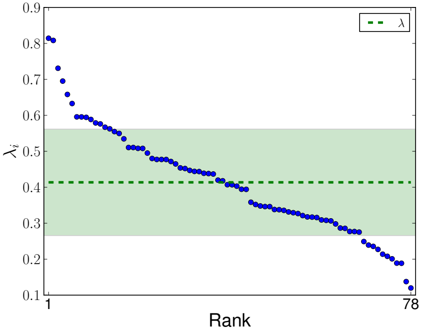
Measures of multi-clustering for real data are systematically higher than the ones obtained for randomised data, where edge correlations are washed out by randomisation

Navigability: shortest paths and interdependence

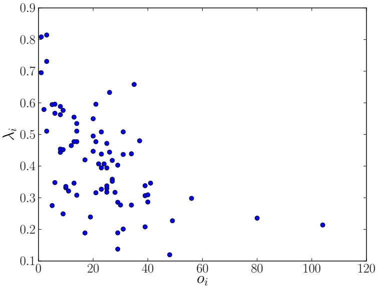
The interdependence λ_i of node i is defined as:

$$\lambda_i = \sum_{j \neq i} \frac{\psi_{ij}}{\sigma_{ij}} \tag{6}$$

where σ_{ij} is the total number of shortest paths between i and j and ψ_{ij} is the number of interdependent shortest paths between node i and node j.



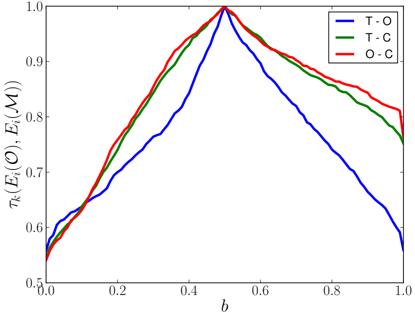
Reachability: shortest paths and interdependence



Centrality

For a duplex we can construct the following adjacency matrix:

$$\mathcal{M}(b) = bA^{[1]} + (1-b)A^{[2]}, \qquad \mathcal{M}(b=0.5) = \mathcal{O} \tag{7}$$



The symmetry/asymmetry of the curves tell us about the interplay between the layers in determining the centrality of the multi-layer system.
Both layers T and O dominate C in determining the centrality of the multiplex.

Summary

- We suggested a comprehensive formalism to deal with systems composed of several layers

- We also proposed a number of metrics to characterize multiplex systems with respect to:

  1. Node degree
  2. Node participation to different layers
  3. Edge overlap
  4. Clustering
  5. Transitivity
  6. Reachability
  7. Eigenvector centrality

  8. You can find more in the paper

"Structural measures for multiplex networks", *Phys. Rev. E* **89** 032804 (2014).
F. Battiston, V. Nicosia and V. Latora,
the LASAGNE team at Queen Mary University of London.

Acknowledge financial support from the