Statistical analysis of network data lecture 10

Sofia Olhede

EPFL

November 20, 2024

1 Barabasi Albert

2 Statistical Temporal Models

3 Multilayer Networks

Barabasi Albert

- We define the collection of edge variables at time step t to be $a_{ij}^{(t)}$.
- A network is sparse if

$$\lim_{t\to\infty} \frac{2}{t(t-1)} \sum_{i<j} a_{ij}^{(t)} = \lim_{n\to\infty} \hat{\rho} = 0.$$

- By the generating dynamics of the BA model there are m new edges at each step so $k_n = mn + k_0$.
- Thus

$$\widehat{\rho} = \frac{2(mn+k_0)}{n(n-1)} \to 0.$$

Barabasi Albert

- The degree distribution counts the relative proportion of vertices of any integer degree

$$p_{A^{(n)}}(k) = \sum_{i=1}^{n} \mathrm{I}\left(d_i^{(n)} = k\right),$$

and $p_{A^{(n)}}(k) \sim k^{-\gamma}$. This corresponds to a power law distribution.

- The generating mechanism of the preferential attachment model is very mechanistic.
- What if we want to generate a set of edges (or contacts) over time?
- We could simply start from the graph limit model

$$\mathbb{E}\{A_{ij} \mid \xi\} = \rho_n f(\xi_i, \xi_j).$$

- We shall assume we observe multiple graphs across time and so for a given edge ij we study edge-variable $A_{ij}(t)$.
- The simplest model takes the form:

$$\Pr\{A_{ij}(t) = 1 \mid A_{ij}(t-1), \ldots\} = h(t,\, A_{ij}(t-1),\, A_{ij}(t-2),\, A_{ij}(t-3), \ldots),$$

for some appropriately chosen function $h()$.

- We shall now look to how the edges can be modelled across times (see Süveges and Olhede (2023)) with label vector z:

$$A_{ij}(t) \mid A_{ij}(t-1), \ldots, A_{ij}(t-K) \sim \mathrm{Bern}\left(\mathrm{logit}^{-1}\left(\sum_{k=1}^{K} b_{z_i z_j k} A_{ij}(t-k) + c_{z_i z_j}(t)\right)\right). \quad (1)$$

- This model defines a correlated process across time, where the correlation is specified by b. Uses the trick of using a modelling framework that naturally limits the success probabilities between zero and one.
- Introduces flexible forms of serial correlation.
- The generating mechanism can still be estimated, and the properties of the network determined.
- Non-stationary models can we included by replacing $c_{ij,g}$ by a time-varying alternatives as we have above.
- Parameters naturally stay within the range of allowed values.

- The simplest version of this model takes $K = 1$.
- This introduces series correlation of length one.
- It is different from just letting the parameters of the stochastic block model change over time.
- Let us discuss how we would expect the realized edges to change?

Multilayer networks

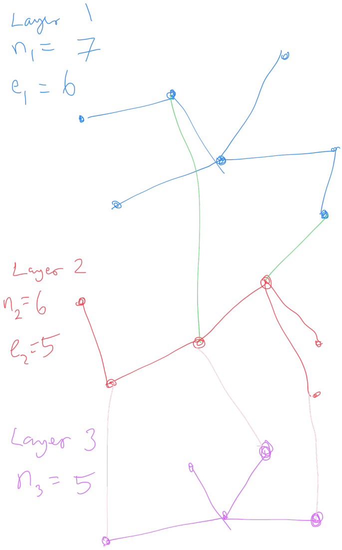
- A **multilayer network** is a pair $M = (G, C)$ where $G = \{G_p; p \in \{1, \dots, P\}$ is a family of simple graphs $G_p = (V_p, E_p)\}$. called the layers of M and C is

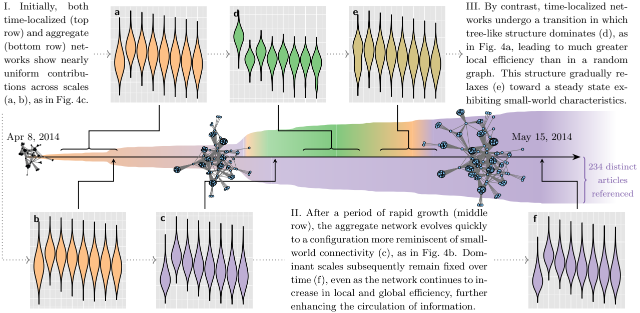
$$C = \{E_{ab} \subset V_p \times V_p, a, b \in 1, \dots, P, a \neq b\},$$

  is the set of **interconnection** between nodes of different layers G_a and G_b with $a \neq b$. The items in C are **crossed layers** and the elements of E_p are **intralayer connections**. Note that the elements of E_{ab} are **interlayer connections**.
- This model is useful to capture phenomena in social systems. The Axelrod model is an example thereof.
- A general multilayer network has potentially different vertices on different layers.
- We can have edges between nodes in different layers.

Multilayer networks

- We can collect all edges in a supra-adjacency matrix, which captures all edges present in the network.
- Obviously too complex.


Layer 1
$n_1 = 7$
$e_1 = 6$
Layer 2
$n_2 = 6$
$e_2 = 5$
Layer 3
$n_3 = 5$

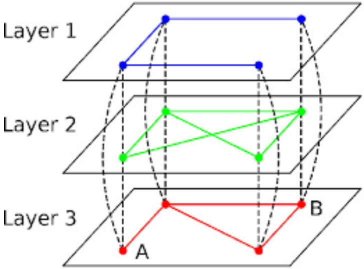

I. Initially, both time-localized (top row) and aggregate (bottom row) networks show nearly uniform contributions across scales (a, b), as in Fig. 4c.
a
d
e
III. By contrast, time-localized networks undergo a transition in which tree-like structure dominates (d), as in Fig. 4a, leading to much greater local efficiency than in a random graph. This structure gradually relaxes (e) toward a steady state exhibiting small-world characteristics.
Apr 8, 2014
May 15, 2014
234 distinct articles referenced
b
c
II. After a period of rapid growth (middle row), the aggregate network evolves quickly to a configuration more reminiscent of small-world connectivity (c), as in Fig. 4b. Dominant scales subsequently remain fixed over time (f), even as the network continues to increase in local and global efficiency, further enhancing the circulation of information.
f

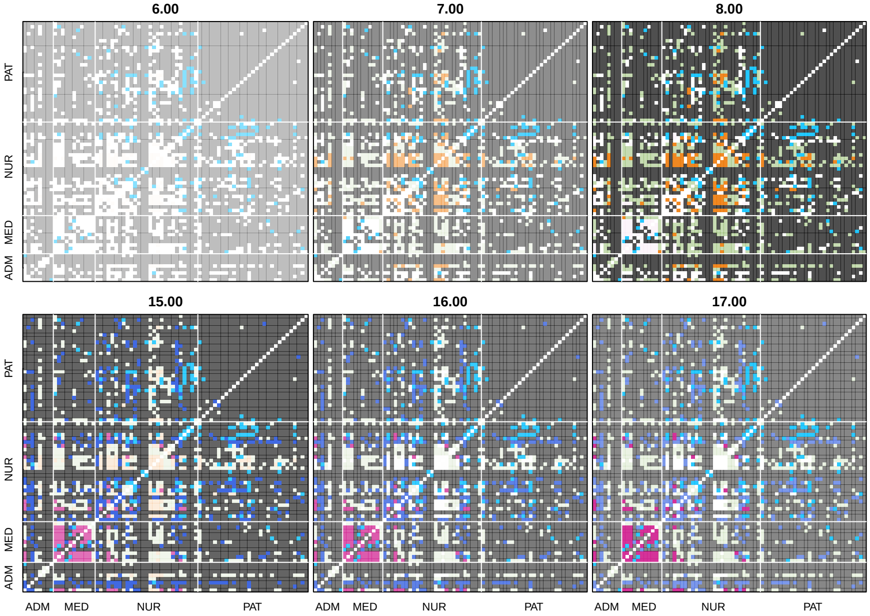
Multilayer networks II

- A special case of the multilayer network is the multiplex network. In this $V_1 = \cdots = V_P$. Thus a multiplex network has a fixed set of nodes that are connected with different edges.
- Simple graphs can be defined from the multiplex network. The projection network proj(M) has an adjacency matrix

$$\bar{A}_{ij} = \mathrm{I}\left(A_{ij}^{(p)} = 1\right), \text{ some } p.$$

- Multilayer networks occur from several different causes: Multiplex networks, Temporal networks, Interacting or interconnected networks, Multidimensional networks or Interdependent (or layered) networks.
- Multiplex networks are several networks with the same nodes, with layers of edges.
- A temporal network is several graphs G_t where it is recorded at time t.
- If we consider a family of networks $\{G_1, \ldots, G_L\}$ that interact, they can be modeled as a multilayer network of layers.


Layer 1
Layer 2
Layer 3
B
A


6.00
7.00
8.00
15.00
16.00
17.00
PAT
NUR
MED
ADM
ADM
MED
NUR
PAT

Multiplex networks

- Thus we can write the adjacency matrix as $A_{ij}(t)$, say but where the t index does not have to be time.
- Then the observations is the tensor $\{A_{ij}(t)\}$.
- We can define the time-varying degree vector $d(t)$, and any other time-varying statistics that we may chose.
- We can also estimate time-varying groups.

Directed Networks

- Social networks are often not symmetric, examples include such as an acquaintance network of university members (Kossinets & Watts (2006)), a large-scale instant-messaging network containing individuals (Leskovec & Horvitz (2008)), friendship networks of a set of American high schools (add-health, Currarini, Jackson & Pin), a social network of a cohort of college students in Facebook (Lewis, Gonzalez & Kaufman).
- We often have directed relationships on email and other modes of communication. Powerline and airline networks are directed, Citation networks likewise, Food webs (foxes eat hares, but hares do not eat foxes), Economic networks etc.

Directed Networks II

- To replicate the patterns we see in practice we need to generalize edges to directed edges.
- A directed edge $\overrightarrow{ij}$ is sometimes called an arc. An arc (i, j) is considered to be directed from i to j; j is called the head and i is called the tail of the arc.
- The adjacency matrix is then defined to be

$$A_{ij} = \mathrm{I}\left(\overrightarrow{ij} \quad \text{is} \quad \text{present.}\right)$$

- It follows that $A \neq A^T$.
- We now need two summaries for the popularity or sociability of a node, an in
- $d_i^{(in)}$ (pointing to i) and an out-degree $d_i^{(out)}$ (pointing away from i).

Directed Networks III

- The in-degrees are defined as the sum of the edges pointing to i:

$$d_i^{(\mathrm{in})} = \sum_{j \neq i} A_{ji}.$$

- The out-degrees are defined as the sum of the edges pointing away from i:

$$d_i^{(out)} = \sum_{j \neq i} A_{ij}.$$

- Furthermore versions of configuration model can be defined for directed graphs with parameters $\boldsymbol{\pi}^{(in)}$ and $\boldsymbol{\pi}^{(out)}$. Thus the degree based model becomes

$$\mathbb{E}\, A_{ij} = \pi_i^{(\mathrm{in})} \pi_j^{(out)}.$$

- Unlike standard configuration model, we now have $2n$ parameters. We could for example postulate $\boldsymbol{\pi}^{(\mathrm{in})} = \rho^{(\mathrm{in})} \boldsymbol{\pi}$ and $\boldsymbol{\pi}^{(out)} = \rho^{(out)} \boldsymbol{\pi}$.

Directed Networks III

- The in-degrees are defined as the sum of the edges pointing to i:

$$d_i^{(\mathrm{in})} = \sum_{j \neq i} A_{ji}.$$

- The out-degrees are defined as the sum of the edges pointing away from i:

$$d_i^{(out)} = \sum_{j \neq i} A_{ij}.$$

- Furthermore versions of configuration model can be defined for directed graphs with parameters $\boldsymbol{\pi}^{(in)}$ and $\boldsymbol{\pi}^{(out)}$. Thus the degree based model becomes

$$\mathbb{E}\, A_{ij} = \pi_i^{(\mathrm{in})} \pi_j^{(out)}.$$

- Unlike standard configuration model, we now have $2n$ parameters. We could for example postulate $\boldsymbol{\pi}^{(\mathrm{in})} = \rho^{(\mathrm{in})} \boldsymbol{\pi}$ and $\boldsymbol{\pi}^{(out)} = \rho^{(out)} \boldsymbol{\pi}$.

Directed Networks IV

EPFL

- We can also do a permutation invariant version of this degree-based model.
- We now need to 1 -d functions $f^{(\text{in})}(x)$ and $f^{(\text{out})}(x)$.
- We define two latent independent vectors ξ and η with uniform entries, and so take

$$\mathbb{E}A_{ij} \mid \xi, \eta = \rho_n \cdot f^{(\text{in})}(\xi_i) f^{(\text{out})}(\eta_j).$$

- We are therefore no longer in the land of symmetric adjacency matrices.

Directed Networks V

- There is a directed version of the graphon model. Specify a graph limit function $f(x, y)$ which is no longer symmetric.
- Specify α as a latent variable, as well as two latent independent vectors ξ and η with uniform entries, and so take

$$\mathbb{E}A_{ij} \mid \xi, \eta = \rho_n \cdot f_\alpha(\xi_i, \eta_j),$$

and conditional on the latent variables generate (A) as independent Bernoulli random variables.

Directed Networks VI

- Some graphs display clear group structure. For each node i we define a random variable $z_i^{(in)}$ that takes the value $\{1, \ldots, k_{(in)}\}$, where this variable is indicating the in-group membership of node i. For each node i we define a random variable $z_i^{(\text{out })}$ that takes the value $\{1, \ldots, k_{(\text{out })}\}$, where this variable is indicating the out-group membership of node i. We additionally define a connection probability matrix Θ which has entries θ_{ab} for $1 \leq a \leq k_{(\text{in })}, 1 \leq b \leq k_{(\text{out })}$. Then

$$A_{ij} \mid z_i^{(\text{in})}, z_j^{(\text{out })} = \text{Bernoulli}\left(\theta_{z_i^{(in)} z_j^{(\text{out })}}\right), \quad 1 \leq j \neq i \leq n.$$

where each realization is independent. Furthermore $A_{ii} = 0$ for $1 \leq i \leq n$, and we complete the matrix by. This is known as the **directed stochastic block model** (Wang & Wong, JASA 1987).

Directed Networks VII

- Wang and Wong proposed maximum likelihood estimation of this set of parameters, but they looked at very small examples.
- We can use the same type of algorithms as for regularly stochastic blockmodels.

$$\ell\left(\boldsymbol{\theta}, \mathbf{z}^{(\text{in})}, \mathbf{z}^{(\text{out})}\right) = \log\left\{\prod_{i\neq j} \theta_{z_i^{(\text{in})} z_j^{(\text{out})}}^{a_{\text{in}}} \left(1 - \theta_{z_i^{(\text{in})} z_j^{(\text{out})}}\right)^{1-a_{ij}}\right\}$$

$$= \sum_{i\neq j}\left\{a_{ij} \log \theta_{z_i^{(\text{in})} z_j^{(\text{out})}} + (1-a_{ij}) \log\left(1 - \theta_{z_i^{(\text{in})} z_j^{(\text{out})}}\right)\right\}.$$

- We can yet again use profile likelihood to estimate the parameters of the model.

Directed Networks VIII

- Having removed the assumption of symmetry, we could also start to look at relationships between different types of objects.
- We could therefore look at generalizations of A_{ij} that are not square, and have relationships between predators and prey for instance.
- We could also use the directed block model to estimate the graph limit without symmetry.