Session II (11 am – 12.30 pm)

Introduction to Network Science

By Yérali Gandica

5th International Conference on Computational Social Science
July 17-20, 2019, University of Amsterdam, The Netherlands

Watts-Strogatz algorithm for Small-world networks

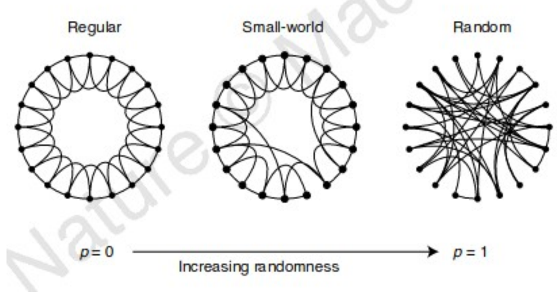
Starting from a ring lattice with n vertices and k edges per vertex, we rewire each edge at random with probability p.

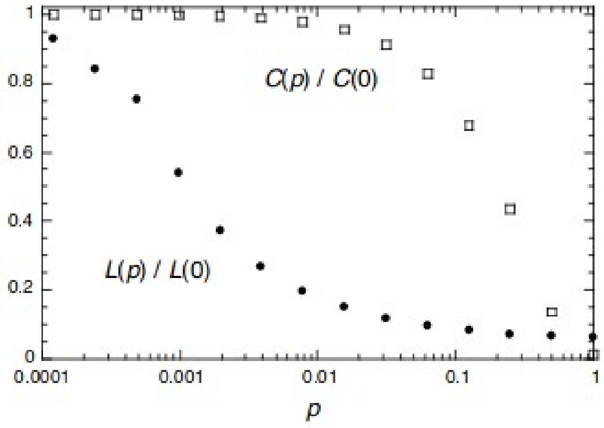
Collective dynamics of 'small-world' networks

Duncan J. Watts* & Steven H. Strogatz

Department of Theoretical and Applied Mechanics, Kimball Hall, Cornell University, Ithaca, New York 14853, USA

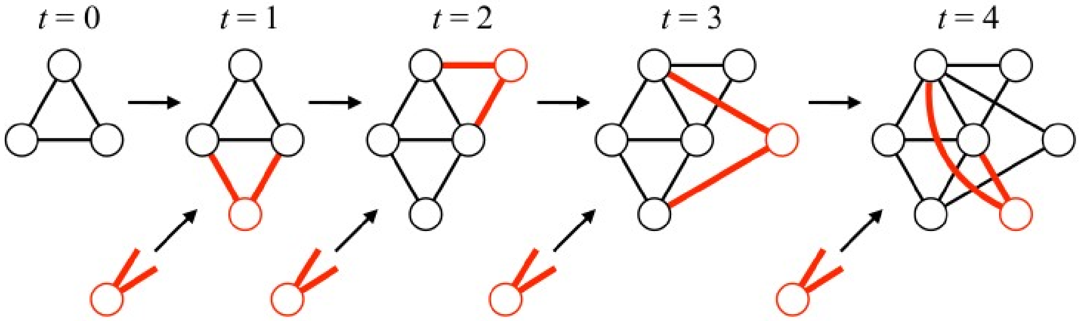
NATURE | VOL 393 | 4 JUNE 1998





Barabasi-Albert algorithm for Scale-free networks

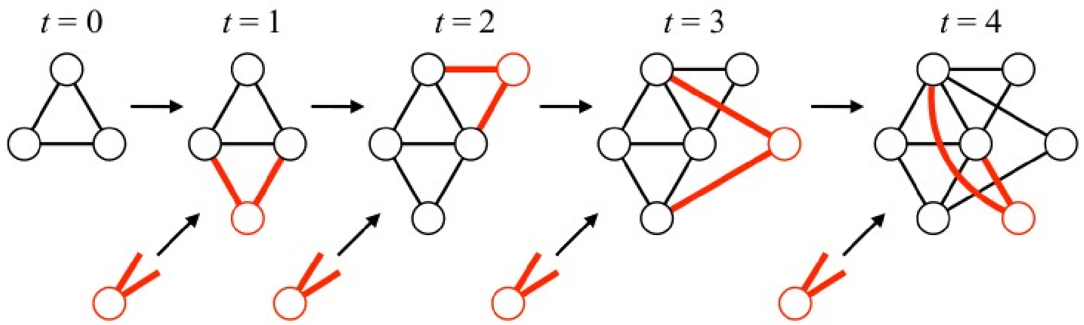
Prepare m_0 nodes, typically fully connected. Add a node with m($\leq m_0$) edges to the connected nodes. A node receives a new link with the probability proportional to its degree (preferential attachment). Continuing growing until achieving N nodes



First several stages of the BA model with $m_0 = 3$ and m = 2.
The bold lines represent new links.

Barabasi-Albert algorithm for Scale-free networks

Prepare m_0 nodes, typically fully connected. Add a node with m($\leq m_0$) edges to the connected nodes. A node receives a new link with the probability proportional to its degree (preferential attachment). Continuing growing until achieving N nodes

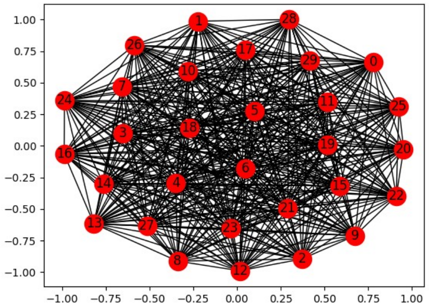


First several stages of the BA model with m_0 = 3 and m = 2.
The bold lines represent new links.

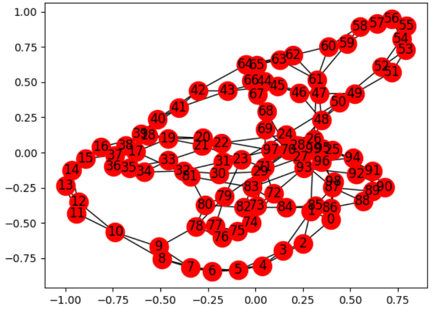
Session II (11 am – 12.30 pm)

Main models of networks

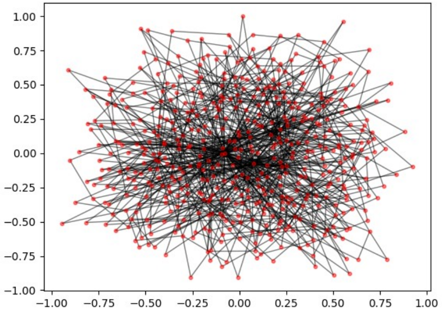
https://networkx.github.io/documentation/stable/reference/generators.html



Fully connected



Watts-Strogatz network
From 1-D



Barabasi-Albert algorithm

Session II (11 am – 12.30 pm)

Synthetic networks

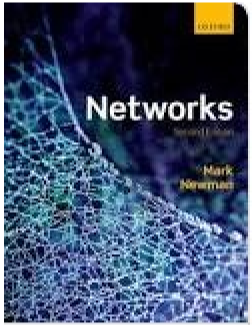
Let us plot:

Complete graphs,

Watts & Strogatz graphs, and

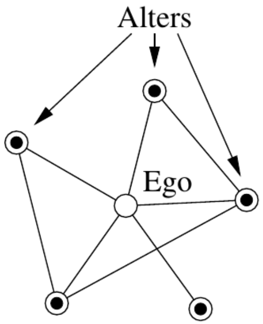
Barabasi & Albert graphs

Session II (11 am – 12.30 pm)



Personal networks or ego-centered networks:

An ego-centered network is a network surrounding one particular individual, meaning, usually, the individual surveyed and her or his immediate contacts. The individual surveyed is referred to as the ego and the contacts as alters.



Let's download some networks:
https://snap.stanford.edu/data/
→ ego-Facebook

Anoter repository:
http://konect.cc/

Two main approaches for community detection:

Structural approach: Based on the arrangement of nodes and links: Cut-based perspectives, clustering perspectives, Stochastically equivalent.

Dynamical approach: It is focused on the processes that take place on the network (For simple diffusion dynamics such as random walk on undirected network, this perspective es similar to the structural ones)

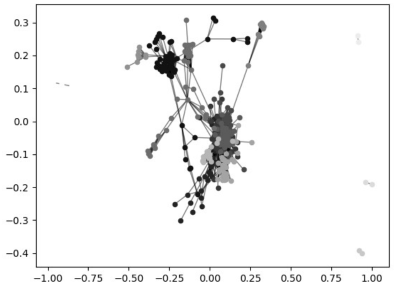
Community detection: I)

The Louvain method maximises Modularity

The underlying idea of modularity is to compare the number of links connecting nodes within communities, with the expected number of links in an appropriate null model.

It measures the density of edges inside communities to edges outside communities

Let’s try community detection over the Ego-network downloaded!!



Community detection: II)

The stochastic block model (SBM) by graph-tool

https://graph-tool.skewed.de/static/doc/index.html#

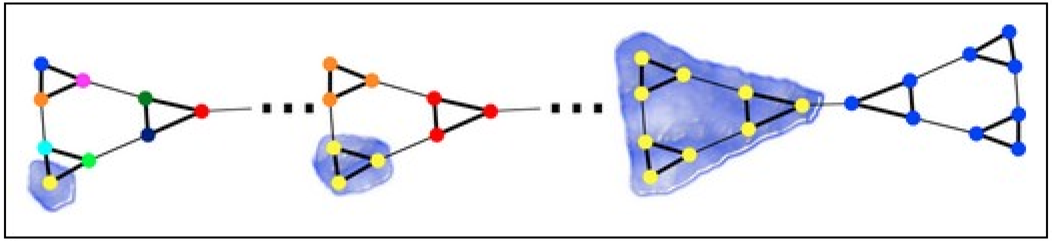
- Premise: The best partition is inferred by maximising the function providing that “nodes that belong to the same group possess the same probability of being connected with other nodes of the network.”

- Root in social network literature: nodes within a community connect with nodes in other communities in a equivalent way

- No interested in maximising some internal density. It is constrained to finding assortative group structure http://wwwf.imperial.ac.uk/~mpbara/Partition_Stability/download.html

Community detection: III)
The stability of a graph partition

Think of the simplest dynamics on the network (a random walk):
If the flow is trapped in a particular region ⇒ relevant community



Diffusion dynamics on the graph evolving over time

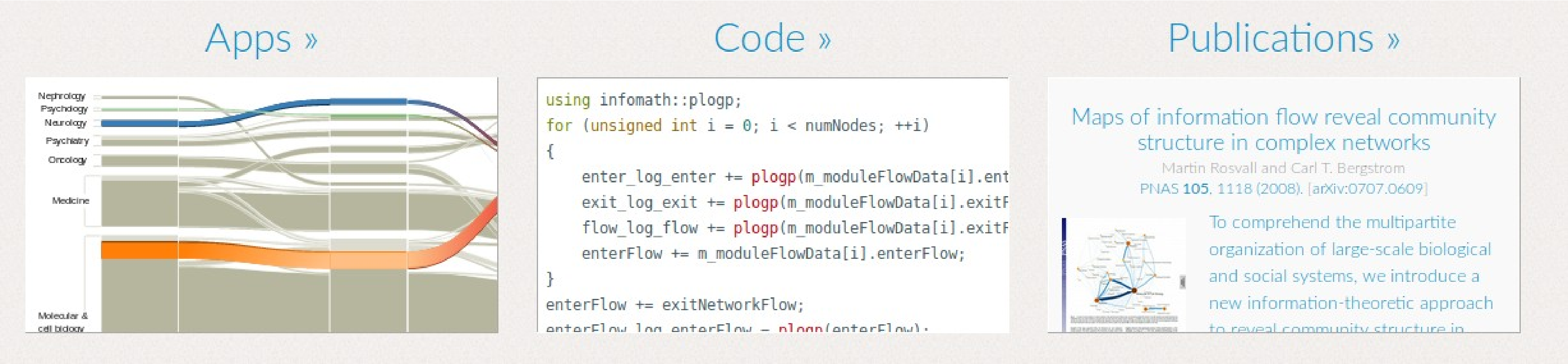
This method has an extra time-scale parameter!!!!

A group of nodes in the network with strong mixing within the group yet weakly coupled to the rest of the network

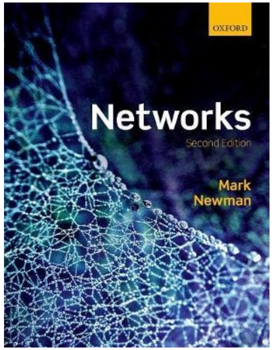
Community detection: IV)

Map equation: It is an information-theoretic approach that uses the duality between compressing data and finding regularities in the data.

https://www.mapequation.org/
https://mapequation.github.io/infomap/

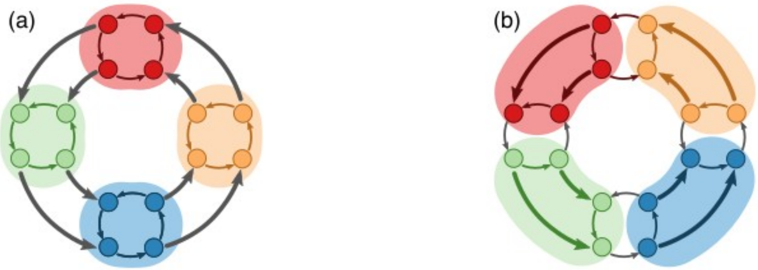


References



In the second edition there is a new part dedicated to Community detection

Different approaches to community detection. Martin Rosvall1, Jean-Charles Delvenne, Michael T. Schaub, and Renaud Lambiotte. https://arxiv.org/pdf/1712.06468.pdf



Dynamical properties (a: map equation) Vs. structural ones (b: modularity)

Session II (11 am – 12.30 pm)

Temporal networks

Evidence suggests that empirical networks observed in a variety of domains are far from static

Many networks grow in time (Example ?)

However, growth is not the only way in which networks evolve.

Nodes and links may emerge and disappear during the lifetime of a network (Example?)

Temporal networks

Evidence suggests that empirical networks observed in a variety of domains are far from static

Many networks grow in time (Example ?)

However, growth is not the only way in which networks evolve.

Nodes and links may emerge and disappear during the lifetime of a network (Example?)

For instance, the same link may be active just for a short period but repetitively with intermittent periods.

Moreover, the rate of link activation may depend on time: circadian and weekly rhythms of actors, interactions between different links, and so forth.

Session II (11 am – 12.30 pm)

Temporal networks

Evidence suggests that empirical networks observed in a variety of domains are far from static

Many networks grow in time (Example ?)

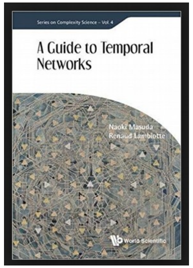
However, growth is not the only way in which networks evolve.

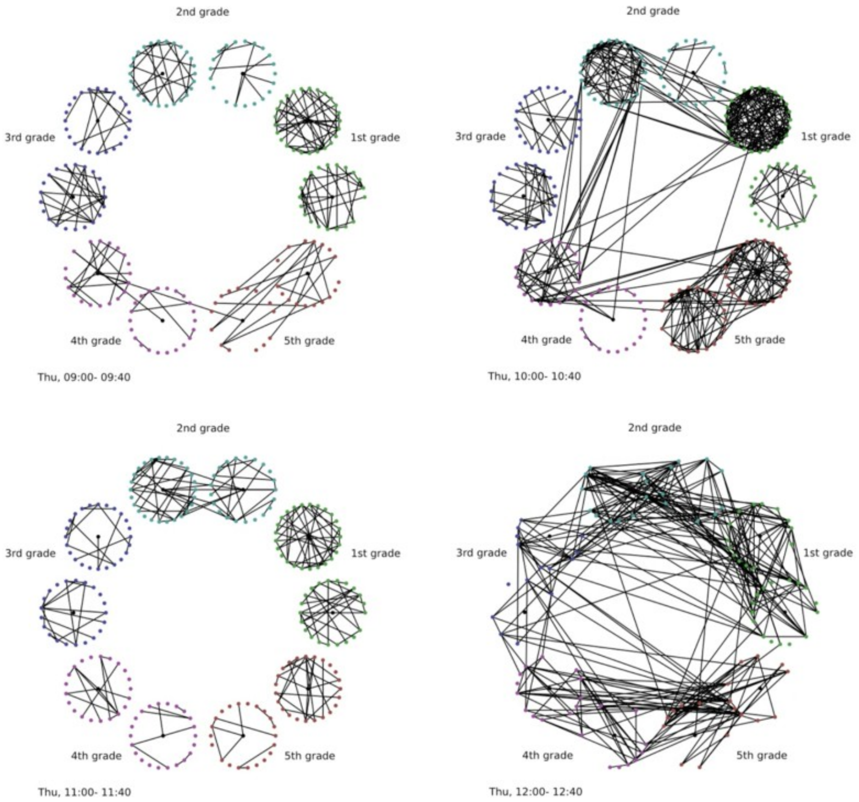
Nodes and links may emerge and disappear during the lifetime of a network (Example?)

These issues are at the core of the study of temporal networks, also called time/temporally varying networks/graphs, evolving graphs and evolutionary network analysis

Session II (11 am – 12.30 pm)

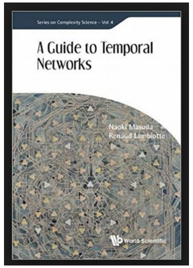
Why does temporality matter?





The spread of rumors, for example.

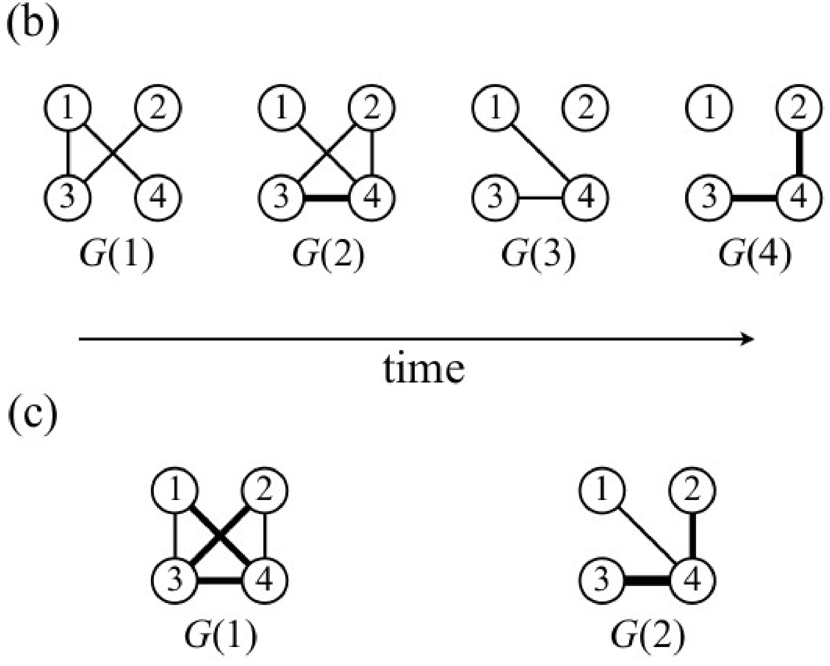
Snapshots of dynamic face-to-face contact patterns in a primary school. A circle represents a school class. A node in the centre of a circle is a teacher. The other nodes are pupils. Networks at four time points are shown.



Two representations of a temporal network

(a)

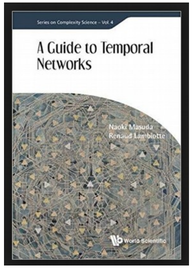
(1,3,1)
(1,4,1)
(2,3,2)
(1,4,4)
(3,4,5)
(2,3,6)
(2,4,6)
(3,4,6)
(1,4,7)
(3,4,9)
(2,4,10)
(3,4,10)
(2,4,11)
(3,4,12)



A network with N = 4 nodes, and an observation period of t_{max} = 12

(a) Event-based representation. (b) Snapshot representation with T_w = 3. It should be noted that, by definition, tmax in the snapshot representation is equal to the number of snapshots, i.e., tmax = 4. (c) Snapshot representation with T_w = 6, for which there are tmax = 2 snapshots. The width of lines in (b) and (c) is proportional to the link weight.

Session II (11 am – 12.30 pm)

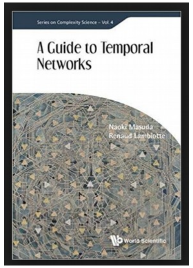


The shortest distance from u to v at time t:

$$d_{\text{short}}(u, v; t) = \min\{n : t_1 \geq t\}.$$

It is equal to the minimum number of hops necessary to travel from u to v along temporal paths. The temporal paths are constrained to start after time t.

Let’s calculate it!!!!



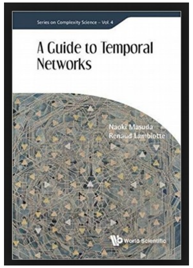
What about the centralities?

Temporal centralities:

A time-independent centrality for temporal networks is defined as a summary over time. It quantifies the overall importance of a node during the observation period. In contract with the time-dependent centrality, which is intended to capture temporal changes in the importance of a node.

Session II (11 am – 12.30 pm)

What about the centralities?



$$\text{betweenness}_i(t) = \frac{1}{(N-1)(N-2)} \sum_{j=1;j\neq i}^{N} \sum_{j'=1;j'\neq i,j}^{N} \frac{\sigma^i_{jj'}(t)}{\sigma_{jj'}}.$$

$\sigma_{jj'}$ is the number of the minimum distance temporal path from the jth to the j'th nodes in the entire observation period [0, tmax]. $\sigma^i_{jj'}(t)$ is the number of the minimum-distance temporal path from the jth to the j'th nodes that passes through the ith node and stays there at time t (i.e., reaches the ith node at time t or earlier and does not move to another node before t). The fraction of the minimum- distance paths that pass through the ith node is averaged over all starting and ending nodes of such paths. This normalisation is the same as that for the static-network counterpart.

Even for undirected temporal networks, the paths from one node to another and those in the opposite direction are separately considered, because the number and the distance of the qualified paths are generally different between the two directions.

$$\text{betweenness}_i = \frac{1}{t_{\max}} \sum_{t=1}^{t_{\max}} \text{betweenness}_i(t).$$

Session II (11 am – 12.30 pm)

References

A Guide to Temporal Networks
Series on Complexity Science: Volume 4
https://doi.org/10.1142/q0033 | September 2016
Naoki Masuda (University of Bristol, UK) and Renaud Lambiotte (University of Namur, Belgium)

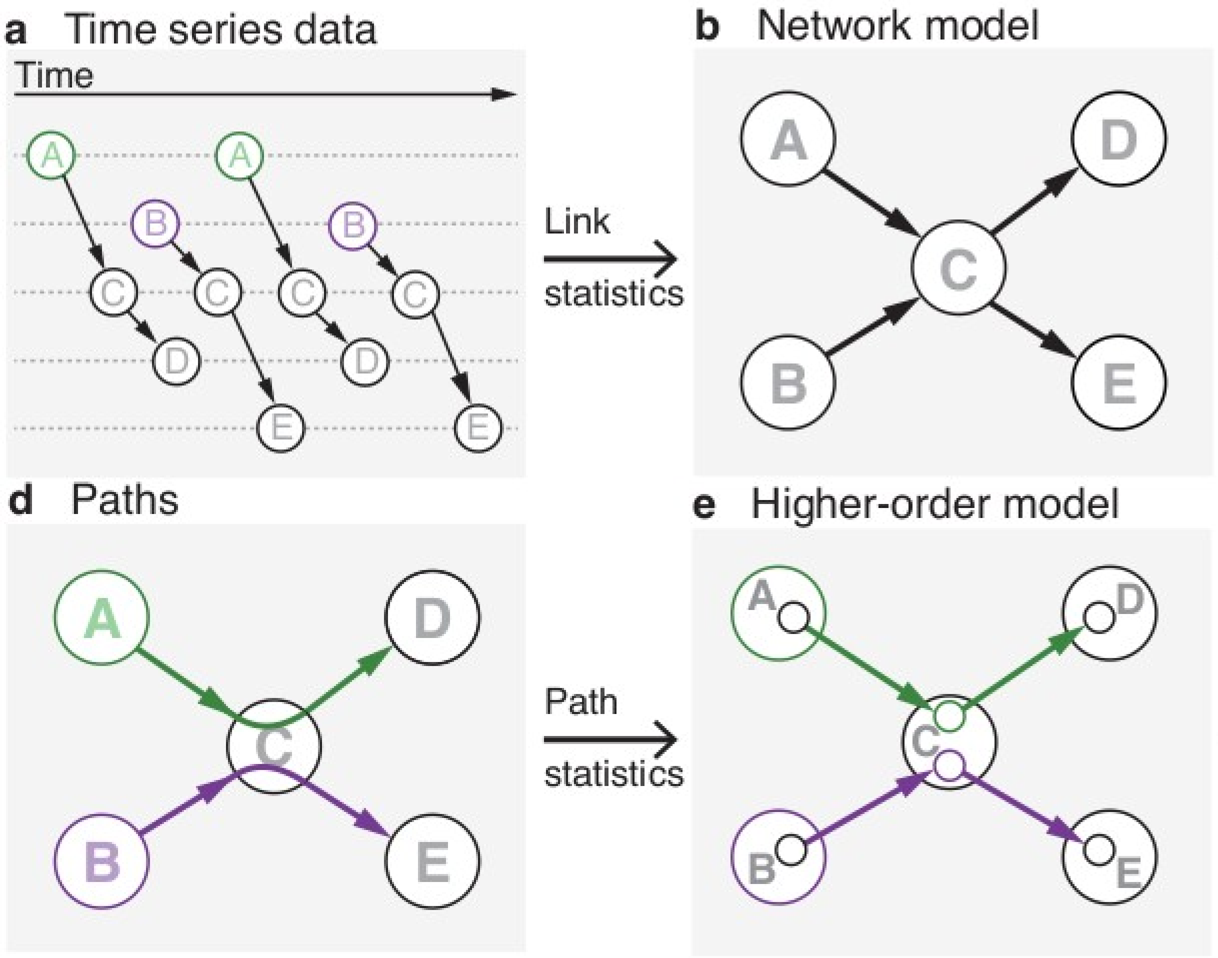
Higher-Order Aggregate Networks in the Analysis of Temporal Networks: Path structures and centralities: Ingo Scholtes, Nicolas Wider, Antonios Garas.
Eur. Phys. J. B (2016) 89: 61. https://doi.org/10.1140/epjb/e2016-60663-0

Second-order metrics: What are they about?

Let us take an idea:

http://www.higherordernetwork.com/

Session II (11 am – 12.30 pm)



References

Understanding Complex Systems: From Networks to Optimal Higher-Order Models

Renaud Lambiotte†, Martin Rosvall⋆, Ingo Scholtes‡

When is a Network a Network? Multi-Order Graphical Model Selection in Pathways and Temporal Networks

Ingo Scholtes

Packages

https://github.com/xyjprc/hon

https://github.com/IngoScholtes/pyTempNets

Thank you very much!!!

That is all for our Introduction

Please send me your comments at
ygandica@gmail.com