Community Detection in Networks with non-Markovian Dynamics

Mauro Faccin
Jean-Charles Delvenne
icteam, Université Catholique de Louvain, Belgium

CompleNet 2107 – Dubrovnik

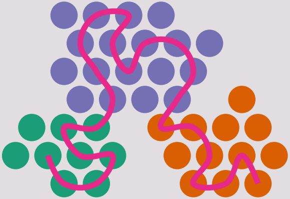
Projected Markov Chain

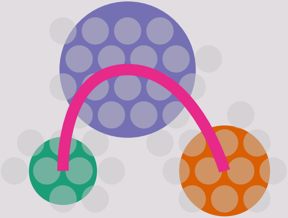
Markov Chain

$$\ldots, x_{\text{past}}, x_{\text{now}}, x_{\text{future}}, \ldots$$

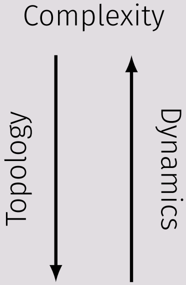
Projection

$$\ldots, Y_{\text{past}}, Y_{\text{now}}, Y_{\text{future}}, \ldots$$





Complexity

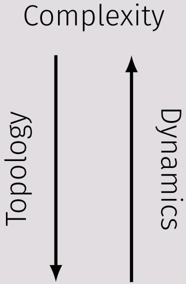


Complexity

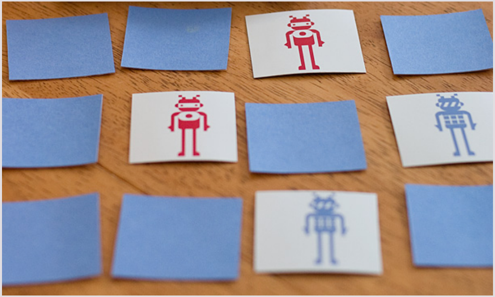
Topology

Dynamics

Complexity

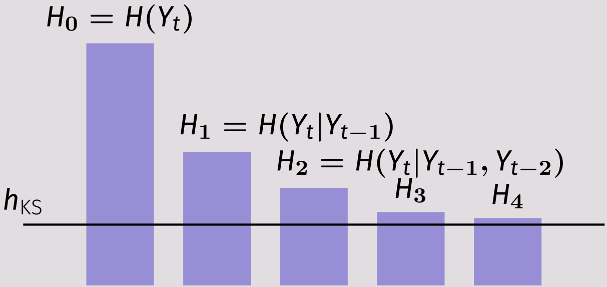


Memories



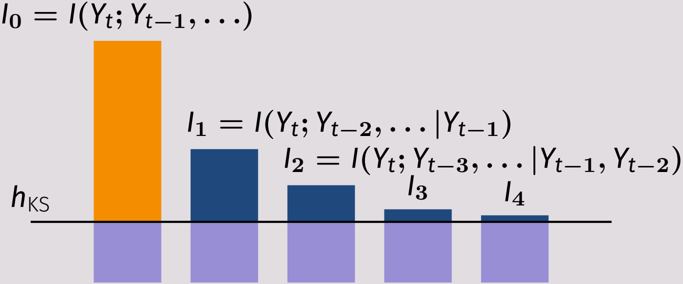
Entrogram

Measure the ***Statistical Complexity*** of the dynamics



Transient Entrogram

Measure the *Statistical Complexity* of the dynamics



$$I_0 = I(Y_t; Y_{t-1}, \ldots)$$

$$I_1 = I(Y_t; Y_{t-2}, \ldots | Y_{t-1})$$

$$I_2 = I(Y_t; Y_{t-3}, \ldots | Y_{t-1}, Y_{t-2})$$

Assess the Partition Quality 1

Projected Markov Chain:

$$\ldots, Y_{\text{past}}, Y_{\text{now}}, Y_{\text{future}}, \ldots$$

Maximize predictability:

$$I_0 = I(Y_{\text{future}}; Y_{\text{past}})$$

amount of information flowing from *past* to *future*.

$$I(Y_{\text{future}}; Y_{\text{past}}) \leq I(x_{\text{future}}; x_{\text{past}})$$

- Increases predictability of future with knowledge of past;
- Favors heterogeneous module-linking;
- Favors homogeneous module size;
- $I(\cdot;\cdot)$ is Mutual Information

Assess the Partition Quality 2

Projected Markov Chain:

$$\ldots, Y_{\text{past}}, Y_{\text{now}}, Y_{\text{future}}, \ldots$$

Minimize memories:

$$I_k = I(Y_{\text{future}}; Y_{\text{past}} | Y_{\text{now}})$$

higher order memory embedded into the process.

$$I(Y_{\text{future}}; Y_{\text{past}} | Y_{\text{now}}) \geq I(x_{\text{future}}; x_{\text{past}} | x_{\text{now}}) = 0$$

- Higher compression can require knowledge of the past dynamics (memories);
- Projected dynamics could differ from dynamics on the projected topology;
- $I(\cdot;\cdot)$ is Mutual Information

Put together:

$$\mathcal{F} = \sum_{i=0}^{k} \alpha_i \, I_i$$

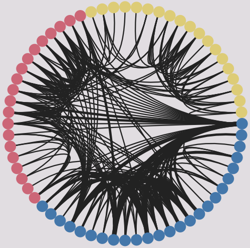
$$\mathcal{F} = \sum_{i=0}^{k} \alpha_i \, I(Y_t; Y_i \ldots | Y_{t-1} \ldots Y_{i+1})$$

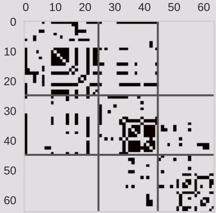
Predictability

$$\boxed{\mathcal{F}_P = I_0 - I_1 = I(Y_t; Y_{t-1})}$$

- Easy to compute
- Under some conditions is the same as DCSBM
- Plenty of algorithms

Blocks on Weighted Networks

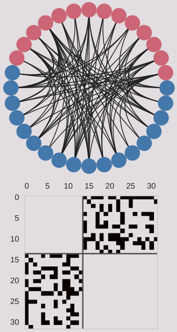


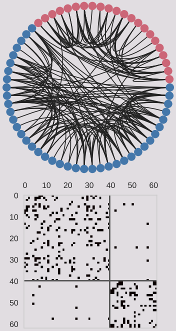


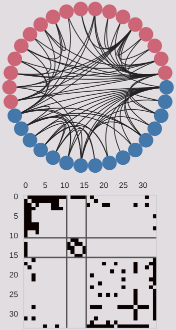
Contacts between suspect terrorists involved in the attack to Madrid station (2004).[1]

[1] *The March 11th Terrorist Network: In its weakness lies its strength, José A. Rodríguez*

Many others


DeepSouth


Dolphins

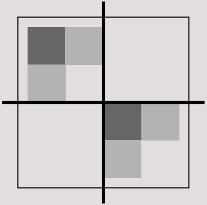

Karate

Predictability vs Markovianity

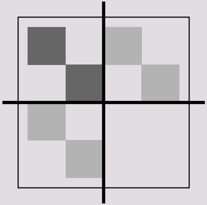
$$\boxed{\mathcal{F}_M = -I_1 = -I(Y_t; Y_{t-2}, \dots | Y_{t-1})}$$

- Markovian model dynamics
- Not biased toward predictable models
- Harder to compute but easy to implement with actual algorithms

What’s the best model?



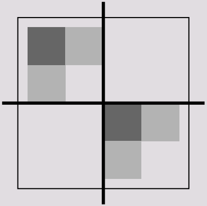
Better predictability of the dynamics



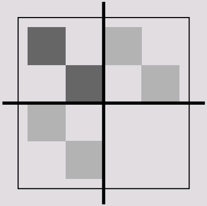
Also this is a perfect description of the system

$$\mathcal{F}_P(C) \quad > \quad \mathcal{F}_P(CP)$$

What’s the best model?



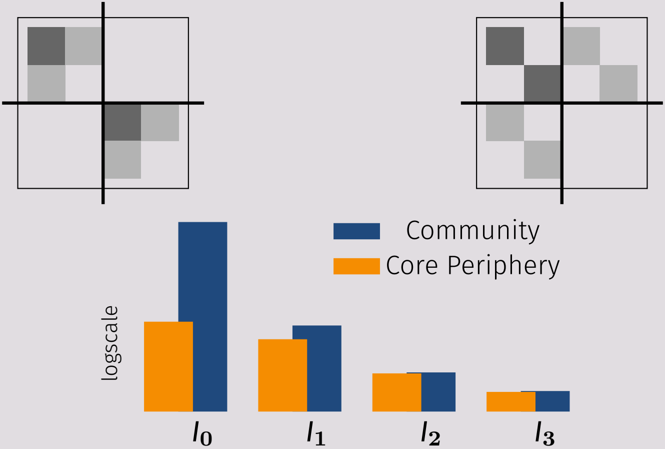
Better predictability of the dynamics



Also this is a perfect description of the system

$$\begin{array}{ccc} \mathcal{F}_P(C) & > & \mathcal{F}_P(CP) \\ \mathcal{F}_M(C) & \sim \mathbf{0} \sim & \mathcal{F}_M(CP) \end{array}$$

What’s the best model?

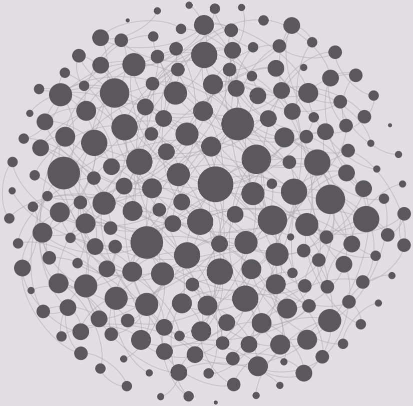


Non Markovian Dynamics

Non Markovian Dynamics

One can image a the Erdős–Rényi city with two quartiers where:

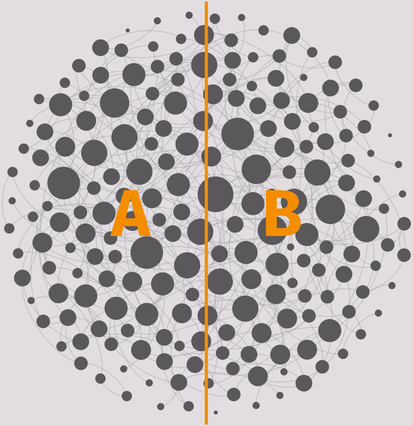
- people from each quartiers go to work on the morning;
- they go back home at night



Non Markovian Dynamics

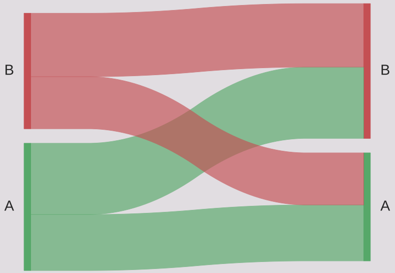
One can image a the Erdős–Rényi city with two quartiers where:

- people from each quartiers go to work on the morning;
- they go back home at night



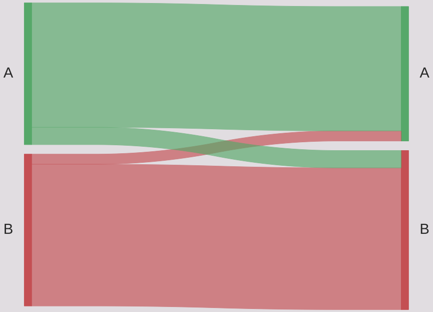
Non Markovian Dynamics

$I(Y_t; Y_{t-1})$
without memories:
no structure



Non Markovian Dynamics

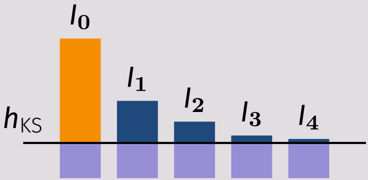
$I(Y_t; Y_{t-1}, Y_{t-2})$ considering memories: structure in dynamics



Concluding

- Entrogram, a entropy fingerprint of the projected dynamics.
- Dynamics define partitioning
- Predictability vs Markovianity
- Only needs dynamics (topology not necessary)
- Extends DCSBM to weighted graphs
- Non-Markovian Systems

Questions?



Joint work with:

JC Delvenne
@ ICTEAM and BigData Group,
UCLouvain.

`https://maurofaccin.github.io`

Good partitioning is the one leading to an interesting reduced model