Non-Markovian Models of Networked Systems

R. Lambiotte
Department of Mathematics
University of Namur

Network science in a nutshell

Say that one wants to model epidemic spreading

a) Construction of a network from empirical data, e.g. airline transportation network

b) Definition of a model for epidemic spreading, e.g. meta-population model

→ Effect of topology on spreading

Modeling step

Exact timings at which users meet are not considered: all that matters are fluxes between node (airports)

Real-world trajectories of the users are not considered: all that matters are fluxes between node (airports)

Now that such data is available: Accuracy of the network paradigm? Other paradigms?

Dynamics on stochastic temporal networks

What is the effect of the temporality of the network on a spreading process?

More and more empirical data incorporate information about the timing of activation of edges (e.g. **when** a phone call is made)

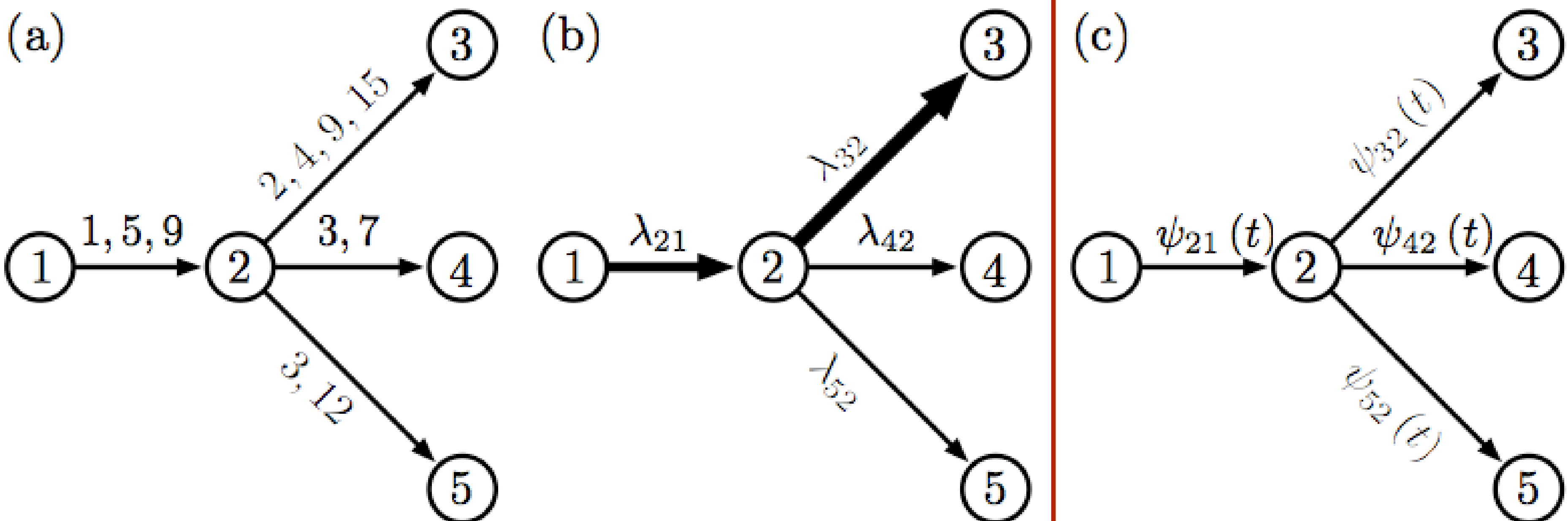
Non-trivial patterns of activation of nodes and edges
Burstiness: intermittent switching between periods of low activity and high activity, and a fat-tailed inter-event time distributions.

Temporal Networks, Petter Holme, Jari Saramäki, Phys. Rep. 519, 97-125 (2012)

Dynamics on stochastic temporal networks

What is the effect of the temporality of the network on a spreading process?

More and more empirical data incorporate information about the timing of activation of edges (e.g. **when** a phone call is made)



Temporal network, usually studied by means of computer simulations

Static network representation where the timing of events are Poisson processes

Stochastic temporal network where the sequence of activation times is a stochastic model that preserves the observed inter-event distribution

Temporal Networks, Petter Holme, Jari Saramäki, Phys. Rep. 519, 97-125 (2012)

From data to models

Let us consider a system of N nodes observed during a time interval T

We focus on the activation of edges going from i to j.

$$t_{ij} = \{t_{ij}^{(1)}, t_{ij}^{(2)}, ..., t_{ij}^{(n_{ij})}\}$$

The exact sequence of activation times is by a random sequence where events take place according to an inter-activation time fitted on the data

$f_{ij}(\tau)d\tau$ → probability to observe a time interval of duration in $[\tau, \tau + d\tau]$ between two activations of the edge

$\int_0^\infty \tau f_{ij}(\tau)d\tau = \langle\tau\rangle_{ij}$ → expected time between two activations of an edge

From data to models

When modeling the diffusion of an entity on the network, the distribution $f_{ij}(\tau)$ only plays an indirect role. The important quantity is instead the waiting time distribution $\psi_{ij}(t)$ that the entity arriving on i has to wait for a duration t before an edge towards j is available.

In epidemic spreading, it is the time it takes for a newly infected node to spread the infection further via the corresponding link.

Assuming that the activations of neighbouring edges are independent

$$\longrightarrow \qquad \psi_{ij}(t) = \frac{1}{\langle \tau \rangle_{ij}} \int_t^{\infty} f_{ij}(\tau) d\tau$$

If the activations of neighbouring edges are independent, *\psi*$_{ij}$*(t)* can be directly measured in empirical data

A.O. Allen. *Probability, Statistics, Queueing Theory: With Computer Science Applications*, 1990.

From data to models

$$\langle t \rangle_{ij} = \int_0^\infty t\psi_{ij}(t)dt = \frac{1}{2}\frac{\langle \tau^2 \rangle_{ij}}{\langle \tau \rangle_{ij}}$$

At a fixed value of the average inter-activation time, the waiting time can be arbitrarily large if the variance of inter-activation times is sufficiently large. This paradox, often called waiting time paradox or bus paradox in queuing theory, is an example of length-biased sampling.

Waiting-times and inter-activation times have the same distribution when the process is Poissonian, in which case

$$\psi_{ij}(t) = f_{ij}(t) = \frac{1}{\langle t \rangle_{ij}} \exp\left(-\frac{1}{\langle t \rangle_{ij}}\right)$$

Their tail has the same nature in the case of power-law tails

$$\psi_{ij}(t) \sim t^{-\alpha} \;\Leftrightarrow\; f_{ij}(\tau) \sim \tau^{-(\alpha+1)}$$

A.O. Allen. *Probability, Statistics, Queueing Theory: With Computer Science Applications*, 1990.

Effect on spreading

In the literature, authors focus on a limited number of families of distributions (gamma, power-law, stretched exponential, log-normal), and on the effect of the tail of the distribution.

Which properties of the waiting time tend to affect (accelerate or slow down) spreading processes:
The average waiting time?
Its variance?
The tail of the distribution?

Effect on spreading

Different temporal properties affect different spreading models, and different spreading properties:

Time ordering of events, how the probability mass of different probabilities are distributed:
- Biases in the trajectories of random walkers (*structural* effect on mixing time)

Effect on spreading: Random Walks

A walker located at a node i remains on it until an edge leaving *i* toward some node j appears. When such an event occurs, the walker jumps to *j* without delay and then waits until an edge leaving *j* appears.

The probability for the walker to jump to *j* depends on $\psi_{ij}(t)$, but also on all $\psi_{ik}(t)$, where *k* are neighbours of *i*, because the walker takes the first edge available for transport. Once a walker has left a node, edges leaving this node become useless for transport. For this reason, the probability to actually make a step from *i* to *j* at time *t* is given by

$$T_{ij}(t) = \psi_{ij}(t) \times \prod_{k \neq j} \int_t^\infty \psi_{ik}(t')\, dt'$$

When two neighbours:

$$T_{ij}(t) = \psi_{ij}(t) \int_t^\infty \psi_{ik}(t')\, dt'$$

The probability for making a jump to node j is given by the effective transition matrix

$$\mathbb{T}_{ij} \equiv \int_0^\infty T_{ij}(t)dt \qquad \sum_j \mathbb{T}_{ij} = 1$$

Generalized Master Equations for Non-Poissonian Dynamics on Networks, Till Hoffmann, Mason Porter and R.L., Physical Review E 2012

Effect on spreading: Random Walks

Generalized Montroll-Weiss Equation (usually for CTRW with non-Poisson inter-event time statistics on lattices)

$$\hat{n}(s) = \frac{1}{s}\left(I - \hat{D}_T(s)\right)\left(I - \hat{T}(s)\right)^{-1} n(0)$$

$$\frac{dn}{dt} = \left(T(t) * \mathcal{L}^{-1}\left\{\hat{D}_T^{-1}(s)\right\} - \delta(t)\right) * K(t) * n(t)$$

Convolution in time

Memory kernel

Effective transition matrix

Time ordering

$$T_{ij}(t) = \psi_{ij}(t) \times \prod_{k \neq i} \chi_{kj}(t)$$

$$= \psi_{ij}(t) \times \prod_{k \neq i}\left(1 - \int_0^t \psi_{kj}(t')\,dt'\right).$$

Generalized Master Equations for Non-Poissonian Dynamics on Networks, T.H., M.A.P. and R.L.

Effect on spreading

Different temporal properties affect different spreading models, and different spreading

Time ordering of events, how the probability mass of different probabilities are distributed:
- biases in the trajectories of random walkers (*structural* effect on mixing time)

Variance, exponential cut-off and spectral gap:
- mixing time for random walker processes

Mixing time

We consider a strongly connected, acyclic underlying network

For discrete time random walks on static networks:

$$p(t+1) = p(t)P,$$

and its solution is given by $p(t) = P^t$.

For a Poisson, continuous-time random walks on static networks:

$$\dot{p}(t) = -p(t)\tau^{-1}(I - P),$$

where I is the identity matrix, and where the solution is given by $p(t) = \exp(-\tau^{-1}(I - P))$. $I - P$ is the normalised Laplacian of the underlying, possibly weighted and directed, network.

The second dominant eigenvalue (spectral gap) determines the characteristic time necessary for the process to reach stationarity

Mixing time

In the case of continuous-time random walks with (the same) arbitrary waiting-time distribution:

$$\mathbf{p}(s) = \frac{1-\rho(s)}{s}(I-\rho(s)P)^{-1}$$

Development in small s gives the long-time behaviour:

$$\tau_{mix} = \max(\tau_0, \tau_{tail})$$

$$\tau_0 \approx \frac{\tau}{\epsilon}(1+\beta\epsilon) = \tau(\epsilon^{-1}+\beta)$$

$$\beta = (\sigma^2-\tau^2)/2\tau^2$$

$$\exp(-t/\tau_{tail})$$

Characteristic time of the exponential tail (cut-off)

Mixing time

$$\tau_0 \approx \frac{\tau}{\epsilon}(1 + \beta\epsilon) = \tau(\epsilon^{-1} + \beta)$$

$$\beta = (\sigma^2 - \tau^2)/2\tau^2$$

Combination of temporal and structural information. Burstiness slows down the walker more efficiently on networks with large epsilon: random networks that have no bottlenecks, such as the Erdos-Renyi and configuration models, or small diameter graphs with no communities.

This factor incorporates the variance of the waiting time, not that of the inter-event time !!!!!!

Mixing time

In Eq. 10, the approximation $\epsilon^{-1} + \beta \approx \max(\epsilon^{-1}, \beta)$ is valid, provided that the two terms are positive and dissimilar in order of magnitude. Under those conditions, the mixing time is given by

$$\tau_{mix} \approx \max(\ \tau/\epsilon,\ \tau\beta,\ \tau_{tail}). \tag{11}$$

Eq. 11 highlights three competing factors regulating the mixing time. While τ/ϵ is essentially a topological factor, capturing the effect of the structural bottleneck, the second term captures the burstiness-driven slowdown (as in [3]) and τ_{tail} quantifies the 'fatness' of the tail of the distribution of waiting times [7, 9]. Burstiness

Either of the three factors may be dominant in real-life data.

Bottlenecks, burstiness, and fat tails regulate mixing times of non-Poissonian random walks, J.-C. Delvenne, Renaud Lambiotte and L. E. C. Rocha, arXiv:1309.4155

Effect on spreading

Waiting time distribution

Different temporal properties affect different spreading models, and different spreading

Time ordering of events, how the probability mass of different probabilities are distributed:
- biases in the trajectories of random walkers (*structural* effect on mixing time)

Variance, exponential cut-off and spectral gap:
- mixing time for random walker processes

Correlations

Correlations between successive events: effect on the spectral gap

From data to models

When modeling the diffusion of an entity on the network, the distribution $f_{ij}(\tau)$ only plays an indirect role. The important quantity is instead the waiting time distribution $\psi_{ij}(t)$ that the entity arriving on *i* has to wait for a duration *t* before an edge towards *j* is available.

In epidemic spreading, it is the time it takes for a newly infected node to spread the infection further via the corresponding link.

Assuming that the activations of neighbouring edges are independent

$$\psi_{ij}(t) = \frac{1}{\langle \tau \rangle_{ij}} \int_t^\infty f_{ij}(\tau) d\tau$$

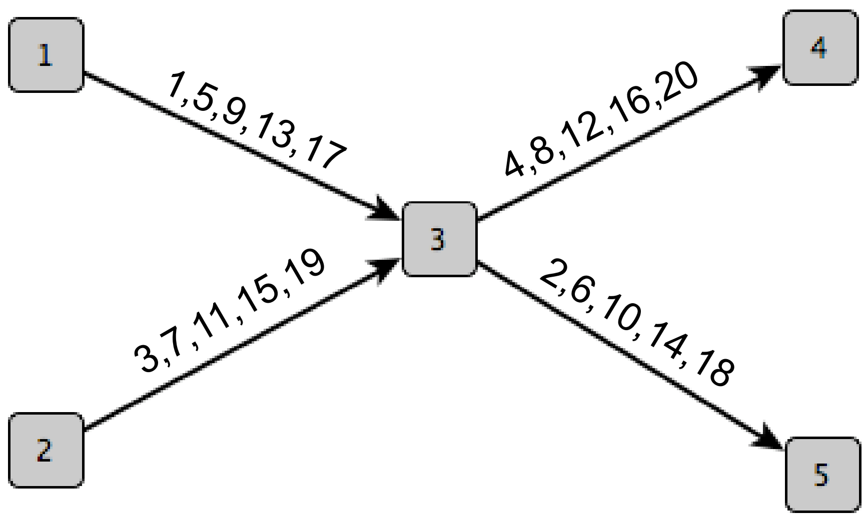
If the activations of neighbouring edges are independent, *\psi*$_{ij}$*(t)* can be directly measured in empirical data

A.O. Allen. *Probability, Statistics, Queueing Theory: With Computer Science Applications*, 1990.

Memory affects mixing time

Correlations between the activation times of neighbouring edges

=> induces non-random pathways: where one goes to depends on where one comes from



Memory network
Betweenness preference → Second-order Markov

Networks with Memory, Martin Rosvall, Alcides V. Esquivel, Andrea Lancichinetti, Jevin D. West, Renaud Lambiotte, arXiv:1305.4807
Slow-Down vs. Speed-Up of Information Diffusion in Non-Markovian Temporal Networks, I Scholtes, N Wider, R Pfitzner, A Garas, C Juan Tessone and F Schweitzer, arXiv:1307.4030

Network science in a nutshell

Say that one wants to model epidemic spreading

a) Construction of a network from empirical data, e.g. airline transportation network

b) Definition of a model for epidemic spreading, e.g. meta-population model

→ Effect of topology on spreading

Modeling step

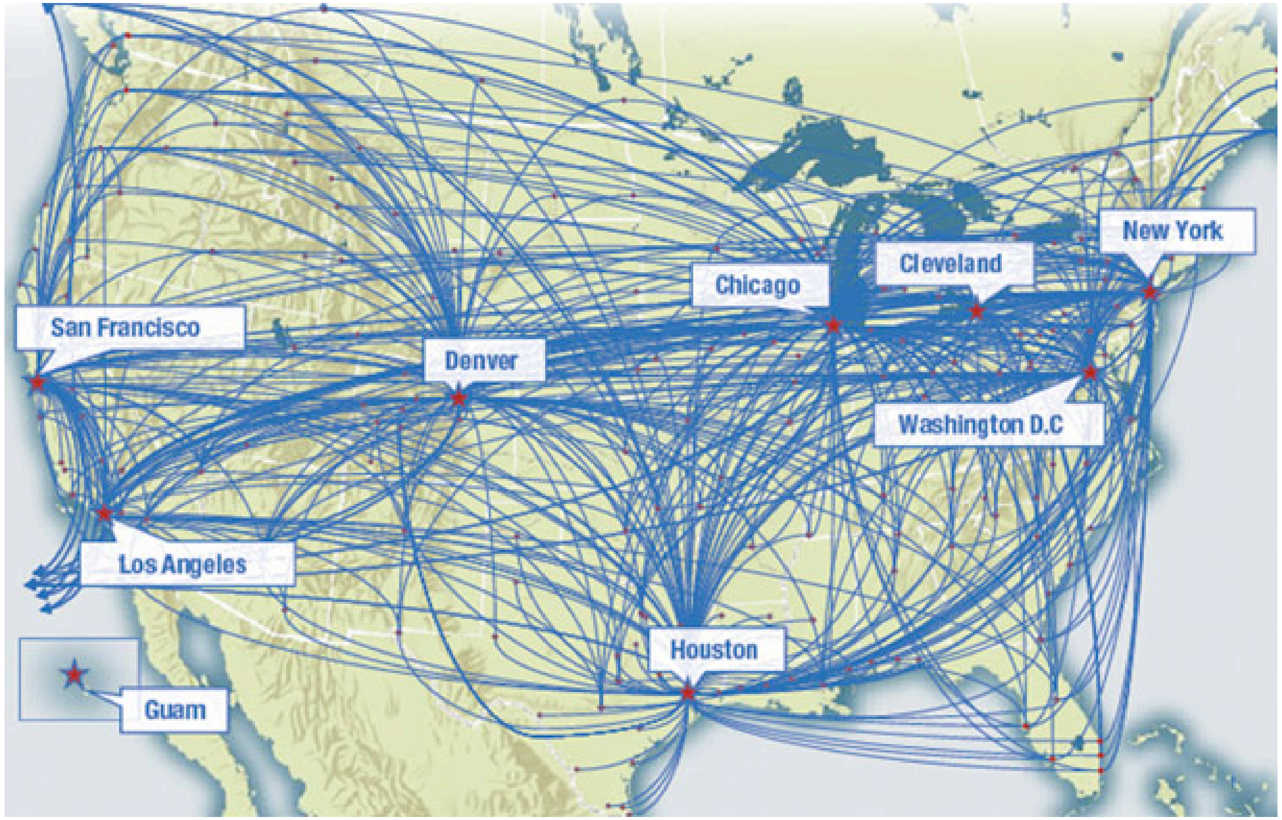
Exact timings at which users meet are not considered: all that matters are fluxes between node (airports)

Real-world trajectories of the users are not considered: all that matters are fluxes between node (airports)

Now that such data is available: Accuracy of the network paradigm? Other paradigms?

Networks with memory

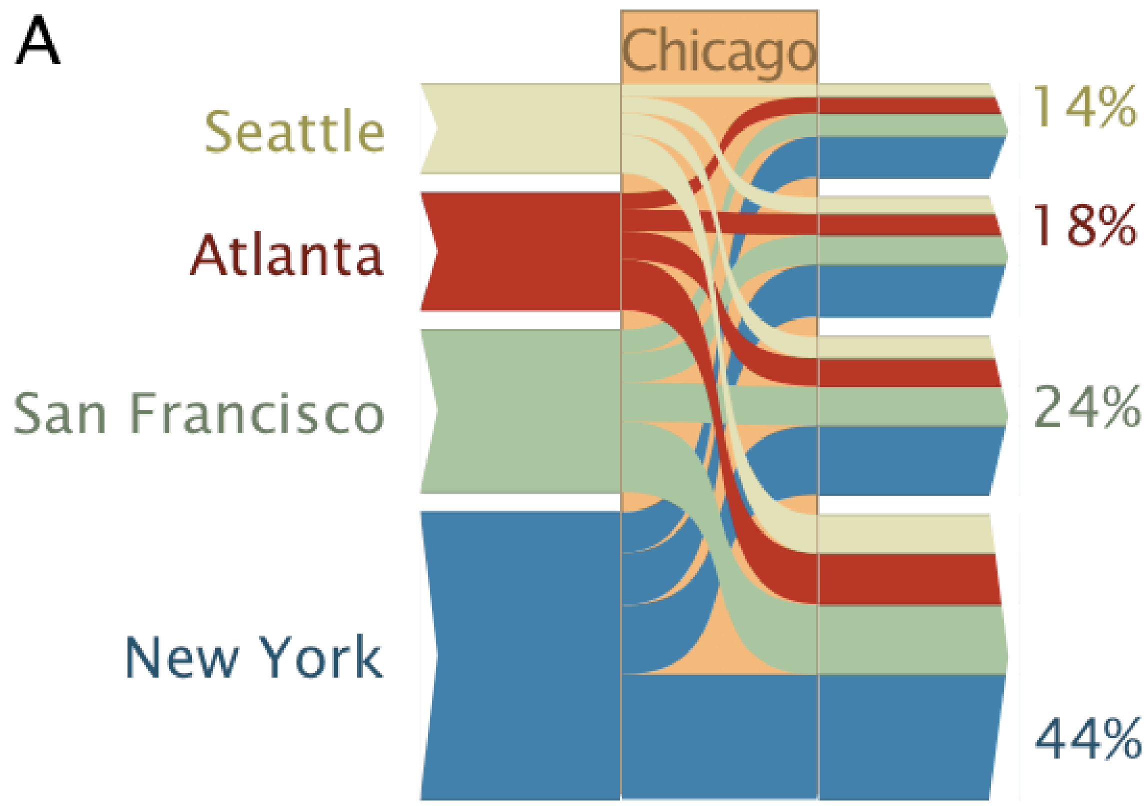
Where you go to depends on where you come from
Mathematics of pathways instead of edges



Networks with Memory, Martin Rosvall, Alcides V. Esquivel, Andrea Lancichinetti, Jevin D. West, Renaud Lambiotte, arXiv:1305.4807

Networks with memory

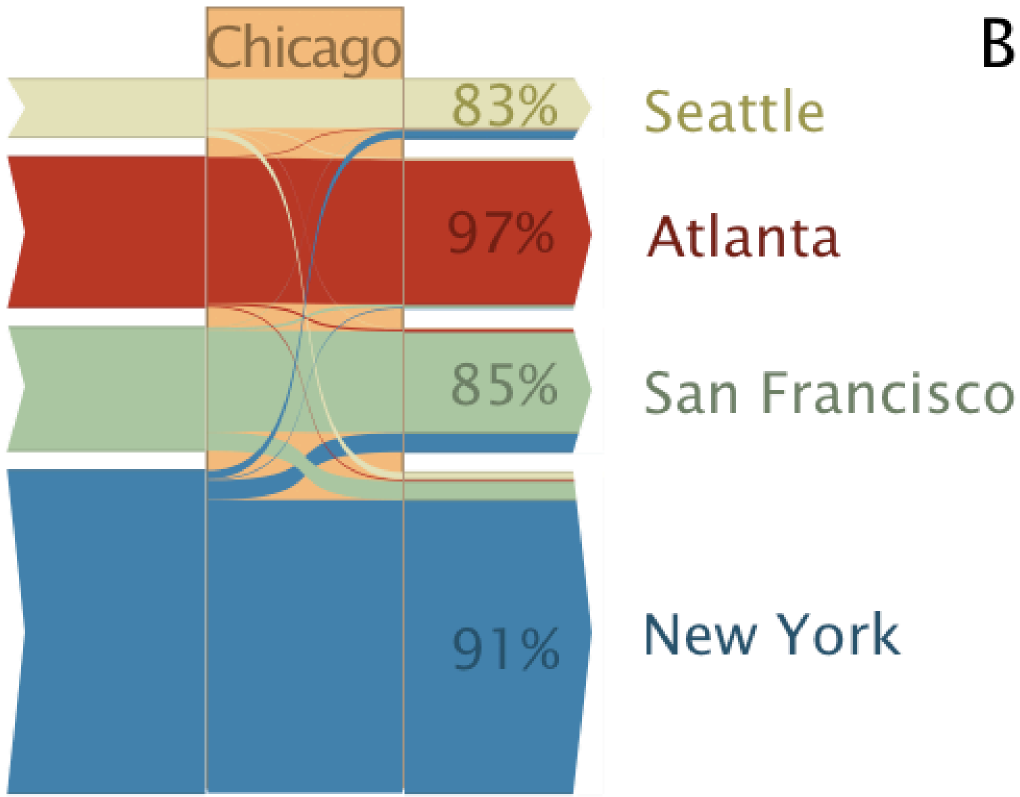
Where you go to depends on where you come from
Mathematics of pathways instead of edges



Networks with Memory, Martin Rosvall, Alcides V. Esquivel, Andrea Lancichinetti, Jevin D. West, Renaud Lambiotte, arXiv:1305.4807

Networks with memory

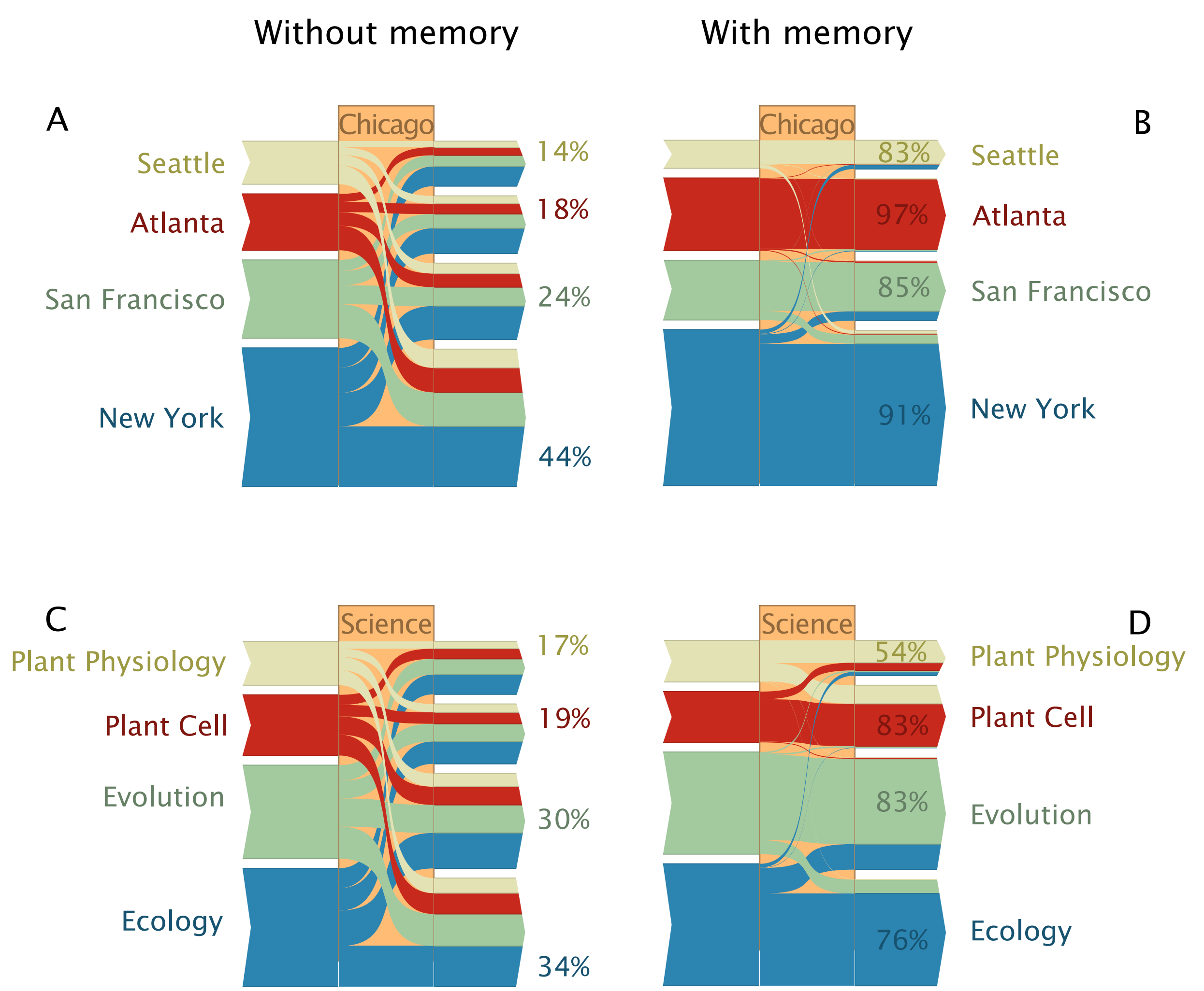
Where you go to depends on where you come from
Mathematics of pathways instead of edges



Networks with Memory, Martin Rosvall, Alcides V. Esquivel, Andrea Lancichinetti, Jevin D. West, Renaud Lambiotte, arXiv:1305.4807

Networks with memory

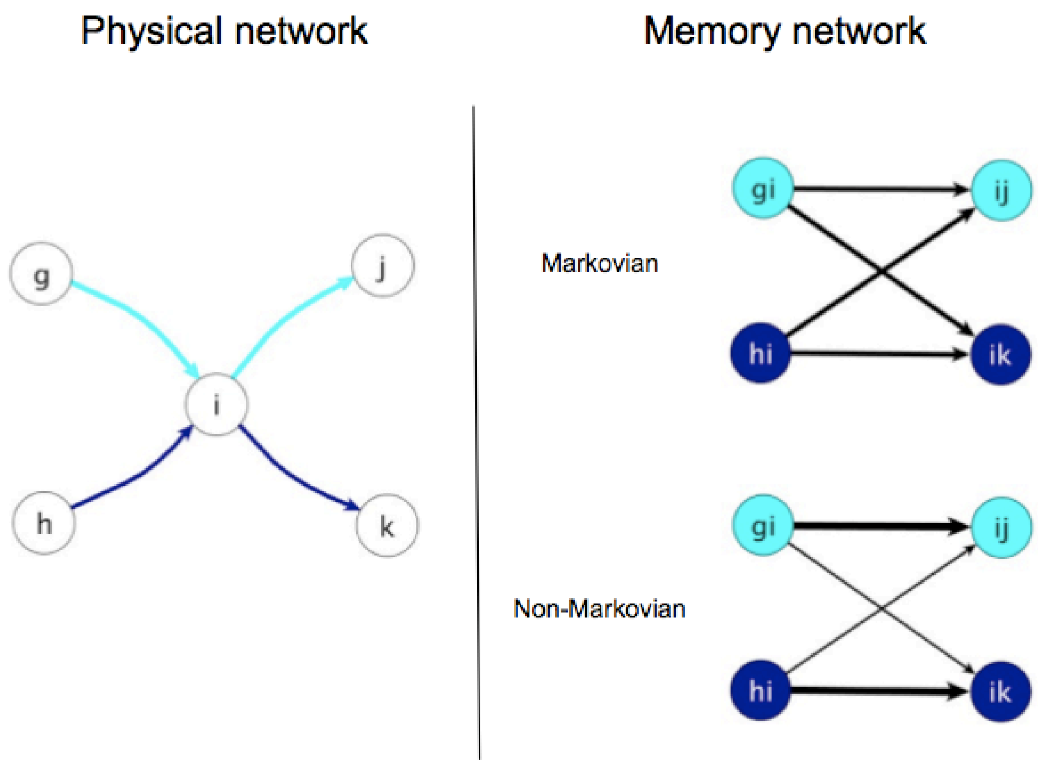
Where you go to depends on where you come from
Mathematics of pathways instead of edges



Networks with Memory, Martin Rosvall, Alcides V. Esquivel, Andrea Lancichinetti, Jevin D. West, Renaud Lambiotte, arXiv:1305.4807

Memory affects mixing time

Second-order Markov: transitions from directed edges to directed edge (memory node)
Memory may induce biases in the transition between memory nodes



Networks with Memory, Martin Rosvall, Alcides V. Esquivel, Andrea Lancichinetti, Jevin D. West, Renaud Lambiotte, arXiv:1305.4807
Slow-Down vs. Speed-Up of Information Diffusion in Non-Markovian Temporal Networks, I Scholtes, N Wider, R Pfitzner, A Garas, C Juan Tessone and F Schweitzer, arXiv:1307.4030

Memory affects mixing time

Does memory accelerate or slow-down diffusion?

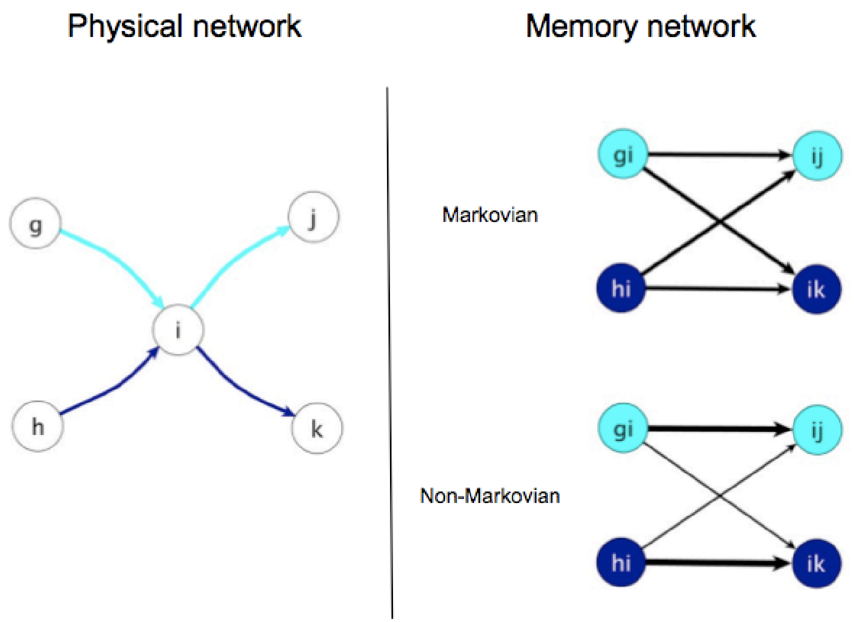
Scholtes et al. show numerically that the spectral gap can either increase or decrease in different real-life and artificial systems
Rosvall et al. show that modularity increases in empirical systems (=> slows down diffusion)

We look for an explicit expression for the effect of memory on the spectral gap (and hence on the mixing time)

Slow-Down vs. Speed-Up of Information Diffusion in Non-Markovian Temporal Networks, I Scholtes, N Wider, R Pfitzner, A Garas, C Juan Tessone and F Schweitzer, arXiv:1307.4030

Effect of Memory on the Dynamics of Random Walks on Networks, R Lambiotte, V Salnikov, M Rosvall, arXiv: 1401.0447

Memory affects mixing time



Random walk on the memory network

$$P(\beta;t+1) = \sum_{\alpha} P(\alpha;t)T_{\alpha\beta}$$

If the dynamics is memoryless, uniform transition:

$$T^{M}_{\alpha\beta} = \begin{cases} 1/k^{\text{out}}_{\alpha} & \text{for } \beta \in \sigma^{\text{out}}_{\alpha}, \\ 0 & \text{otherwise,} \end{cases}$$

(Left and right) eigenvectors of the spectral gap associated to the best bi-partition of the network (Fiedler)

$$\mathrm{uT} = \lambda_2 \mathrm{u}$$

Small deviation to the Markovian case and perturbation analysis:

$$T = T^{M} + \Delta T$$

Memory affects mixing time

$$\Delta\lambda_2 = -\frac{\sum_{\alpha\beta} u_\alpha^M \Delta T_{\alpha\beta} v_\beta^M}{\sum_\alpha u_\alpha^M v_\alpha^M}$$

Interplay between memory and the (dominant) bi-modular structures:

- if memory enhances flows inside communities => slowing down of diffusion

- if memory enhances flows across communities => acceleration of diffusion

Memory affects mixing time

Physical network

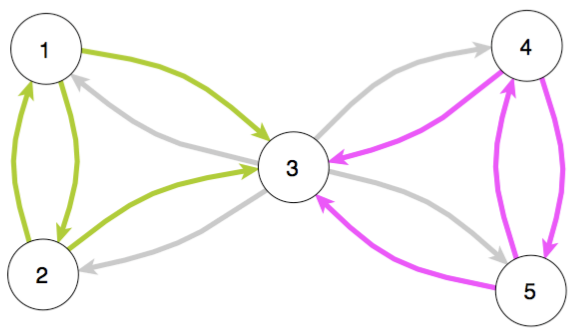


FIG. 2: Illustration of the bow tie network studied in detail in the main text. When the process is Markovian and that the transitions between memory nodes are uniform, the second dominant (left) eigenvector, of eigenvalue $1/2$, is represented by the color code: 1 for green memory nodes, -1 for pink ones, and 0 for grey ones.

Memory network

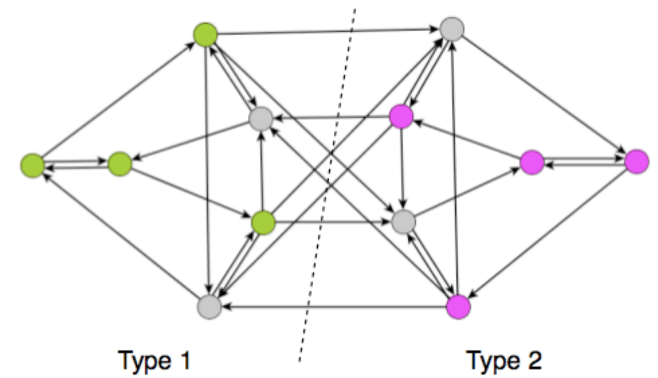
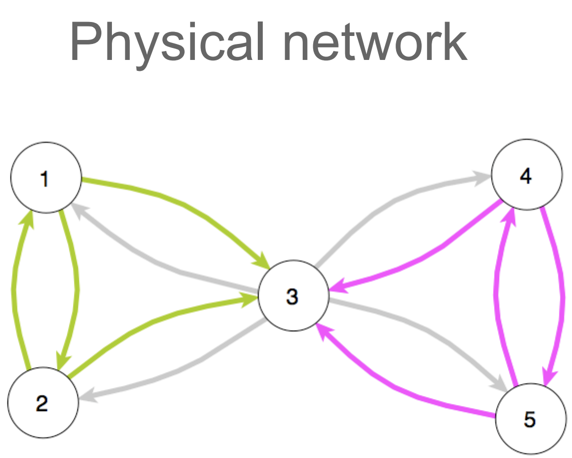
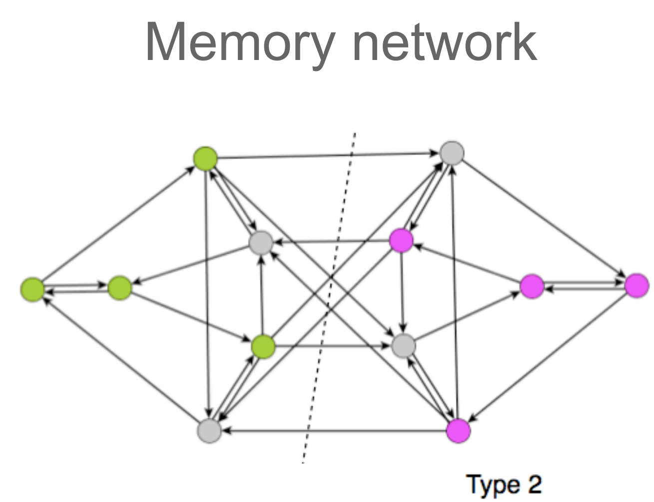


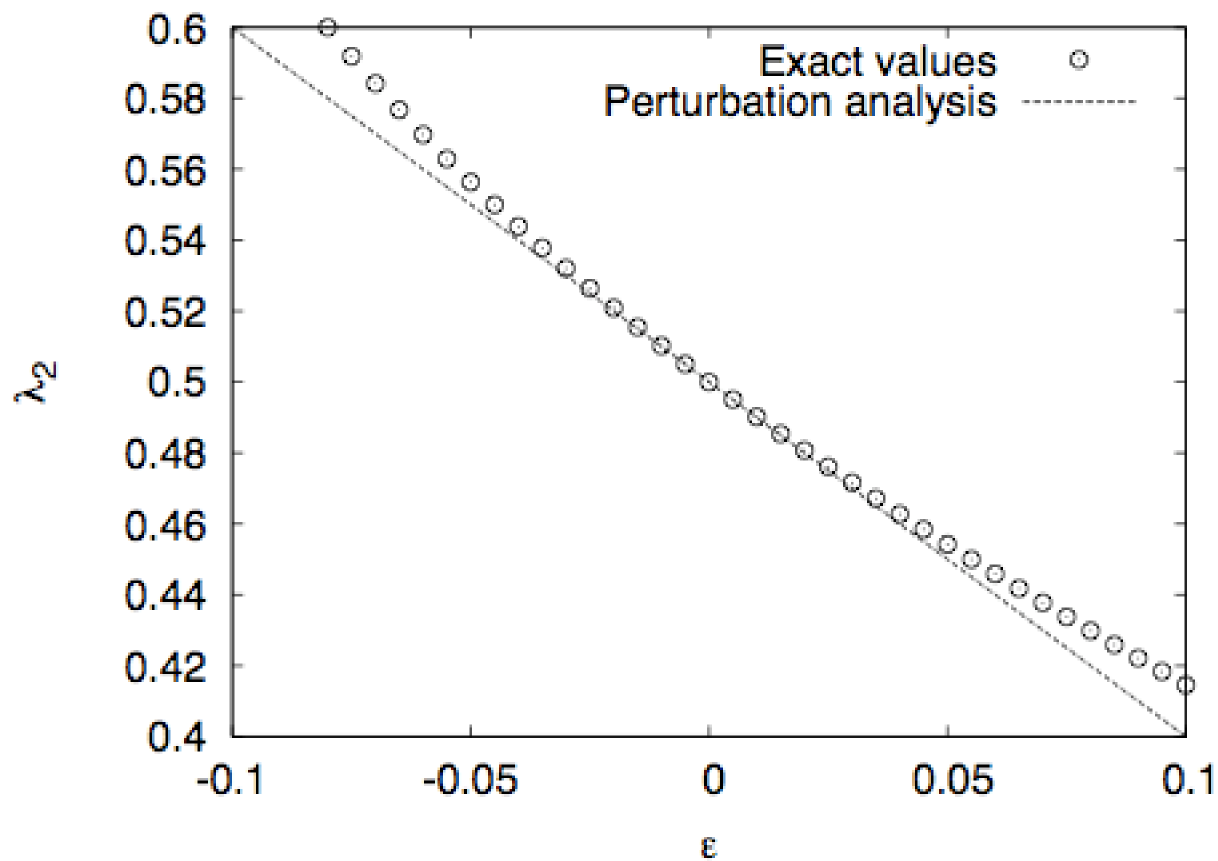
FIG. 3: Representation of the memory network associated to Fig. 2. The same color code has been used. The non-Markovian process is defined by partitioning the network into two groups, and by assigning different types of transitions across and within groups.

The non-Markovian dynamics is modelled as follows: memory nodes are partitioned into two groups. The weight of a transition between nodes of the same type (different types) is $1+\varepsilon$ $(1-\varepsilon)$.

Memory affects mixing time





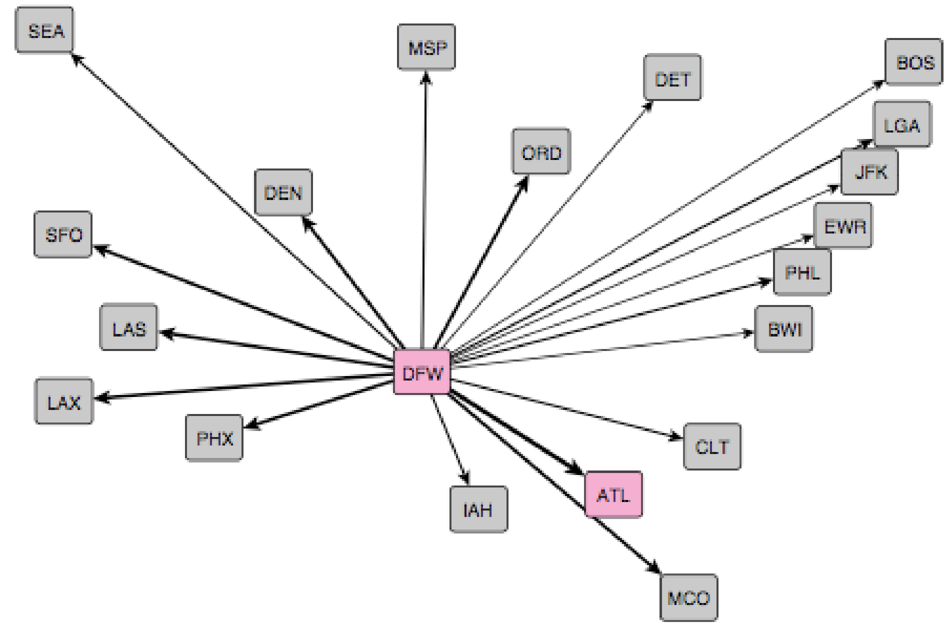


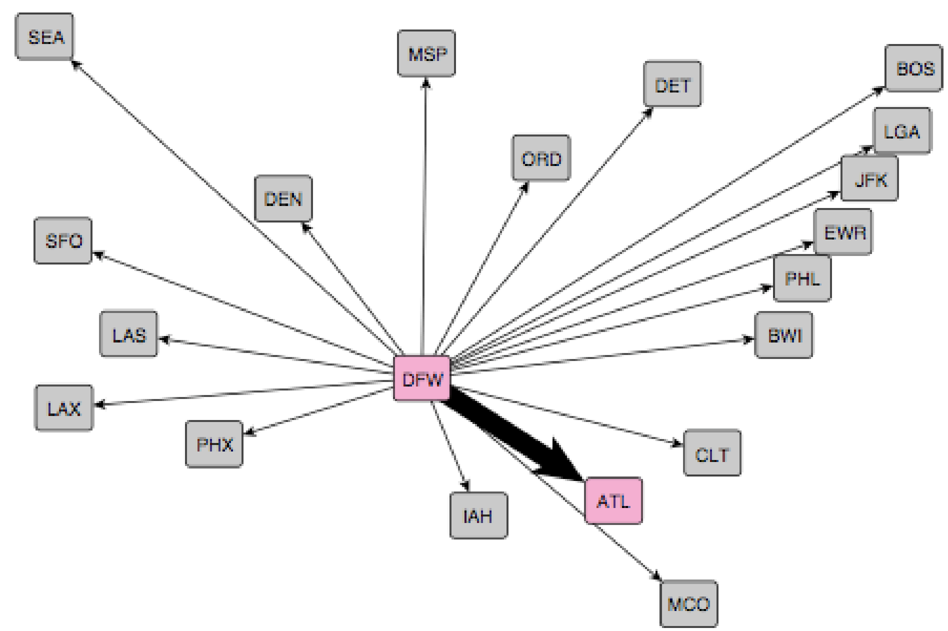
Memory affects mixing time

Airline pathways of passengers between US airports.

Modeled as a memory network, memory nodes (ATL, DFW) represent flight legs from airport ATL to DFW, and links represent connected flight legs, e.g. (ATL,DFW) -> (DFW,DEN) has a weight given by the number of passengers going to DEN from ATL, passing by DFW.

We restrict the scope to a subset of the top 20 airports in terms of traffic, in order to ensure that the memory network is strongly connected





Memory affects mixing time

Airline pathways of passengers between US airports.

Modeled as a memory network, memory nodes (ATL, DFW) represent flight legs from airport ATL to DFW, and links represent connected flight legs, e.g. (ATL,DFW) -> (DFW,DEN) has a weight given by the number of passengers going to DEN from ATL, passing by DFW.

We restrict the scope to a subset of the top 20 airports in terms of traffic, in order to ensure that the memory network is strongly connected

Because the airport network is weighted, in contrast with the previous example, the transition matrix of the Markovian process is now given by

$$T^{M}_{\alpha\beta} = \frac{w_{\beta}}{\sum_{\beta} w_{\beta}}$$

in order to provide an adequate baseline, and not uniform before. w_ \beta is the total number of passengers along flight leg β, independently on where they come from.

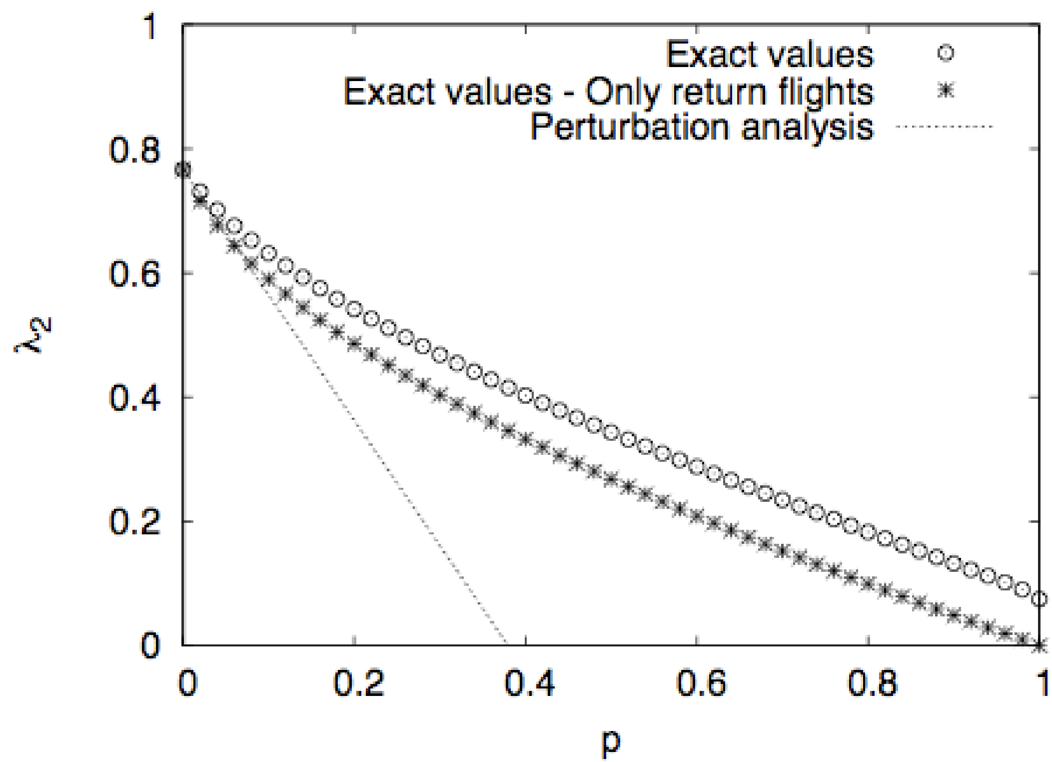
Memory affects mixing time

Random process driven by the tunable transition matrix

$$T^{(p)} = pT + (1 - p)T^M$$

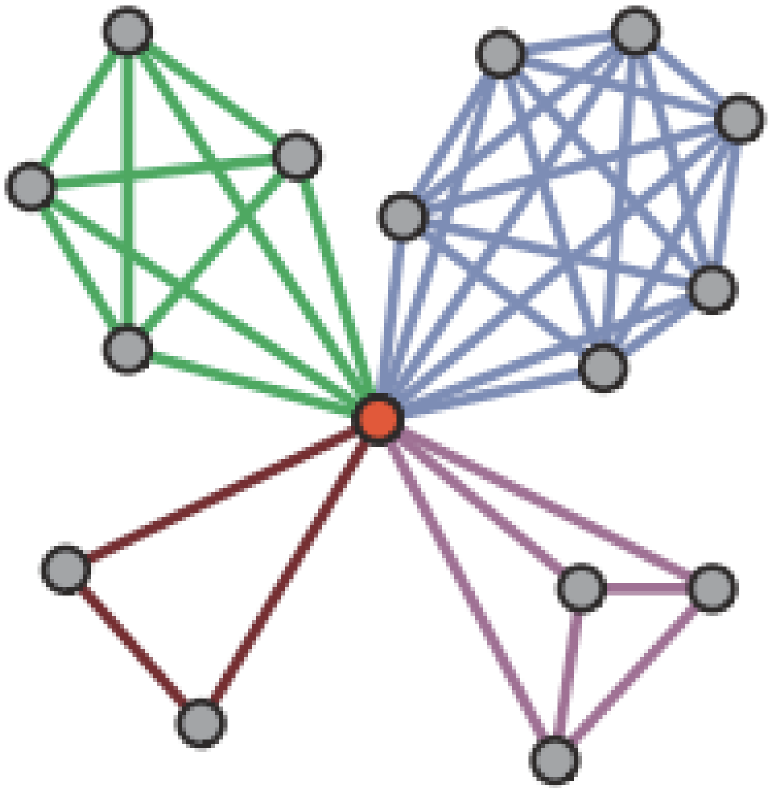
First-order Markov process with probability (1-p), and second-order Markov process with probability p.

This hybrid process models the diffusion of an item, e.g. a virus or a bank note, which travels with passengers and changes owner inside cities with probability p.



The effective size of the system is multiplied by 10 => the network is topologically small **but dynamically large.**

Diffusion in Multiplex Multiplex Networks



Overlapping community structure of social networks

Information spreads differently in different circles

Conclusion

Modeling step

Exact timings at which users meet are not considered: all that matters are fluxes between node (airports)

Real-world trajectories of the users are not considered: all that matters are fluxes between node (airports)

Now that such data is available: Accuracy of the network paradigm? Other paradigms?

Generalized Master Equations for Non-Poissonian Dynamics on Networks, Till Hoffmann, Mason Porter and R.L., Physical Review E 2012
Burstiness and spreading on temporal networks, R.L., L. Tabourier and J.C. Delvenne, EPJB 2013
Networks with Memory, Martin Rosvall, Alcides V. Esquivel, Andrea Lancichinetti, Jevin D. West, Renaud Lambiotte, arXiv:1305.4807
Bottlenecks, burstiness, and fat tails regulate mixing times of non-Poissonian random walks, J.-C. Delvenne, Renaud Lambiotte and L. E. C. Rocha, arXiv:1309.4155
Effect of Memory on the Dynamics of Random Walks on Networks, R Lambiotte, V Salnikov, M Rosvall, arXiv: 1401.0447

Thanks to:

Till Hoffmann and Mason Porter (Oxford): temporal networks

Martin Rosvall (Umea) and Andrea Lancichinetti (Northwestern): pathways

J.C. Delvenne (Louvain), Luis Rocha (Namur) and Lionel Tabourier (Namur): diffusion

OPEN positions with JC Delvenne!

Financial support of Optimizr, IAP DYSCO from Belspo, Région Wallonne and FNRS

More info: http://xn.unamur.be