From a static to a dynamic analysis of complex networks

Lionel Tabourier

LIP6 (Sorbonne Université / CNRS)

Soutenance de HDR

Outline

About complex networks analysis

What are complex networks?

Extracted from real-world data
ex: computer networks, social networks, biological networks, ...

- agents interacting, no "central brain"
- $\Rightarrow$ emerges from local interactions

Studied as graphs

- system common features visible with graphs
  ex: heterogeneous degree distribution, high clustering, ...
- similar methods and algorithms

Accounting for temporality

About complex networks analysis

What are complex networks?

Extracted from real-world data
ex: computer networks, social networks, biological networks, . . .

- agents interacting, no "central brain"
- $\Rightarrow$ emerges from local interactions

Studied as graphs

- system common features visible with graphs
  ex: heterogeneous degree distribution, high clustering, . . .
- similar methods and algorithms

Accounting for temporality

About complex networks analysis

What are complex networks?

Extracted from real-world data

ex: computer networks, social networks, biological networks, . . .

- agents interacting, no "central brain"
- $\Rightarrow$ emerges from local interactions

Studied as graphs

- system common features visible with graphs
  ex: heterogeneous degree distribution, high clustering, . . .
- similar methods and algorithms

Accounting for temporality

Scientific interest

Attractive because ...

- Variety of objects
- Create bridges between different topics

Getting the “big picture”

... and in retrospect

- Variety of communities to learn about
- Highly active research fields

⇒ Stimulating but demanding

Focus on three main axes of research

Scientific interest

Attractive because ...

- Variety of objects
- Create bridges between different topics

Getting the "big picture"

... and in retrospect

- Variety of communities to learn about
- Highly active research fields

⇒ Stimulating but demanding

Focus on three main axes of research

Scientific interest

Attractive because . . .

- Variety of objects
- Create bridges between different topics

Getting the "big picture"

. . . and in retrospect

- Variety of communities to learn about
- Highly active research fields

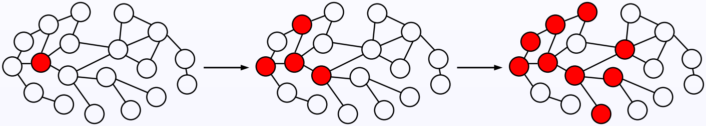
$\Rightarrow$ Stimulating but demanding

Focus on three main axes of research

Describing spreading processes

Dynamical processes on complex networks

Synchronization, opinion dynamics, routing, **spreading**, ...



Models which represent:

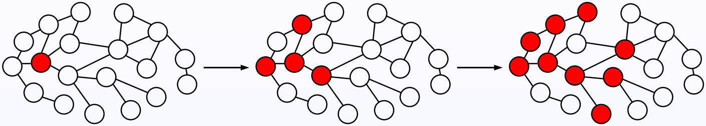
- an infection propagating in a population
- a packet spreading in a computer network
- ...

Also a means to investigate the dynamical structure

Describing spreading processes

Dynamical processes on complex networks

Synchronization, opinion dynamics, routing, **spreading**, . . .



Models which represent:

- an infection propagating in a population
- a packet spreading in a computer network
- . . .

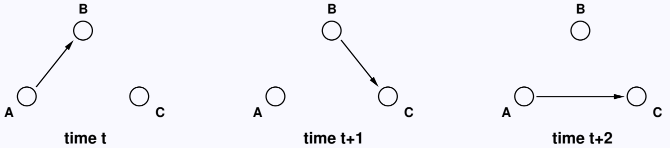
Also a means to investigate the dynamical structure

Predicting and recovering links

Understanding the microdynamics

Unknown mechanisms ⇒ learn from data

ex: triadic closure



Specificities of the problem

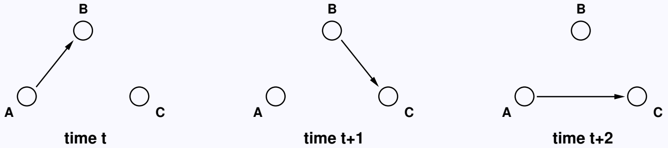
- tuning number of predictions
- on large graphs
- how to incorporate temporal information?

Predicting and recovering links

Understanding the microdynamics

Unknown mechanisms ⇒ learn from data

ex: triadic closure



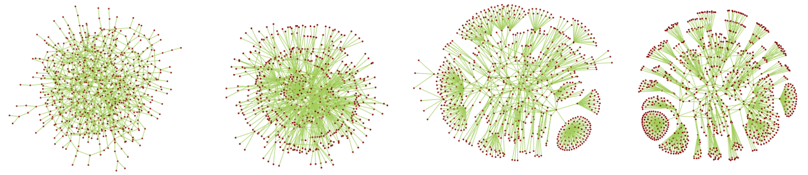
Specificities of the problem

- tuning number of predictions
- on large graphs
- how to incorporate temporal information?

Model the interaction structure and dynamics

What is modeling here?

Generating artificial structures resembling real data



⇒ allows to uncover possible explanations

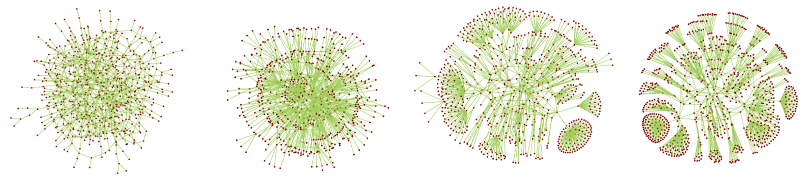
Challenges

- realistic structural constraints on graphs
- towards modeling temporal networks

Model the interaction structure and dynamics

What is modeling here?

Generating artificial structures resembling real data



$\Rightarrow$ allows to uncover possible explanations

Challenges

- realistic structural constraints on graphs
- towards modeling temporal networks

Presentation outline

- **Describing spreading dynamics**
- Predicting links
- **Modeling the structure and dynamics**

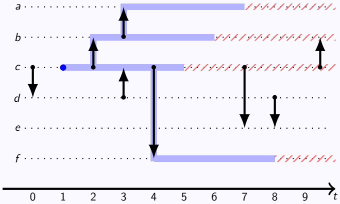
Highlight the connexion between these topics

Outline

Spreading cascade in a temporal network

Model analyzed

SIR model: infection probability 1, deterministic recovery after τ



Spreading cascade (but many other names)

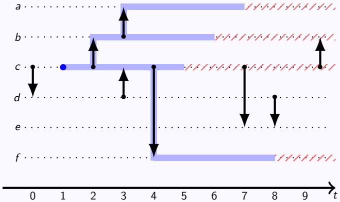
Advantages:

- simplicity of the model
- allows to tune the time scale
- probe of the temporal network (motif)

Spreading cascade in a temporal network

Model analyzed

SIR model: infection probability 1, deterministic recovery after τ



Spreading cascade (but many other names)

Advantages:

- simplicity of the model
- allows to tune the time scale
- probe of the temporal network (motif)

Role of in/out-degree correlations

Peruani, Tabourier, 2011

Phonecall dataset

$\sim$14 millions phone calls between $\sim$1 million users during 1 month
European mobile phone provider

Strongly asymmetric roles (super-spreaders, super-receivers)

Role of in/out-degree correlations

Peruani, Tabourier, 2011

Phonecall dataset

$\sim$14 millions phone calls between $\sim$1 million users during 1 month
European mobile phone provider

Strongly asymmetric roles (super-spreaders, super-receivers)

Critical recovery time τ_c

τ such that cascade size diverges

$$\tau_c = \frac{< k_{out} >}{< \rho >< k_{out} \cdot k_{in} >}$$

with some assumptions (no cycles)

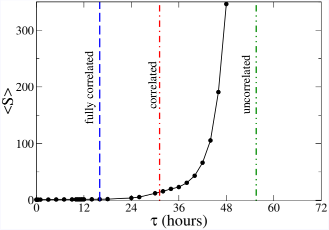
Role of in/out-degree correlations

Peruani, Tabourier, 2011

Experimental observations

- real (k_{in}, k_{out}) distribution: 32 hours
- fully-correlated data $k_{in} = k_{out}$: 14 hours
- uncorrelated data: 55 hours

$\Rightarrow$ significant differences



underestimation of the threshold

Role of individual temporal patterns

Measuring the characteristics of the cascades

- cascade = tree $\Rightarrow$ size (σ), depth (δ)
- abundance of each type of cascades

To what can we compare these measures?

Comparing to null models

- *time mixing model*: random shuffle of all time labels
- *correlation mixing model*: random shuffle of time labels for interactions originating from same source

Role of individual temporal patterns

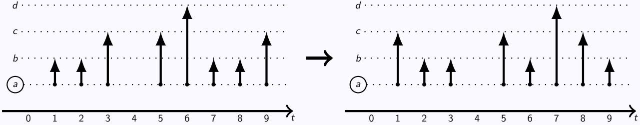
Measuring the characteristics of the cascades

- cascade = tree ⇒ size (σ), depth (δ)
- abundance of each type of cascades

To what can we compare these measures?

Comparing to null models

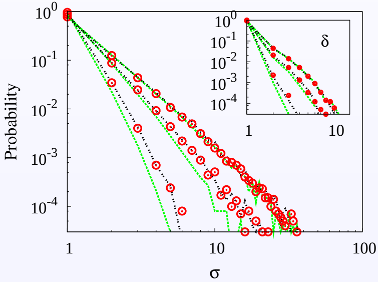
- *time mixing model*: random shuffle of all time labels
- *correlation mixing model*: random shuffle of time labels for interactions originating from same source



Experimental results on phonecall data

Tabourier, Stoica, Peruani 2012

Distributions of σ (size) and δ (depth) for $\tau = 30$mins, 3hrs, 12hrs



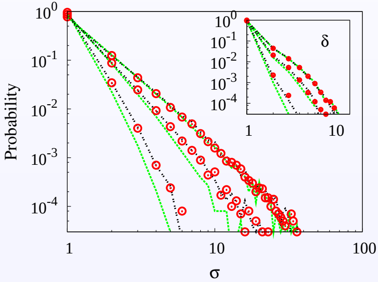
time mixing model
corr. mixing model
real data

time mixing model breaks bursty temporal activity patterns
$\Rightarrow$ Bursty patterns accelerates spreading at short timescales

Experimental results on phonecall data

Tabourier, Stoica, Peruani 2012

Distributions of σ (size) and δ (depth) for τ = 30mins, 3hrs, 12hrs



time mixing model
corr. mixing model
real data

time mixing model breaks bursty temporal activity patterns
⇒ Bursty patterns accelerates spreading at short timescales

Experimental results on phonecall data

Tabourier, Stoica, Peruani 2012

About the *correlation mixing model*:

- σ and δ insufficient to distinguish real data from *cmm*
- but comparing cascades and other patterns with *cmm*
  $\rightarrow$ *cmm* underestimates short cyclic patterns

cmm breaks correlation between temporal activity and structure
$\Rightarrow$ correlations create a trapping effect

Consistent with *"while bursts hinder propagation at large scales, conversations favor local rapid cascades"* Miritello et al., 2011

Experimental results on phonecall data

Tabourier, Stoica, Peruani 2012

About the *correlation mixing model*:

- σ and δ insufficient to distinguish real data from *cmm*
- but comparing cascades and other patterns with *cmm*
  $\rightarrow$ *cmm* underestimates short cyclic patterns

cmm breaks correlation between temporal activity and structure
$\Rightarrow$ correlations create a trapping effect

Consistent with *"while bursts hinder propagation at large scales, conversations favor local rapid cascades"* *Miritello et al., 2011*

Conclusion and prospect

- Schematic picture of spreading:
  rapid cascades at small scales, slower on larger scales
- Due to individual activity patterns, correlations between structure and temporality
- Understanding one step further how the temporal network structure affects spreading phenomena
  $\rightarrow$ new null models for temporal networks

Outline

Motivation

To explain the observed structure

- generate a structure that resembles real data
- unbiased interpretation assumes uniform generation

With graphs

- either biased generation model
  ex: Barabási-Albert, Watts-Strogatz, . . .
- or relatively simple properties
  ex: Erdős-Rényi, configuration model, . . .

Versatile method to generate **uniformly** graphs
with complex constraints

Motivation

To explain the observed structure

- generate a structure that resembles real data
- unbiased interpretation assumes uniform generation

With graphs

- either biased generation model
  ex: Barabási-Albert, Watts-Strogatz, . . .
- or relatively simple properties
  ex: Erdős-Rényi, configuration model, . . .

Versatile method to generate **uniformly** graphs
with complex constraints

Motivation

To explain the observed structure

- generate a structure that resembles real data
- unbiased interpretation assumes uniform generation

With graphs

- either biased generation model
  ex: Barabási-Albert, Watts-Strogatz, ...
- or relatively simple properties
  ex: Erdős-Rényi, configuration model, ...

Versatile method to generate **uniformly** graphs
with complex constraints

Switching methods principle

Standard switch method

ex: generate a random graph with a fixed degree distribution

- start from a graph of the set
- iterate: select two edges randomly, exchange the ends
  - if new graph $\in$ set, continue
  - if new graph $\notin$ set, exchange back, continue

After enough iterations $\rightarrow$ random element of the set

More complex constraints

Problem: standard method does not guarantee ergodicity
$\rightarrow$ biased generation process

Switching methods principle

Standard switch method

ex: generate a random graph with a fixed degree distribution

- start from a graph of the set
- iterate: select two edges randomly, exchange the ends
  - if new graph $\in$ set, continue
  - if new graph $\notin$ set, exchange back, continue

After enough iterations $\rightarrow$ random element of the set

More complex constraints

Problem: standard method does not guarantee ergodicity

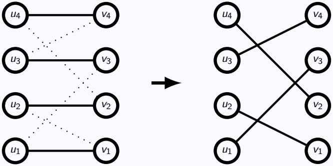
$\rightarrow$ biased generation process

Obtaining ergodicity

Tabourier, Roth, Cointet, 2011

k-switch method

- start from a graph of the set
- iterate: select k edges randomly, permute the ends randomly
  - if new graph $\in$ set, continue
  - if new graph $\notin$ set, exchange back, continue



Experimentally
low values of k sufficient with most constraints

Illustration

Tabourier, Cointet, Roth 2017

Scientific coauthoring networks

Scientists publishing papers together
underlying bipartite structure
$\rightarrow$ interested in the structure **projection** on authors

- Usual models for social event-based networks:
  - standard configuration model M
  - bipartite configuration model B
- More complex model:
  - bipartite + monopartite configuration model on authors MB

MB demands k-switch generation

Illustration

Tabourier, Cointet, Roth 2017

Scientific coauthoring networks

Scientists publishing papers together
underlying bipartite structure
$\rightarrow$ interested in the structure **projection** on authors

- Usual models for social event-based networks:
  - standard configuration model M
  - bipartite configuration model B
- More complex model:
  - bipartite + monopartite configuration model on authors MB

MB demands k-switch generation

Illustration

Tabourier, Cointet, Roth 2017

Scientific coauthoring networks

Scientists publishing papers together
underlying bipartite structure
$\rightarrow$ interested in the structure **projection** on authors

- Usual models for social event-based networks:
  - standard configuration model M
  - bipartite configuration model B
- More complex model:
  - bipartite + monopartite configuration model on authors MB

MB demands k-switch generation

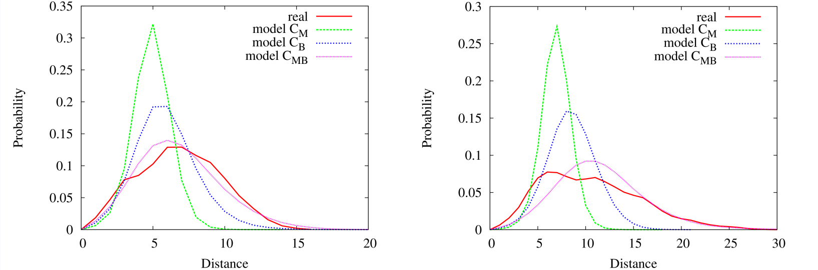
Illustration

Tabourier, Cointet, Roth 2017

Experimental results

Measure: distance distribution in the models vs real data

datasets: archeology in Britain, in Europe



⇒ more accurate model for coauthoring network

Generalization to temporal networks: principle

Motivation

Most models focused on aggregated properties *Gauvin et al., 2018*

$\rightarrow$ alternate models preserving structuro-temporal properties

ex: fixed number of temporal motifs

Applying the k-switch method on temporal networks

time label used as an **attribute constraint**

$\Rightarrow$ same randomization technique

Generalization to temporal networks: principle

Motivation

Most models focused on aggregated properties *Gauvin et al., 2018*
$\rightarrow$ alternate models preserving structuro-temporal properties
ex: fixed number of temporal motifs

Applying the k-switch method on temporal networks

time label used as an **attribute constraint**
$\Rightarrow$ same randomization technique

Generalization to temporal networks: challenges

- **technical problem:**
  more elaborate the constraint $\Rightarrow$ slower generation process

- **interpretation problem:**
  how to have more flexibility on the temporal granularity?
  example of strict vs approximate simultaneity

Outline

Conclusion

Common thread

- describing the temporal network
- predicting links
- modeling the structure of interactions
- on specific applications

serve to identify dynamical rules of the network
$\rightarrow$ **reconstruct the general evolution of the system**

Thank you for your attention !

My collaborators