Dynamics of and on complex networks

Clémence Magnien
Complex Networks team
http://complexnetworks.fr

LIP6 laboratory (CNRS, Université Pierre et Marie Curie)

October 15th, 2010

Complex Networks

computer science: web, internet, peer-to-peer, usages, etc.

social sciences: friendships, communications, collaborations, exchanges, economics, etc.

biology: brain, genes, proteins, ecosystems, etc.

linguistics: synonymy, co-occurrences, etc.

transportation: road, air, electricity, etc

etc, etc

networks

very different contexts

no mathematical definition

Complex Networks

most complex networks share

non-trivial properties [WS98]

common properties:

- low average distance (*small-world*)
- heterogeneous degrees (power-law, *scale-free*)
- low global density *vs* high local density (triangles, *clustering*)

Large body of works on complex networks
$\longrightarrow$ acknowledged tools and notions

Complex Networks

most complex networks share

non-trivial properties [WS98]

common properties:

- low average distance (*small-world*)
- heterogeneous degrees (power-law, *scale-free*)
- low global density *vs* high local density (triangles, *clustering*)

Large body of works on complex networks
$\longrightarrow$ acknowledged tools and notions

What about dynamics?

(All) complex networks are dynamic

Appearance and disappearance with time of:

- nodes
- links

Few or no established notions for the study of these dynamics

What about dynamics?

(All) complex networks are dynamic

Appearance and disappearance with time of:

- nodes
- links

Few or no established notions for the study of these dynamics

What about dynamics (bis)?

Dynamics on networks:

Spreading phenomena (epidemics, rumours, . . .)
Nodes' states change

Team activities – Dynamics

Measurement

Acquiring informations on the dynamics
(of the network or on the network)

Description

Statistical properties (life duration, correlations, . . .)
Structural properties (communities)

Structural dynamics :
normal vs abnormal
Spreading phenomena :
correlations with community structures

Outline

A radar for the internet

[Latapy, Magnien, Ouédraogo 2008]

Internet topology

Routers, IP-level links

Mesurement: `traceroute` tool

- Partial (and biased) information
- Periodical measurements: costly $\longrightarrow$ low frequency

Orthogonal approach: ego-centred view

- a monitor
- several destinations
- well-defined object
- high frequency: radar

A radar for the internet

[Latapy, Magnien, Ouédraogo 2008]

Internet topology

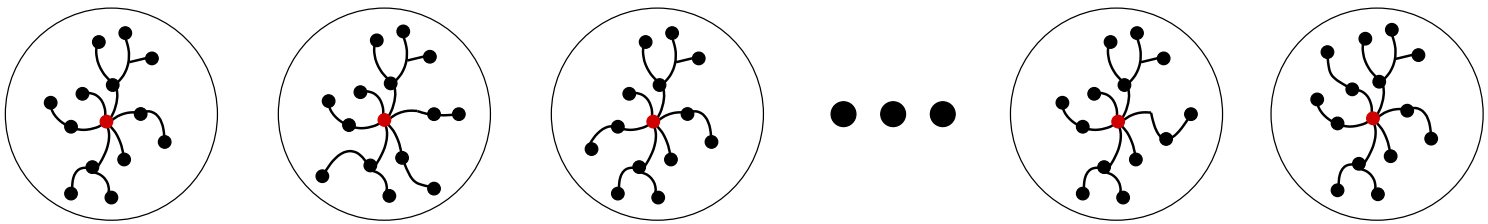
Routers, IP-level links

Mesurement: `traceroute` tool

- Partial (and biased) information
- Periodical measurements: costly $\longrightarrow$ low frequency

Orthogonal approach: ego-centred view

- a monitor
- several destinations
- well-defined object
- high frequency: radar



Tool and measurements

- Design of a dedicated tool: `tracetree`
- Calibration

  $\longrightarrow \sim 100$ rounds / day

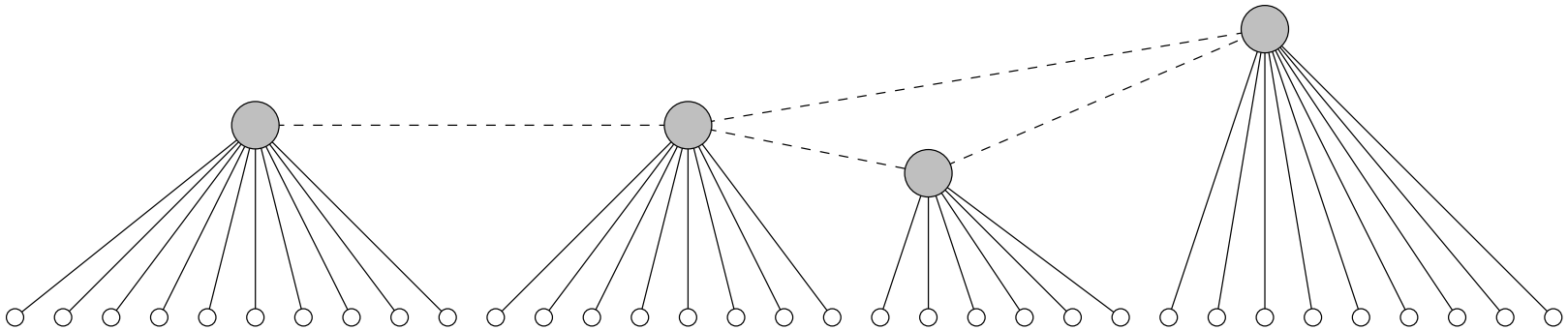
Measurements

- from > 100 monitors (planetlab, . . .)
- for several months

Tool and data are publicly available

eDonkey P2P system

[Allali, Latapy, Magnien 2009]



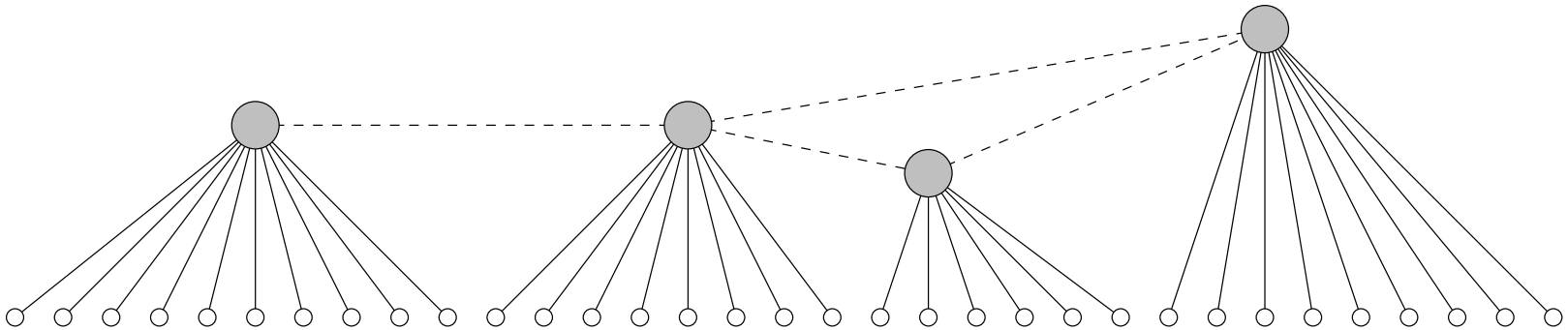
peer		**server**
keywords	$\longrightarrow$	
	$\longleftarrow$	filelist
file	$\longrightarrow$	
	$\longleftarrow$	providers

Measurements

- server
- client, honeypot

eDonkey P2P system

[Allali, Latapy, Magnien 2009]



peer		**server**
keywords	$\longrightarrow$	
	$\longleftarrow$	filelist
file	$\longrightarrow$	
	$\longleftarrow$	providers

Measurements

- server
- client, honeypot

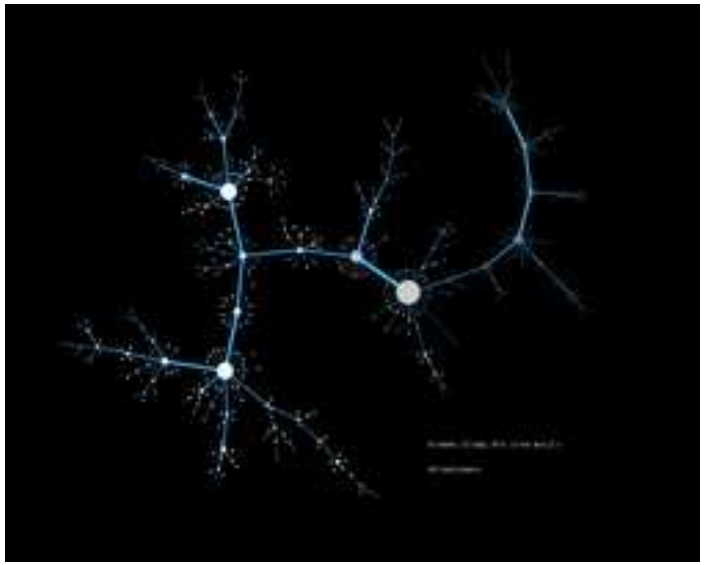
Spreading phenomenon – Happy flu experiment

[Friggeri, Cointet, Latapy 2010]

Happy flu: dedicated experiment

web applet copied from page to page
+ centralised data (copies and sightings)

2 months measurements $\longrightarrow$ 492 participants, $\sim$ 98 000 viewers



Heterogeneous influence

Measurement – summary

Measurements of several networks

Publicly available

Massive datasets

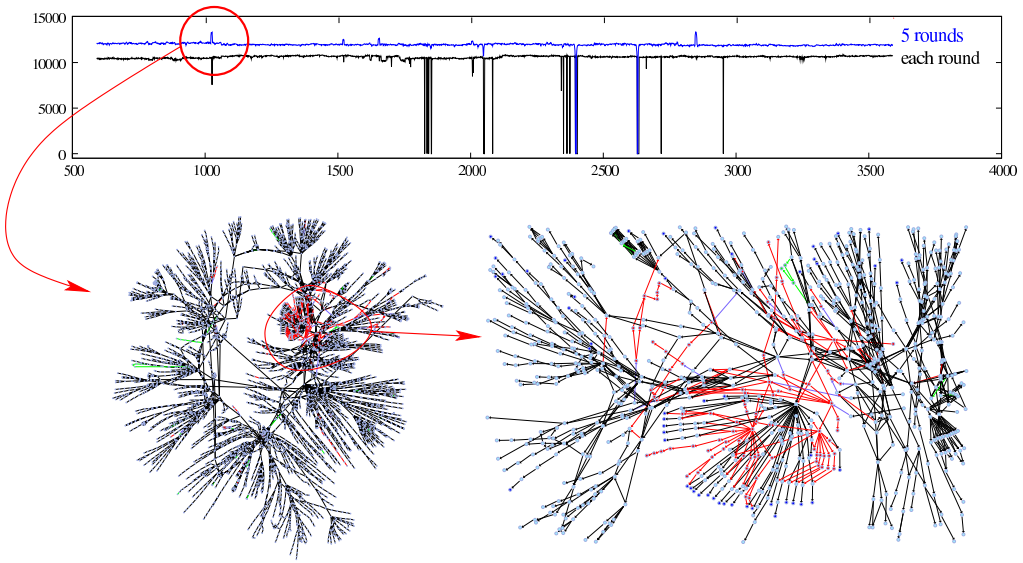
(size of network | measurement duration | frequency)

Outline

Event detection

[Hamzaoui, Latapy, Magnien 2010]

Radar measurements: identify abnormal events



Event detection

Automatic methods

Identify homogeneous statistics with outliers

Interpretation

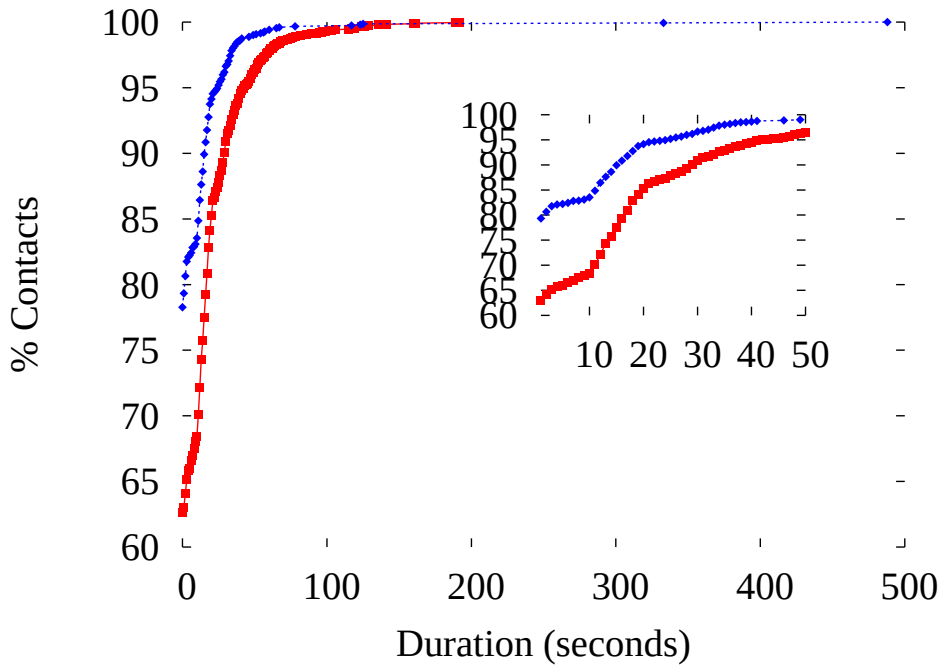
Need to validate with other methods/observations

$\longrightarrow$ Apply to other datasets

Characterise expected behaviour

Contact duration in a population.

- how much time do two people stay physically close?
- do different nodes have different contact duration patterns?



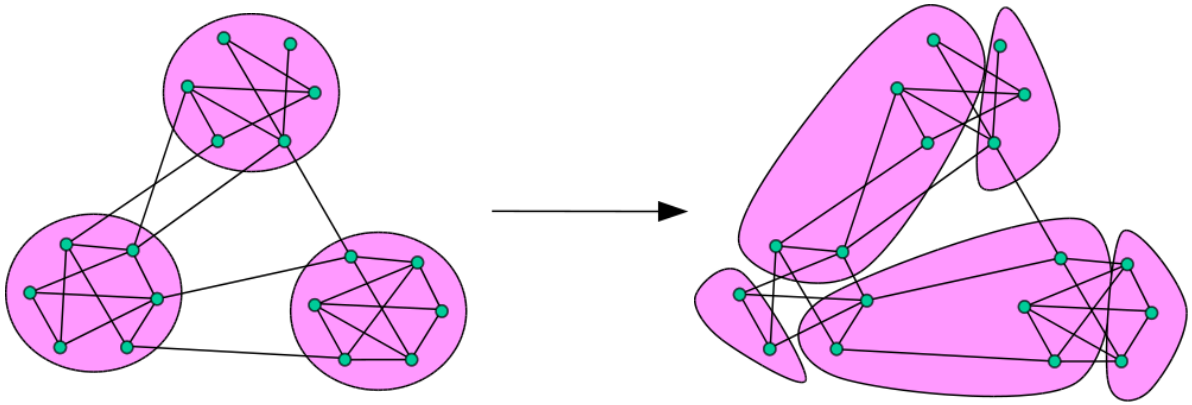
Blue node : longer maximal duration

Red node : longer contacts on average

Dynamic Communities

[Aynaud, Guillaume 2010]

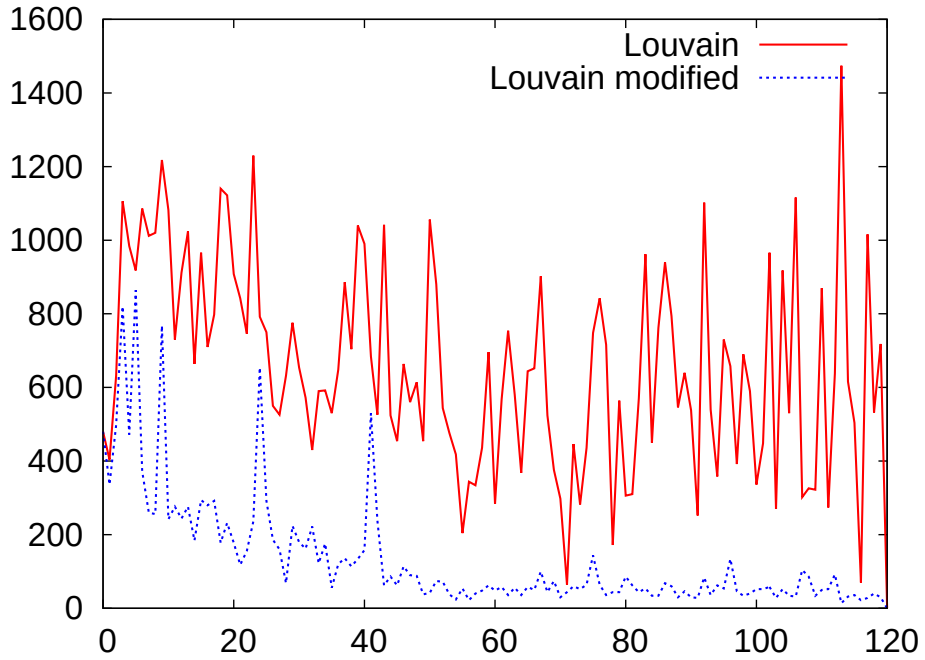
Community : group of nodes with many links inside, few outside



Detection: Louvain method
Problem: existing detection algorithms are unstable
(one change $\rightarrow$ very different partitions)

Dynamic Communities

Stabilise the algorithm: start from partition of previous time step (instead of each node in a different partition)



Links with event detection

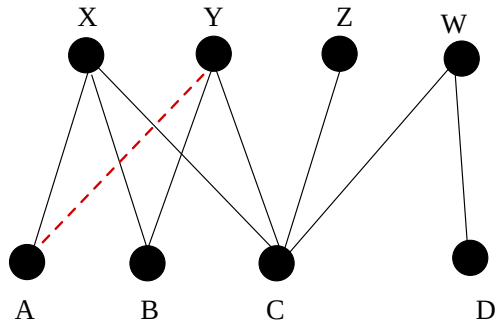
Link prediction

Bipartite networks

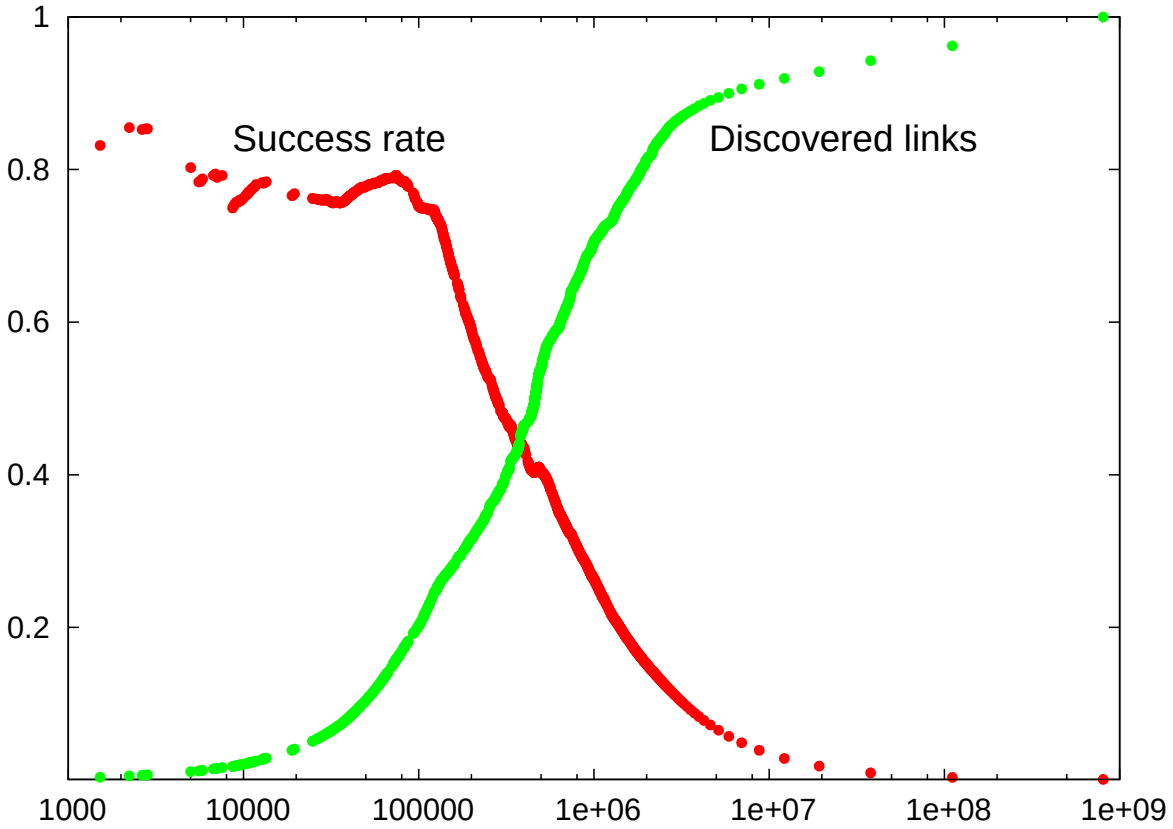
- P2P users/files
- Buyers/products
- ...

Notion of similarity between users

Based on common neighbourhood
Predict links which strengthen similarity



Prediction

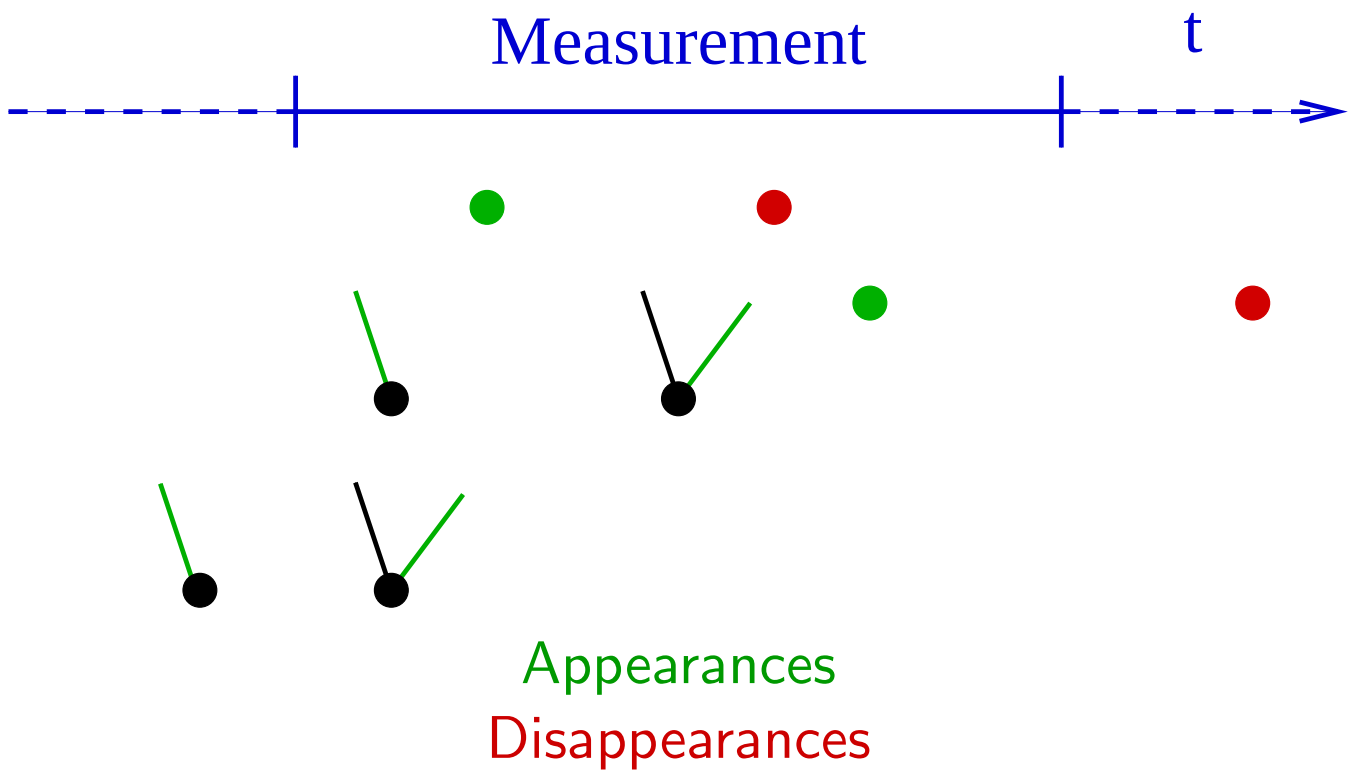


- Influence of learning / prediction period
- Different notions of similarity

Bias in dynamic measurements

[Benamara, Magnien 2010]

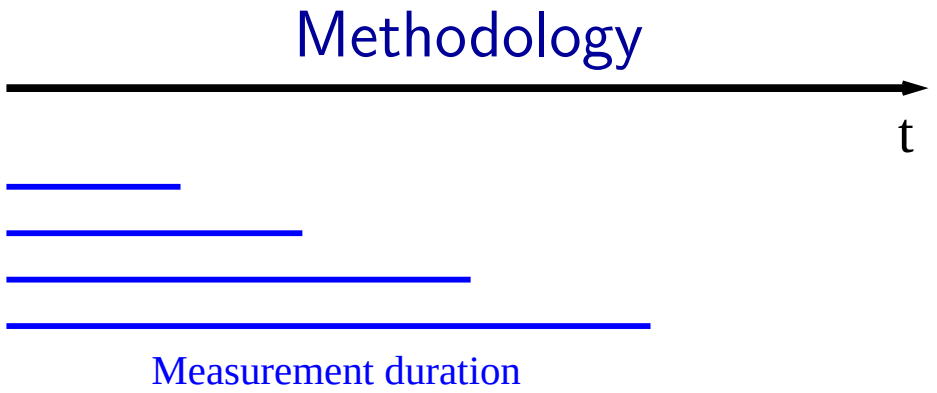
Finite measurement period $\longrightarrow$ unobserved events



Bias in the estimation of dynamic properties

Longer measurement $\longrightarrow$ smaller bias

Bias in dynamic measurements

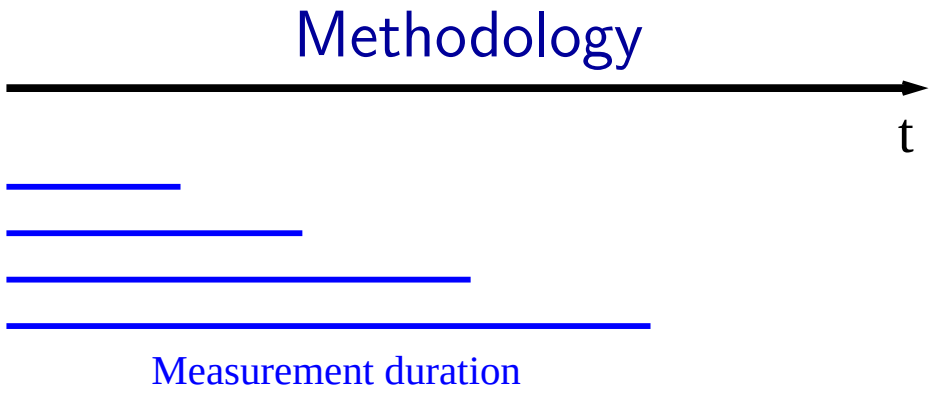


Evolution as a function of measurement length

- Fluctuations $\Longrightarrow$ no conclusion
- Stable $\Longrightarrow$ independant of measurement length (a priori).

Distinction between different behaviours
(normal vs abnormal)

Bias in dynamic measurements

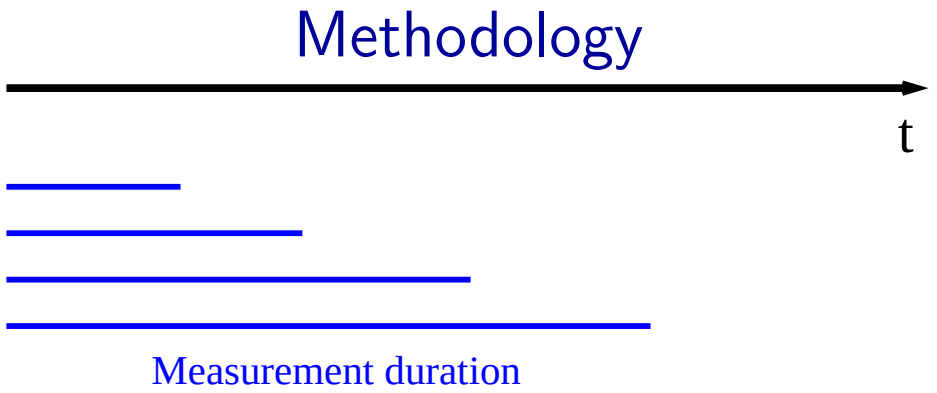


Evolution as a function of measurement length

- Fluctuations $\Longrightarrow$ no conclusion
- Stable $\Longrightarrow$ independant of measurement length (a priori).

Distinction between different behaviours
(normal vs abnormal)

Bias in dynamic measurements



Evolution as a function of measurement length

- Fluctuations $\Longrightarrow$ no conclusion
- Stable $\Longrightarrow$ independant of measurement length (a priori).

Distinction between different behaviours
(normal vs abnormal)

Spreading phenomena

Spreading in blogs

A post citing another post
$\longrightarrow$ information spreading

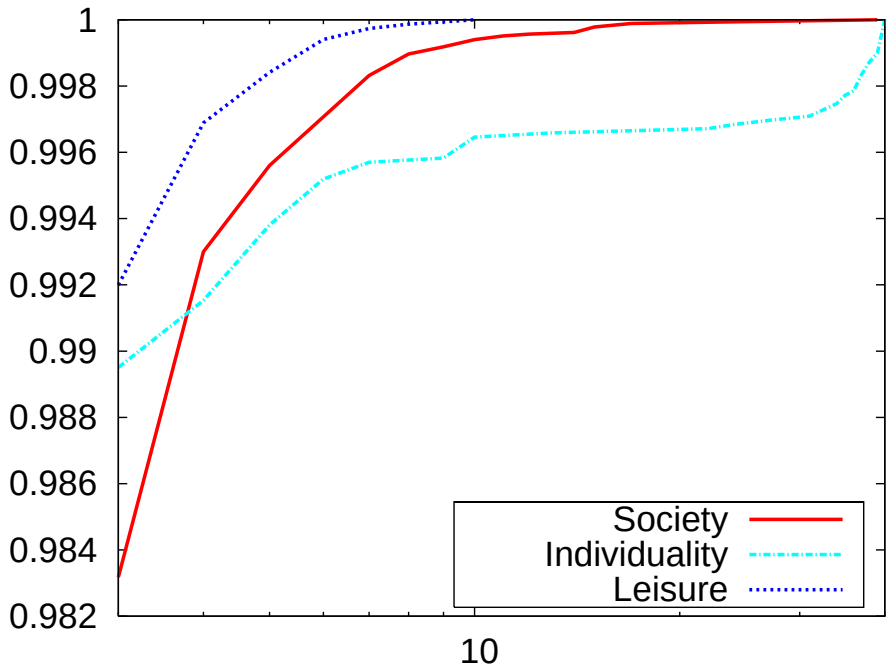
File spreading in a P2P system

Massive data
Large cascades

- Links between spreading and communities
  (structural, thematic)
- Confrontation with classical hypotheses

Spreading phenomena and thematic communities

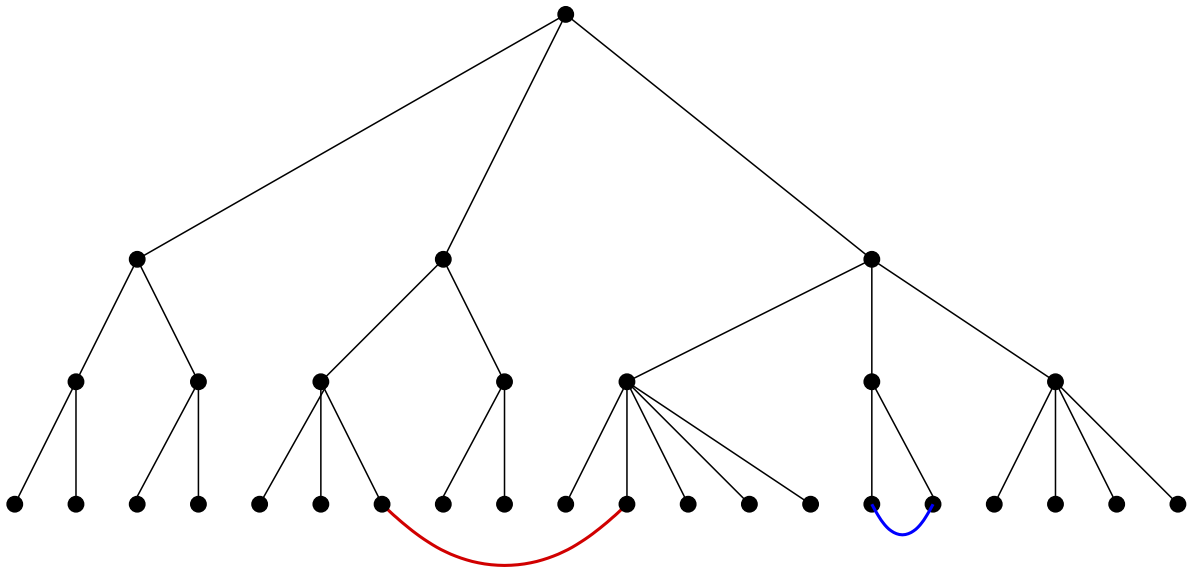
Popularity of different thematic communities



$\longrightarrow$ impact of thematic communities on spreading

Spreading phenomena and structural communities

Hierarchical community structure



Spreading from node to node:

how far in the community structure?

Conclusion

dynamics of and on networks
measurement and description

Our specificities:

- studies based on real-world cases
- design generic notions/methods

Questions

- events vs normal dynamics
- normal vs abnormal behaviour
- bias induced by measurement

- standard notions and tools for description
- modelling / formalisms