10. Dynamic networks driven by human mobility

Fabien Tarissan

CNRS

École Normale Supérieure Paris-Saclay

Plan

1. Mobile Networks
   - Mobile Networks
   - Capacity
   - Different Protocols
   - Some empirical results on mobility
2. Empirical study
   - Markovian models
   - Dataset
   - Methodology
   - Analyzing real traces
   - Impact of the markovian model

Mobile Networks

Dynamical aspect of networks

Motivation

- Development of wireless devices
- A lot of new open dataset
- Dynamics ON and **OF** the network
- New structural properties
- Redefining usual metrics (graphs)

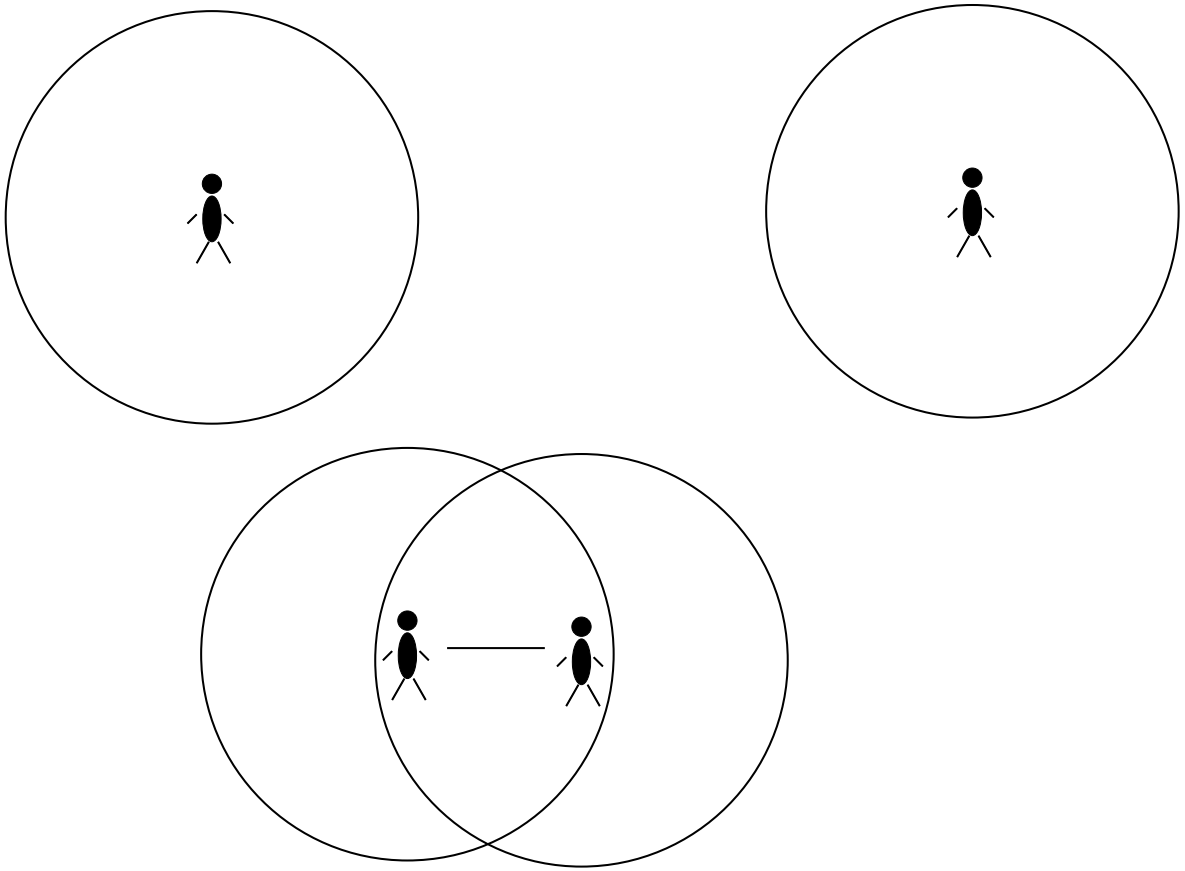
Issues

- How acquire knowledge from this object ? **(measure)**
- Which notable properties ? **(analyze)**
- *Which models best capture those properties ?* **(modelling)**

Object under study

Human-contact networks

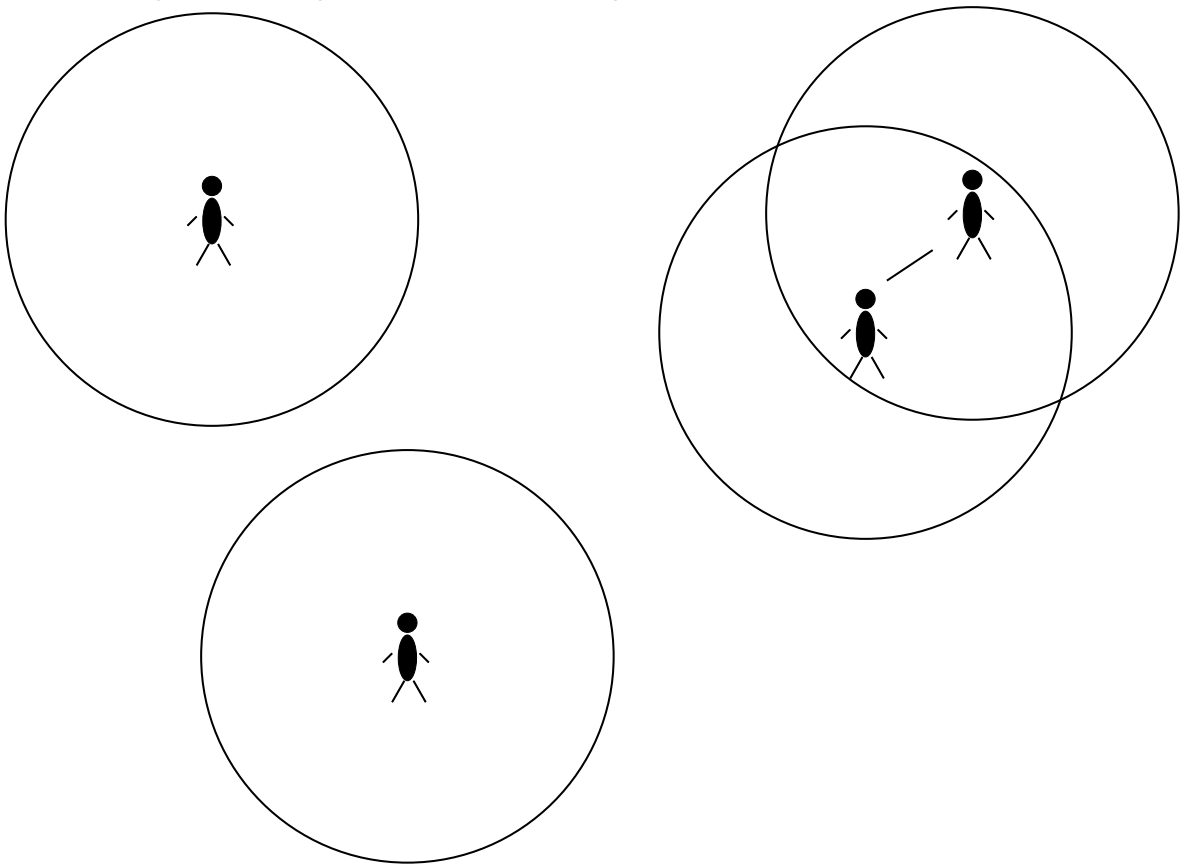
Mobility modify the proximity between users, hence wireless communications



Object under study

Human-contact networks

Mobility modify the proximity between users, hence wireless communications



Motivations

Ad-hoc communications : transmission hop-by-hop

- Reduce the cost (eg. broadcast)
- Extreme events (eg. earth quake)
- Military applications

Virus propagation

Radio proximity $\leftrightarrow$ physical proximity

Efficiency ?

Capacity :

- for a pair source/destination : the amount of data (bits) that can be transmitted (per sec.)
- of the network : average capacity for a random pair source/destination

Need to model :

- who communicates with whom ?
  - when
  - how many times (frequecy)
- which protocol of communication ?

Point-to-point

For each packet, the source waits until it meets the target.

Advantage :

Drawbacks :

Point-to-point

For each packet, the source waits until it meets the target.

Advantage :

simple

Drawbacks :

Time to send the paquet ?

Flooding protocol

Each node send all its packets to all the nodes it meets

Avantages

Drawbacks

Flooding protocol

Each node send all its packets to all the nodes it meets

Avantages

- time to reach every one : proxy for a lower bound in terms of times required for the transmission

Drawbacks

- Waste of memory
- Interference

Two-hop relay

Each source transmits its packets to the first node it meets :

- either directly to the target
- either to a relay

Each relay waits until it meets the target

Avantages/drawbacks

Protocol used in an article. Detailed later

Static nodes

[Gupta, Kumar, 2000]

Settings :

- n nodes are randomly placed on a disc whose surface is 1.
- the nodes don't move
- for each source : one destination is randomly chosen
  one sends an infinity of packets

Results

- Capacity in $cte/\sqrt{n \log n}$ possible
- Capacity in $cte/\sqrt{n}$ impossible

Discussion

Discussion

The capacity desceases while the number of nodes increases
$\longrightarrow$ the more the number of nodes, the less the efficiency

Discussion

Important theoretical result
What happens if :

- nodes moves ?
- pattern of communication changes ?
  (who wants to communicate with whom)

With mobile nodes

[Grossglauser, Tse, 2002]

Settings :

- The nodes move in a disc whose surface is 1.
- $X_i(t)$: position of i at time t.
- Uniform distribution over the disc
- One node $\longrightarrow$ one destination
- One sends an infinity of packets

First idea

Each pair source/destination is infinitly often close

$\longrightarrow$ direct communications when *distance* $<$ *threshold*

Problem

Better compromise : impossible to reach $1/\sqrt{n}$

First idea

Each pair source/destination is infinitly often close
$\longrightarrow$ direct communications when *distance* $<$ *threshold*

Problem

- high threshold $\rightarrow$ lot of interferences
  few communication at the same time
- short threshold : short distances $\rightarrow$ needs to wait for longer

Total capacity low in both cases

Better compromise : impossible to reach $1/\sqrt{n}$

Second idea : relay

At each time step : one chooses randomly θn sources
($\theta < 1$, parameter)
the other nodes are receptors

2 modes

- if t odd : each source sends its packet to the closest receptor
  - either the destination
  - either a relay (unique)
- if t even : each source sends a relayed packet, if possible.
  - transmission only to the destination

Results

The capacity of each pair is constant

Discussion

Important theoretical result :

Discussion

Important theoretical result : the mobility plays a great impact

Only one relay : surprising ?

Discussion

Important theoretical result : the mobility plays a great impact

Only one relay : surprising ?

The modelling of the mobility assumes a trajectory :

- stationnary
- uniformly distributed on the disc
- independant in regards to the nodes

$\longrightarrow$ each node has the same probability to meet the destination

Is it a reallistic model ?

Let's find out

Contact and inter-contact duration times

For radio badges networks :

Contacts duration times

How long two nodes stay in contact

Inter-contact duration times

How long before two given nodes meets again

$\longrightarrow$ Distribution

One example : rollerblade tour

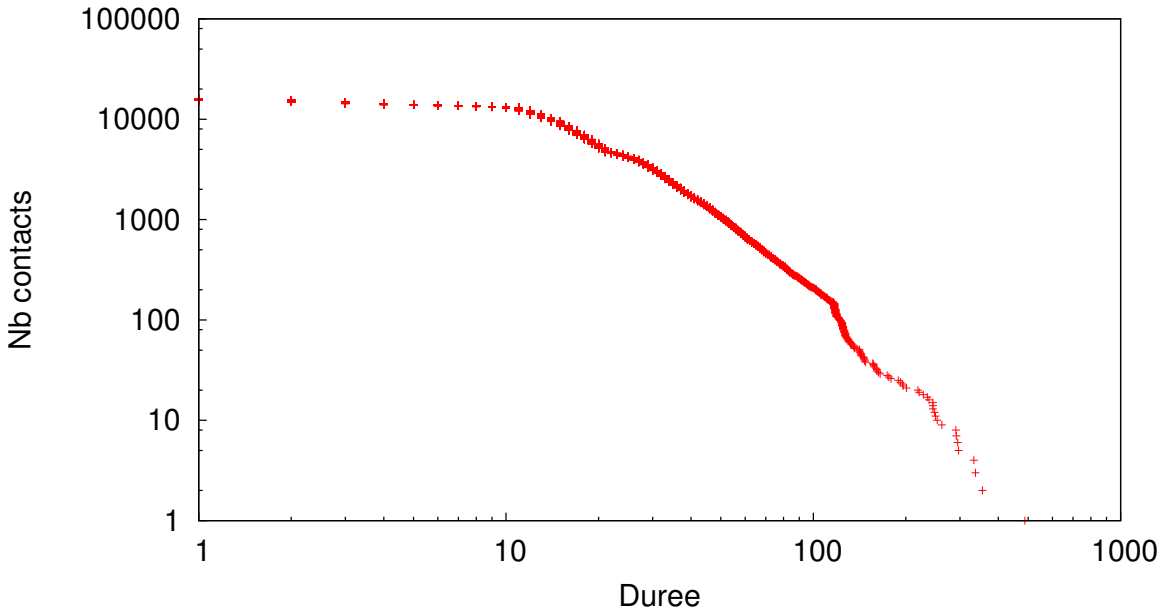
[Tournoux, Leguay, Benbadis, Conan, Dias de Amorim, Whitbeck, 2009]

Rollernet

- Rollerblade tour in Paris
- 62 nodes
- 3 hours

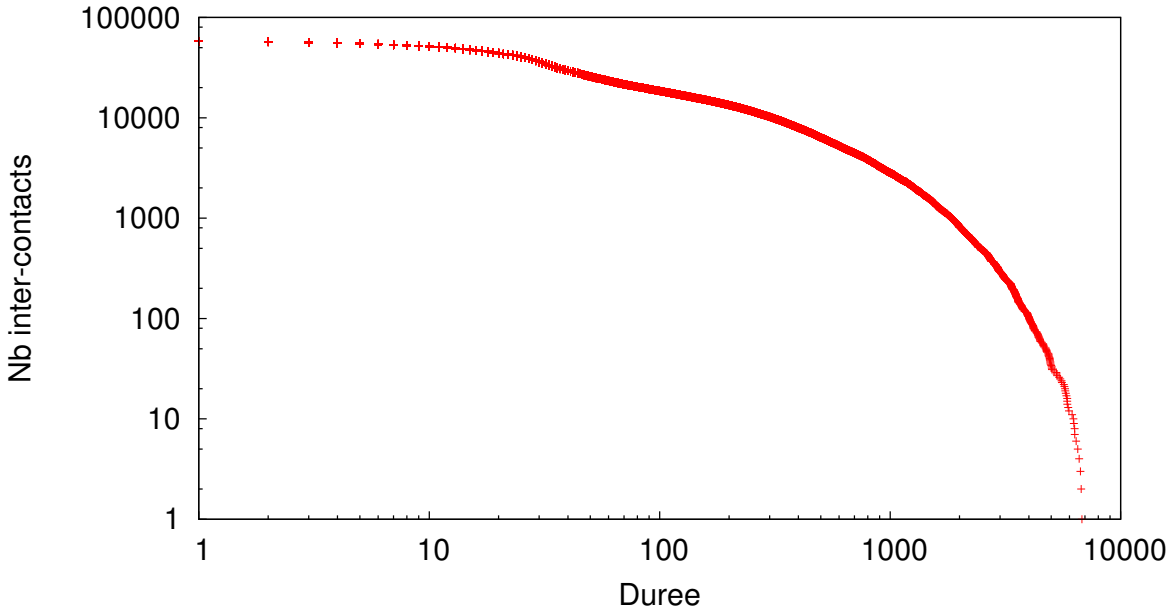
Contacts duration times

Inverse CDF (cumulative distribution frequency)



Inter-contacts duration times

Inverse CDF



Discussion

Distributions not homogeneous

(more or less heterogeneous depending on the dataset)

Mobility

Consequences

[Chaintreau, Hui, Crowcroft, Diot, Gass, Scott, 2006]

Inter-contacts duraction times following a powerlaw

$\Longrightarrow$

Discussion

Distributions not homogeneous
(more or less heterogeneous depending on the dataset)

Mobility

Observations incompatible with hypothesis made in the previous article :
It is not reallistic to model a regular and uniform mobility

Consequences

[Chaintreau, Hui, Crowcroft, Diot, Gass, Scott, 2006]

Inter-contacts duraction times following a powerlaw
$\Longrightarrow$ no efficient communications possible

Empirical study

Goal

The existing :

- Studies analyzing mobile networks [CHA07, TOU09, ...]
- Few models but recently : [CLE08, CLE09]

Used to study diffusion protocols :

- *flooding* protocol [BAU09, CLE10]
- *push* protocol [CLE13]

Goal

The existing :

- Studies analyzing mobile networks [CHA07, TOU09, ...]
- Few models but recently : [CLE08, CLE09]

Used to study diffusion protocols :

- *flooding* protocol [BAU09, CLE10]
- *push* protocol [CLE13]

In this course : is the model realistic ?

Models for evolving graphs

Background :

- **Evolving graph** model : recent [FER02]
- Evolving graph = Succession of distinct graphs G_0, G_1, ... with V given
- Capture all types of dynamics

Models for evolving graphs

Background :

- **Evolving graph** model : recent [FER02]
- Evolving graph = Succession of distinct graphs G_0, G_1, ... with V given
- Capture all types of dynamics

Variant of edge-markovian evolving graph :

- Temporal dependency in the evolution of the graph
- G_{t+1} determined by G_t and 2 parameters :
  - p : probability of creation of a *non-existing* link
  - d : probability of deletion of an *existing* link

Example

Example with 4 nodes, $p = 0.3$, $d = 0.2$ and 5 time steps.

1 3 2 3
1 4 1 3
2 3 1 2
2 3 4 4
2 4 1 2
2 4 4 4
3 4 1 3

- 1st and 2nd column : identifiers of nodes involved in the contact
- 3rd column : starting time of contact
- 4th column : ending time of contact

Advantages / drawbacks

Interest is twofold :

- $\forall G_0, p, d$: converge towards an Erdös Rényi graph with $\hat{p} = \frac{p}{p+d}$
- Few parameters $\Longrightarrow$ theoretical results

But it is also its weakness :

Advantages / drawbacks

Interest is twofold :

- $\forall G_0, p, d$: converge towards an Erdös Rényi graph with $\hat{p} = \frac{p}{p+d}$
- Few parameters $\Longrightarrow$ theoretical results

But it is also its weakness :

- 2 parameters to rule **all** creations/deletions
- Suppose that those 2 values are representative for the l'**entire** evolution of the de network

Methodology

Goal of the course :

Conduct a study to see if it is true.

- Analyze properties of the dynamics as observed in several dataset
- Comparison with the markovian model

Elements of response

- Yes for [WHI11] (and [VOJ11]) but ...
- ... study over 1 dataset
- ... the criteria is weak : time needed to flood the network

Rollernet

- Rollerblade tour in Paris
- Date : August 2006.
- Duration : 3h with a break (30 min) couvering approx. 30km,
- Location : street of Paris
- Technology : *iMotes* (bluetooth)
- Size : 62 participants
- Frequency : every 15s.

Infocom06

- Experiment made during Infocom conference at Barcelona.
- Date : April 2006
- Duration : 3 days
- Technology : *iMote*
- Size : 98 iMotes (78 participants, 17 static, and 3 in elevators)
- Frequency : every 120s.

Sociopattern

- Exhibition in at a gallery (deseases propagations).
- Date : 2009
- Duration : 3 months
- Technology : *radio bagdes*
- Size : 88 to 410 (depends on the day)
- Frequency : every 20s.

6 case studies

Dataset	RollerNet	Infocom05	Infocom06	HT09	Socio	PMTR
Duration	3 hours	4 days	4 days	2,5 days	1 day	10 days
Participants	62	41	98	113	151	44
Contacts	60 146	17 682	148 784	9 865	2 051	11 895
Frequency (sec.)	15	120	120	20	20	1

For each :

- "Physical" contact network among individuals
- Each individual is equipped with a sensing device
- Detection between devices if proximity between individuals (2 to 10 m.)
- Frequency of detection varies, as well as duration of the experiments

In the rest of the presentation, 3 dataset only :

- `RollerNet`
- `Infocom06` : similar to `Infocom05`
- `SocioPattern` : similar to `HT09` and `PMTR`

Methodology

For each dataset and for each time step

- Fraction of created links (over possible new links)
- Fraction of deleted links (over existing links)

Corresponds to the parameters p and d of the model

Analyze :

- Evolution over time
- Distribution of the values
- Generation of artificial graphs according to the markovian model
- Comparison between real/artificial graph

Created links

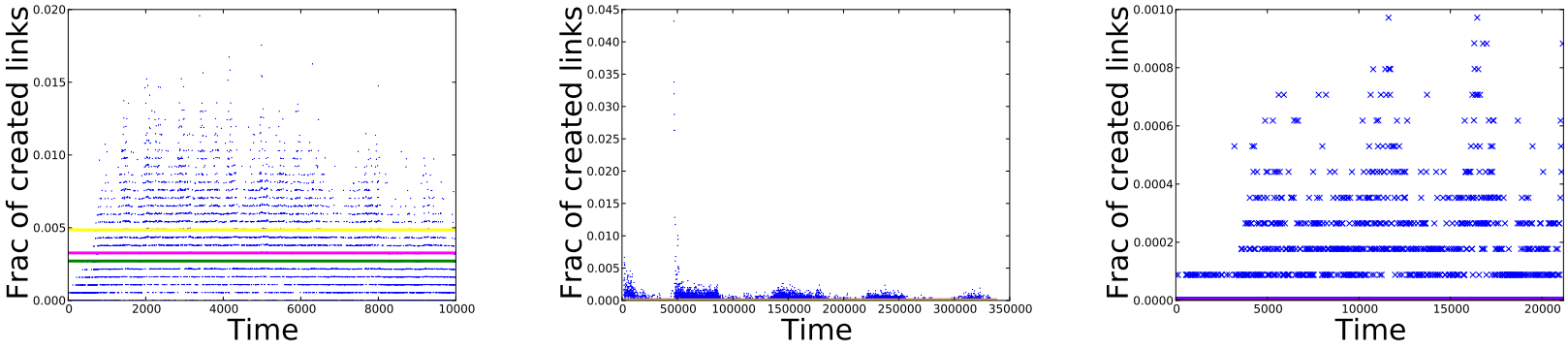


FIGURE : Evolution of the proportion de created links over time

Results

Created links

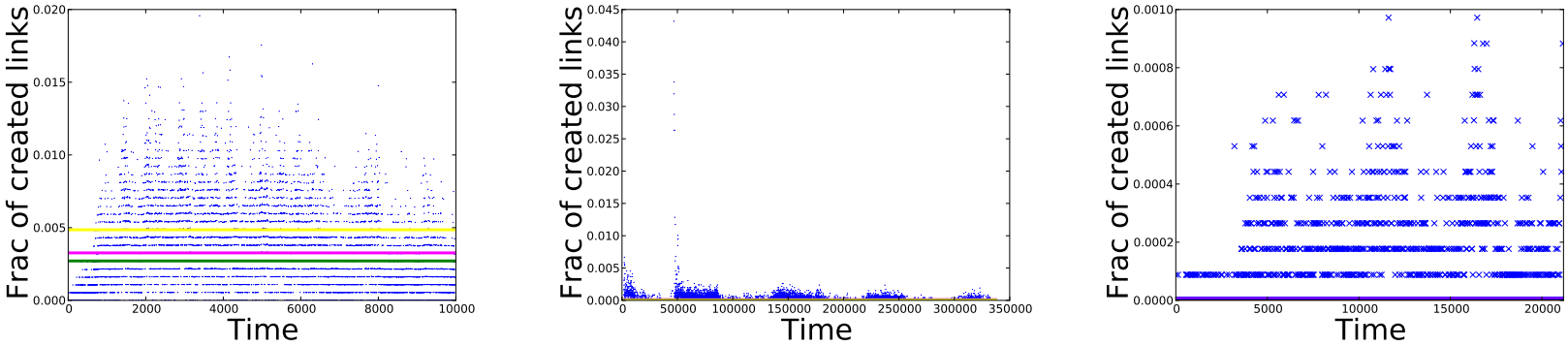
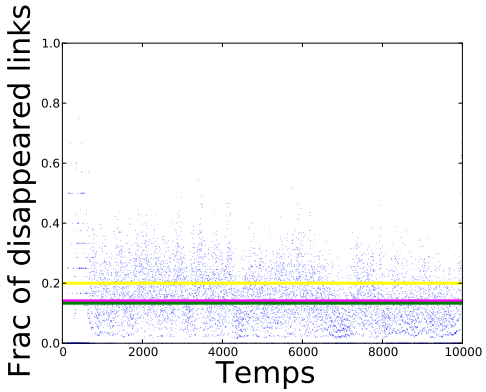


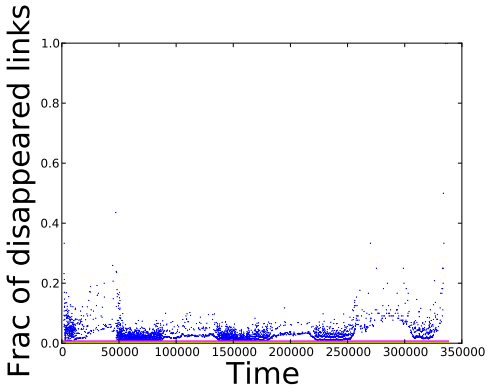
FIGURE : Evolution of the proportion de created links over time

Results

- `RollerNet` : notion of average is relevant
- `Infocom06`, `SocioPattern` : wide range of values
- `Infocom06`, `SocioPattern` : average, median and 75th percentile overcome by weak values
- $\Longrightarrow$ `Infocom06`, `SocioPattern` : non realistic

Deleted links





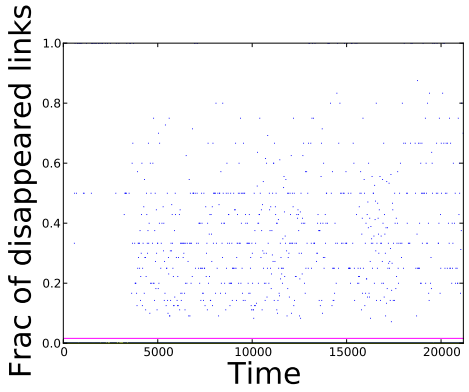
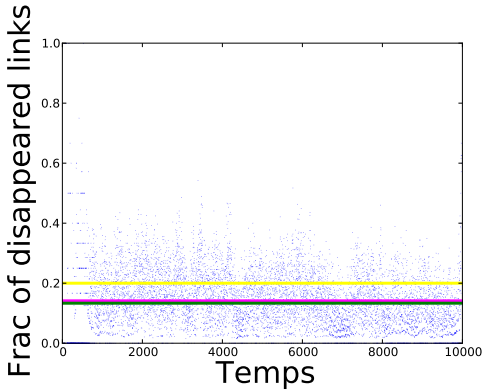
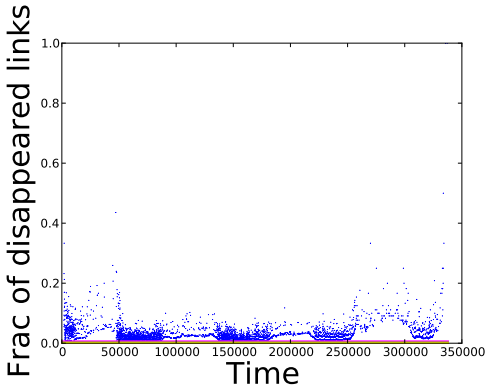


FIGURE : Evolution of the proportion of deleted links over time

Results

Deleted links





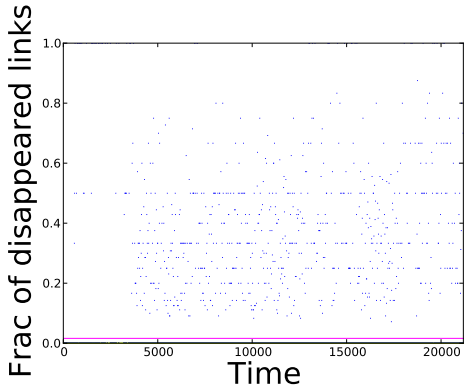


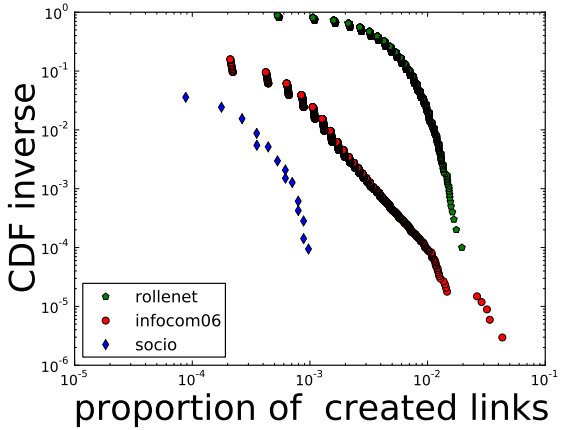
FIGURE : Evolution of the proportion of deleted links over time

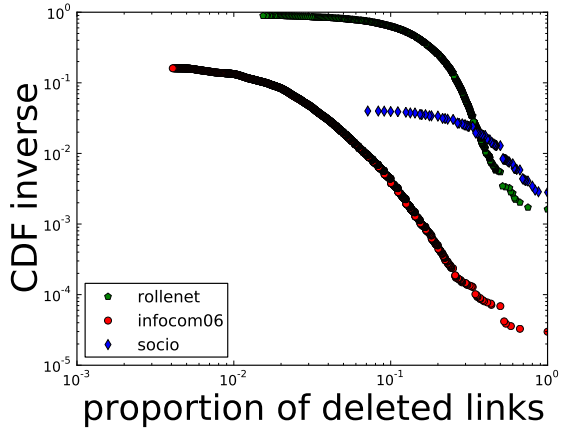
Results

- Same observation but amplified
- Range of values is covered ([0 : 1])
- Particular case for $d = 1$

Distribution of p and d values

Dataset	RollerNet	Infocom06	Socio
Fractions of created links (average)	3.2 (10^{-3})	9.5 (10^{-5})	9 (10^{-6})
Fractions of deleted links (average)	1.4 (10^{-1})	4.5 (10^{-3})	1.6 (10^{-2})

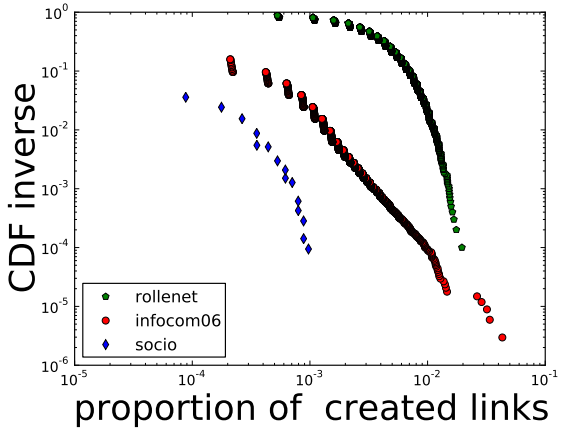


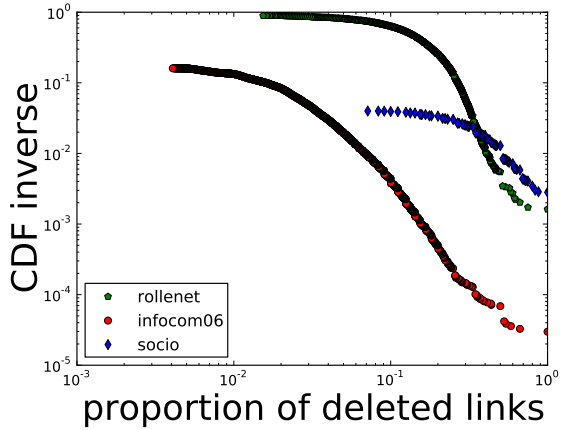


Results

Distribution of p and d values

Dataset	RollerNet	Infocom06	Socio
Fractions of created links (average)	3.2 (10^{-3})	9.5 (10^{-5})	9 (10^{-6})
Fractions of deleted links (average)	1.4 (10^{-1})	4.5 (10^{-3})	1.6 (10^{-2})





Results

- Clearly heterogeneous for `Infocom06`
- and on several order of magnitudes
- `RollerNet` : sudden slope around the average value

Methodology

So far :

- Studied the dynamics related to creation and delation of links
- Provided evidences that the models *is probably not suited* to particular dataset

How to demonstrate that the model is not pertinent ?

Methodology

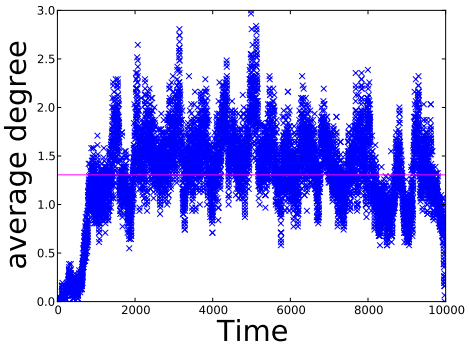
So far :

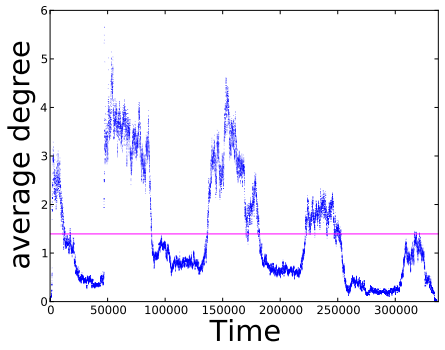
- Studied the dynamics related to creation and delation of links
- Provided evidences that the models *is probably not suited* to particular dataset

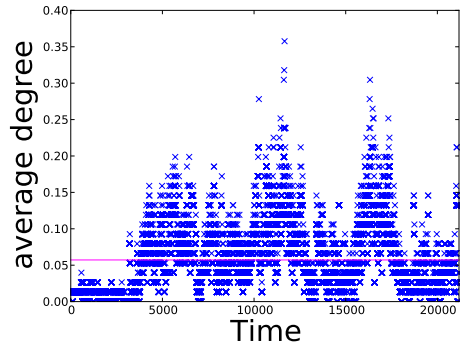
How to demonstrate that the model is not pertinent ?

- Choose an external criteria (ie not the fraction of created and deleted links) ...
- ... but close enough the meaning of p and d (for fairness)
- Compute the value of the criteria for the real and the artificial graphs.
- Comparison between real/artificial graph.

Evolution of mean degree

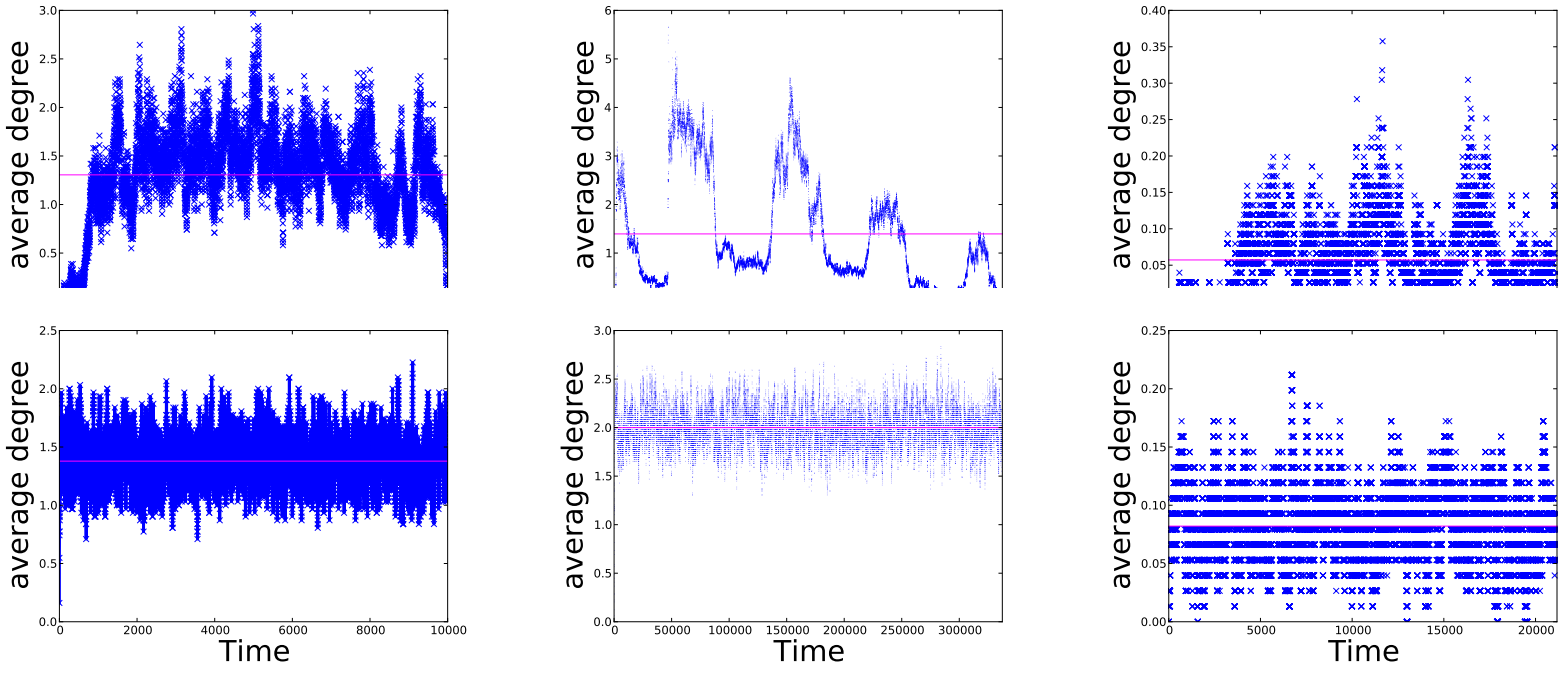






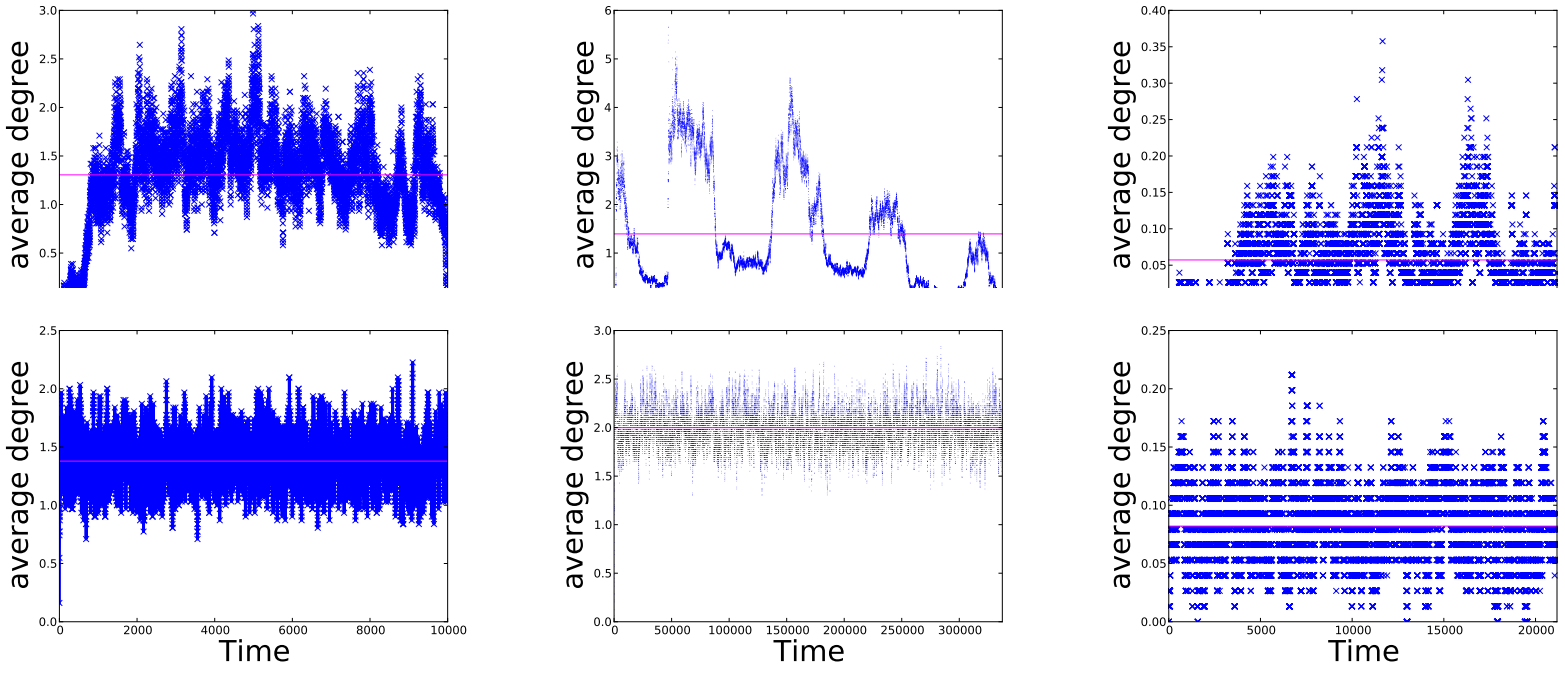
Results

Evolution of mean degree



Results

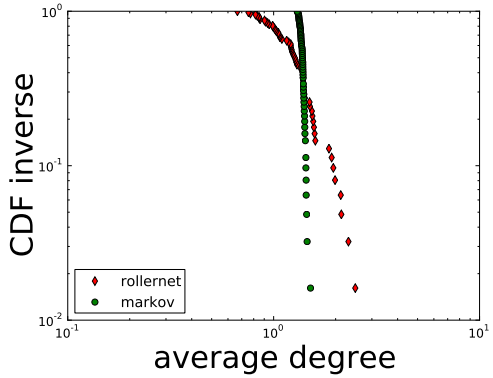
Evolution of mean degree

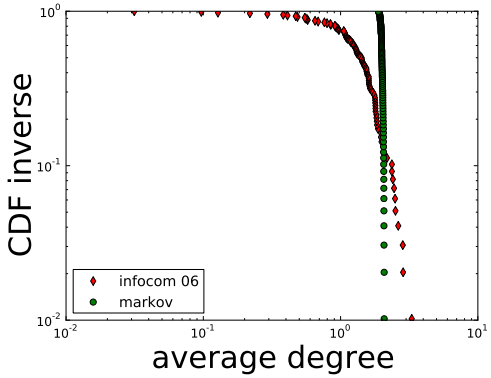


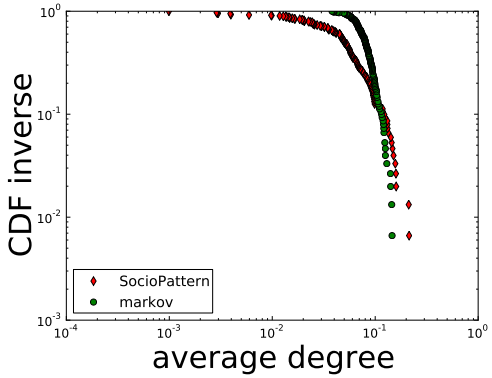
Results

- "Uniformization" for `Infocom06` and `SocioPattern` (not the same range of values !)
- Seems to have little impact on `RollerNet`
- Except at the beginning (expected)

Average degree distribution

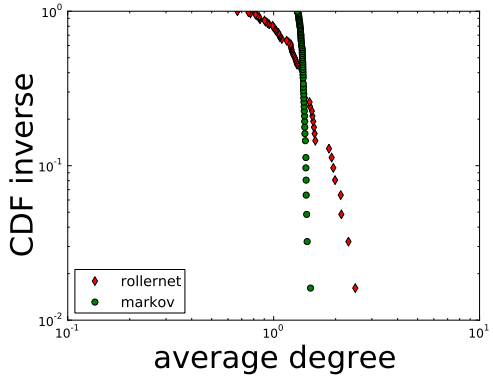


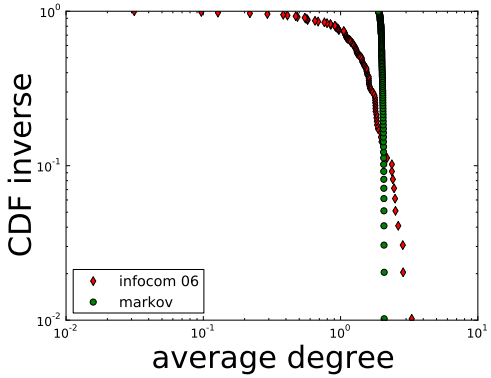


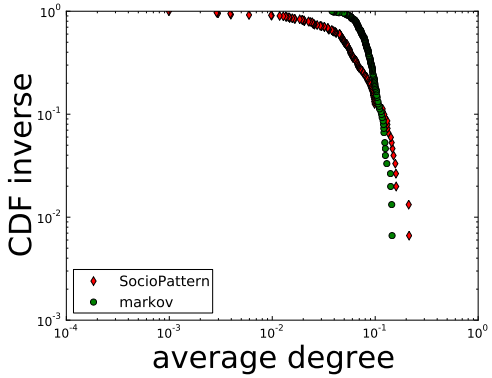


Results

Average degree distribution



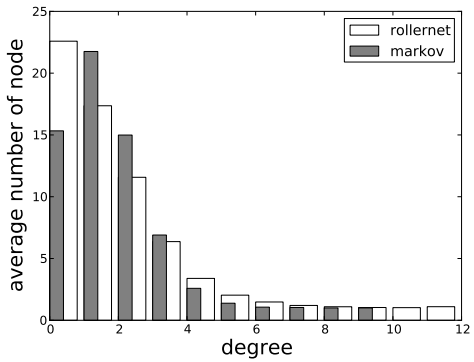


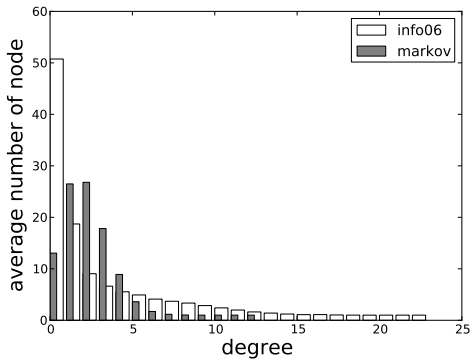


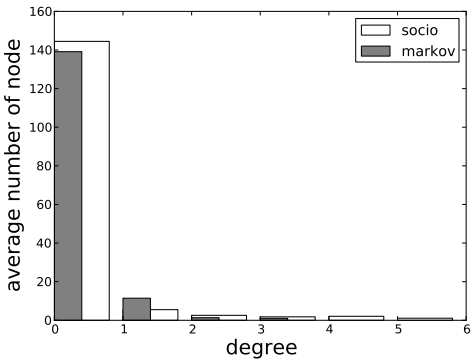
Results

- `Infocom06` : clear differences between model and real data (expected)
- `RollerNet` and `SocioPattern` : also different, although less obvious

Degree distribution

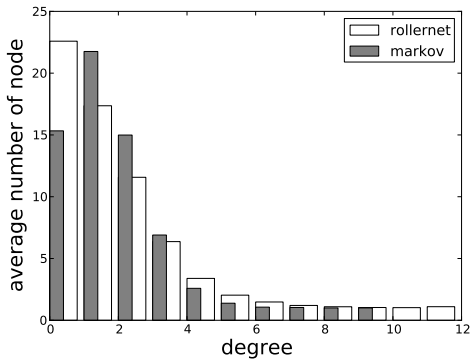


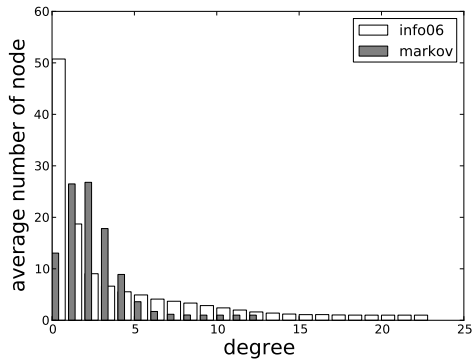


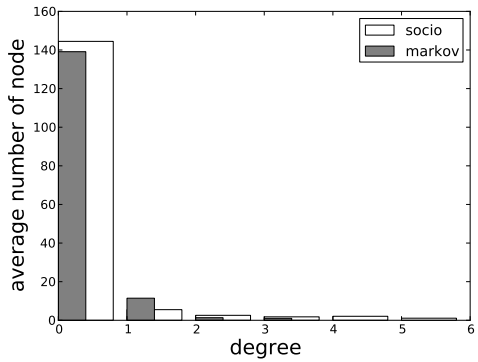


Results

Degree distribution



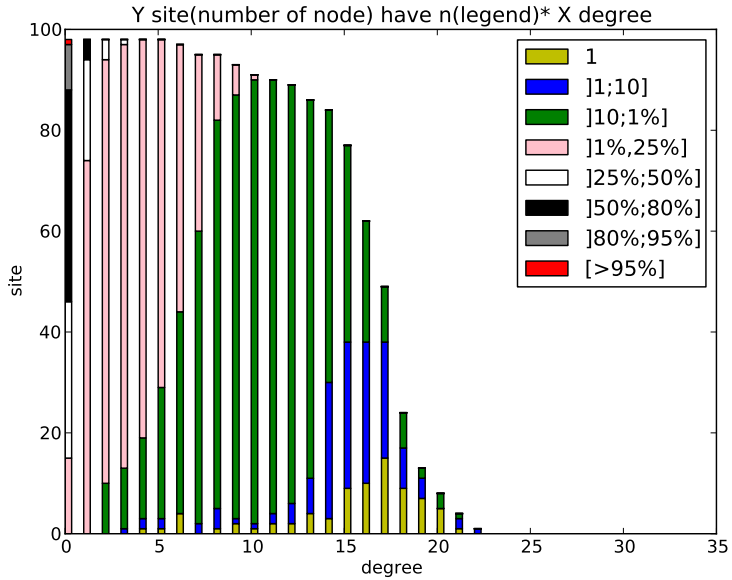


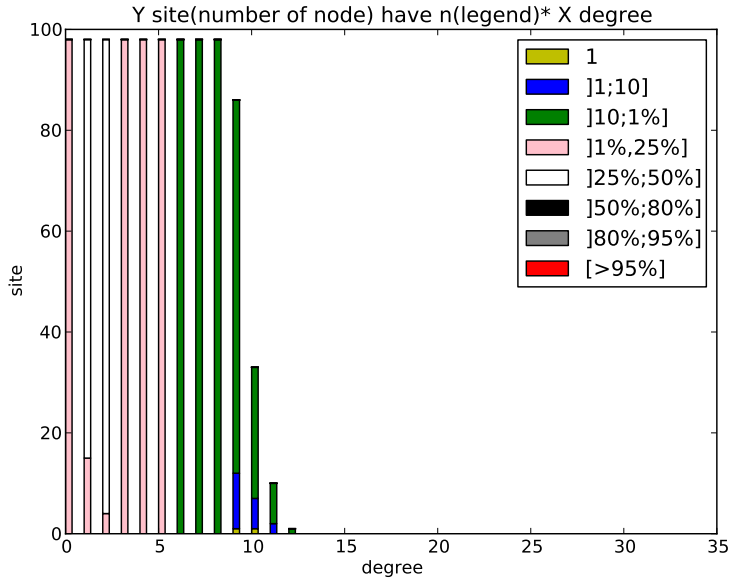


Results

average value relevant $\not\Longrightarrow$ the model reproduces well the **global** properties of the networks

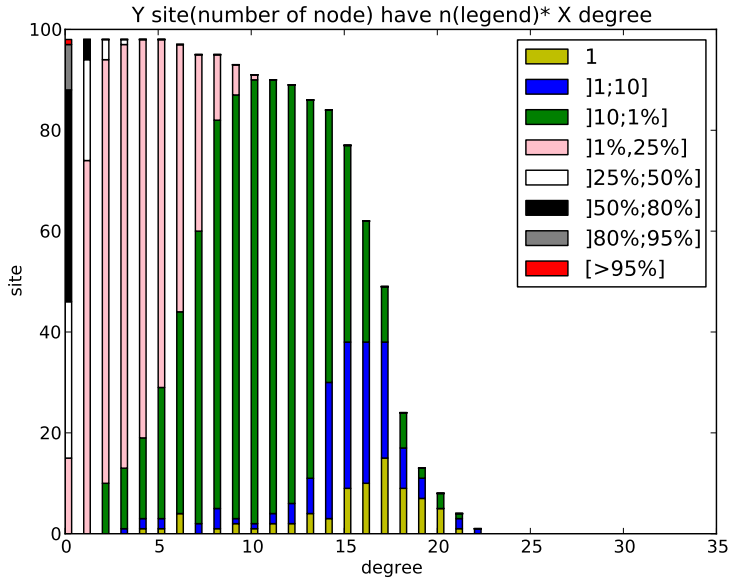
Distribution and frequency of the degrees (Infocom06)

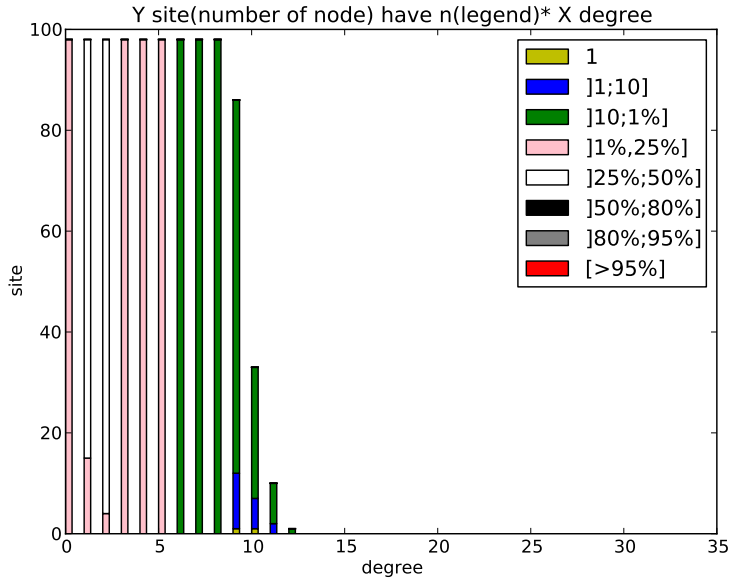




Results

Distribution and frequency of the degrees (Infocom06)

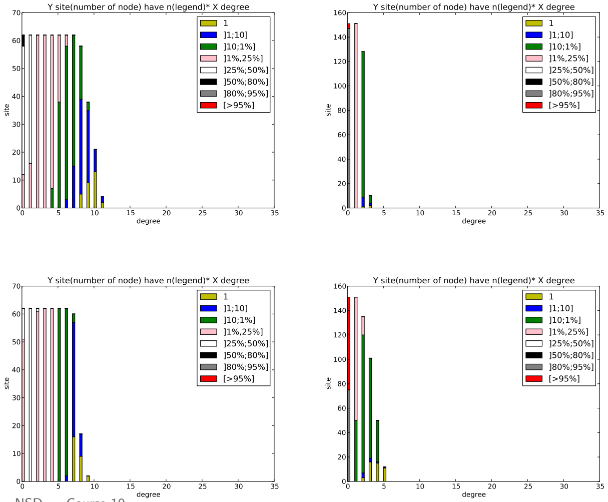




Results

- Nodes are more degree-stable in real networks
- Small degrees are over-represented
- *No node with the same degree more than 50 % of the time in the model*

Distribution and frequency of the degrees (RollerNet, SocioPattern)



Conclusions and perspectives

Conclusions

- Confrontation markovian model vs. real data
- Hypothesis of homogeneity does not stand in most of the cases
- Even in favourable case, it **does not reproduce the dynamics**
- *Still remain useful* : cf [WHI11, VOJ11]

Perspectives

- Consider other way to define p and d (following an heterogeneous distribution ? different for each nodes ? depending on the graph state ? ...)
- Study refined properties (repartition of connexions)
- Analyze correlation between creations and deletions
- Take into account the local density
- Study *gossip* protocols of diffusion on **real data**