Characteristics of the Dynamic of Mobile Networks

P. Borgnat, E. Fleury, J.-L. Guillaume, C. Robardet, A. Scherrer
http://perso.ens-lyon.fr/eric.fleury/
http://www.ens-lyon.fr/LIP/D-NET/
mailto://Eric.Fleury@inria.fr

ENS Lyon/LIP – INRIA/D-NET

Workshop: Dynamical Complex Systems
December 2009, Buenos Aires, Argentina
http://cnet.fi.uba.ar/wdcs/

Outline

Outline

- **MOSAR Project**
  - Project overview
- Dynamic Network Characterization
  - Motivation
  - Statistical analysis of snapshots of graphs
  - Towards a global analysis of the dynamics
  - Modeling of the dynamics
- Conclusion

Deployment of a large-scale dynamic networks

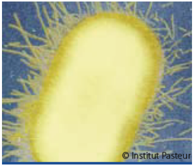
Control of antimicrobial resistance of bacteria responsible for major and emerging nosocomial infections.

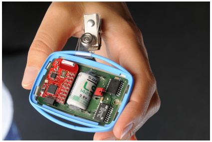
MOSAR Experiment

- Medical / staff / Patients (500 people)
- Individual antibiotic use;
- Characterization of the isolates bacteria and their epidemicity;
- 7/24 during **6 month** long period

Document contact frequencies

- Associate 1 sensor with each actor
- monitor the dynamic (inter & intra contact)



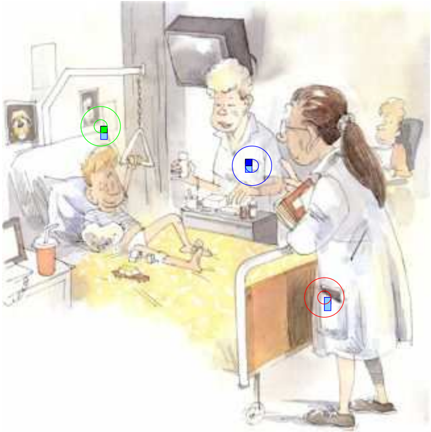


Patient room

Patient room

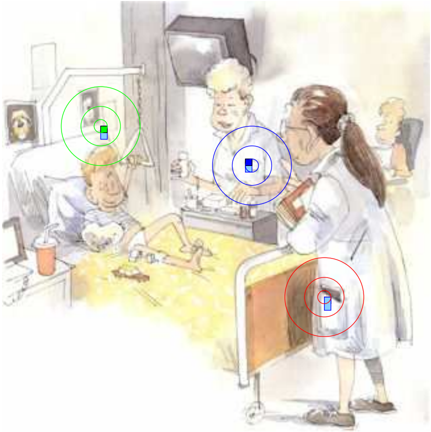
Patient room

Patient room



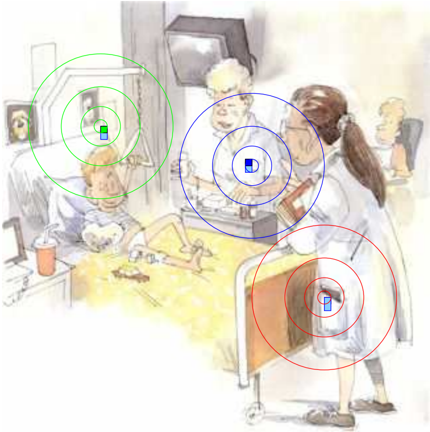
Patient room

Patient room



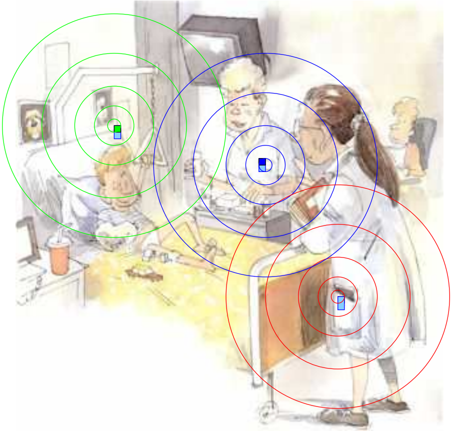
Patient room

Patient room



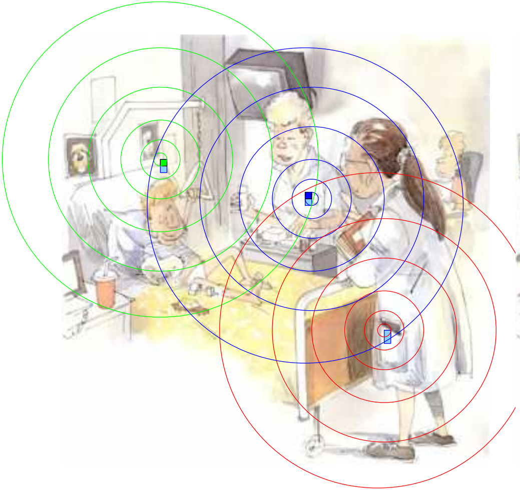
Patient room

Patient room



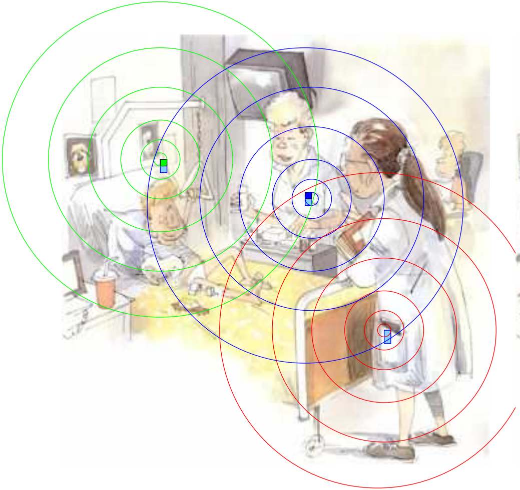
Patient room

Patient room

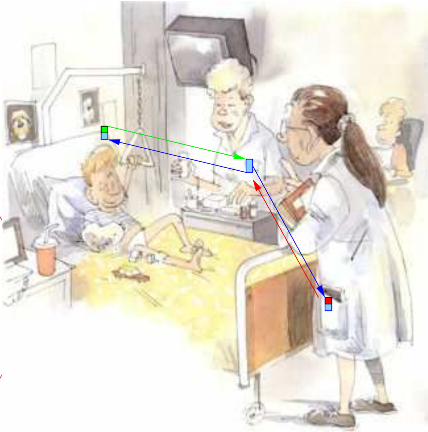


Patient room

Patient room

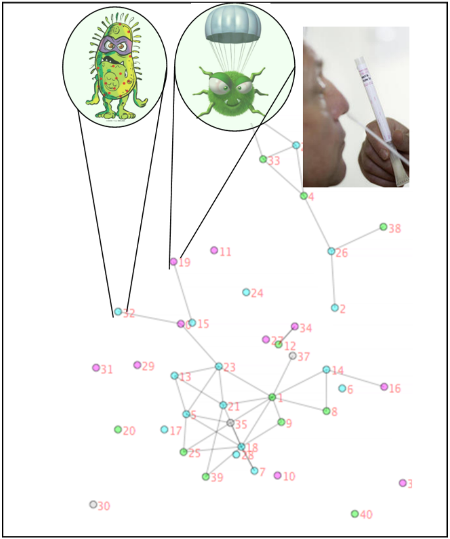


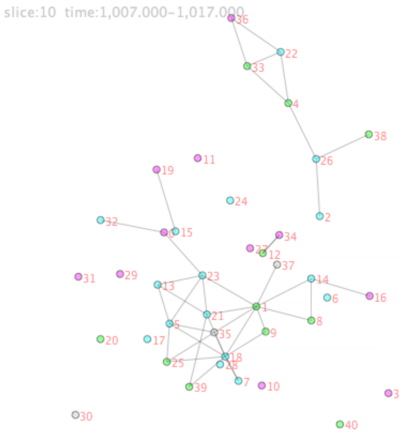
Patient room

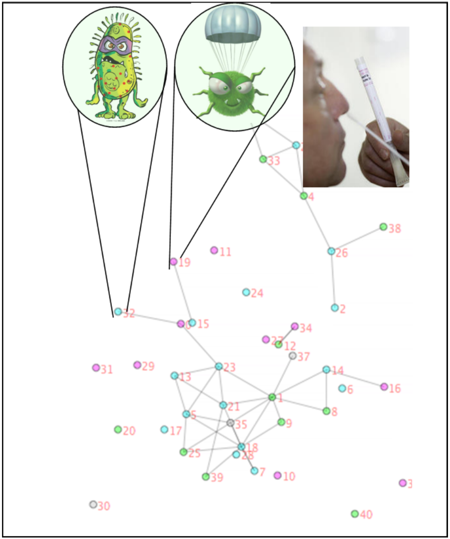


Patient room

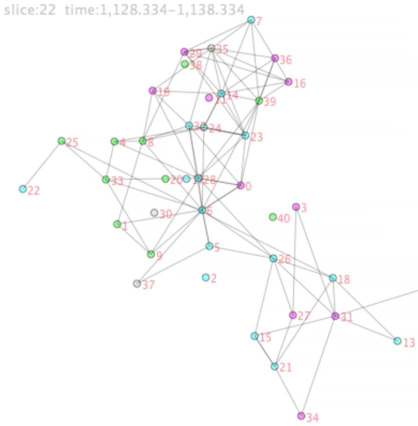
Multi modal / multi time scale

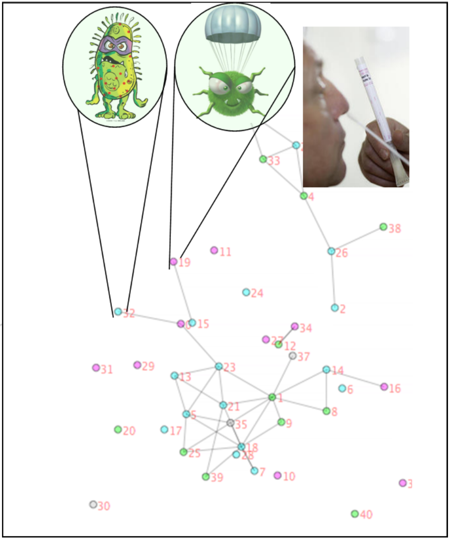






Multi modal / multi time scale





Multi modal / multi time scale

Outline

Objectives

MOSAR project

- Better understand the intrinsic characteristics / **properties** of dynamic networks
- Model / analyze interaction between node/users
- **Describe accurately the dynamics**

Two central questions:

- Obtaining random models that reproduce "*these*" properties
- How do their functionalities constrain the structures of real network?

Objectives

MOSAR project

- Better understand the intrinsic characteristics / **properties** of dynamic networks
- Model / analyze interaction between node/users
- **Describe accurately the dynamics**

Two central questions:

- Obtaining random models that reproduce "*these*" properties
- How do their functionalities constrain the structures of real network?

Objectives

MOSAR project

- Better understand the intrinsic characteristics / **properties** of dynamic networks
- Model / analyze interaction between node/users
- **Describe accurately the dynamics**

Two central questions:

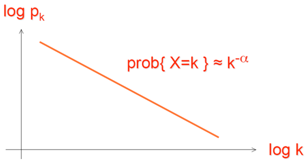
- Obtaining random models that reproduce “*these*” properties
- How do their functionalities constrain the structures of real network?

Preliminary data[1]

Mosar experiment: May 2009 – Nov 2009.

"Toy" traces are now available

- 41 nodes, 3 days (254 151 sec), every 120sec
- 820 possible links,
- *"[...] inter contact time distribution can be compared to the one of power law [...]"*



Power law...

- What do power law really signify?
- Is it the ultimate argument?

[1] A. Chaintreau and J. Crowcroft and C. Diot and R. Gass and P. Hui and J. Scott, *Impact of Human Mobility on the Design of Opportunistic Forwarding Algorithms*, INFOCOM 2006

Methodology

Descriptive: Standard graph properties

1. as a function of time to to provide an empirical statistical characterization of the dynamics.
2. temporal evolution of the snapshots
3. statistical signal processing

Analysis: global indicators

- connected components, triangles, and communities

Model

- We propose models to perform random dynamic networks simulations.

Standard graph properties

Snapshots $G_t = (V^0, E_t)$

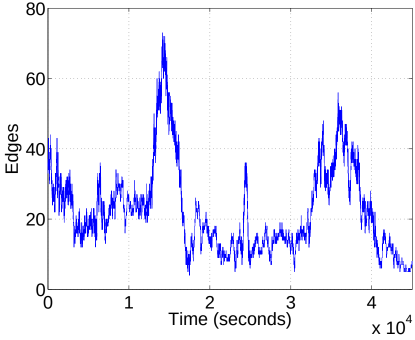
- Active links: $E(t) = |E_t|$
- Connected vertcices: $V(t) = |\{u \in V^0, d_{G_t}(u) > 0\}|$
- Average degree of connected vertices is $D(t) = \sum_{u \in V^0} d_{G_t}(u) / V(t)$
- Number of connected components (*maximal subgraph such as every node of the subgraph is connected to each another node*): $N_c(t) = |C_{G_t}|$
- Number of triangles: $T(t) = |T_{G_t}|$

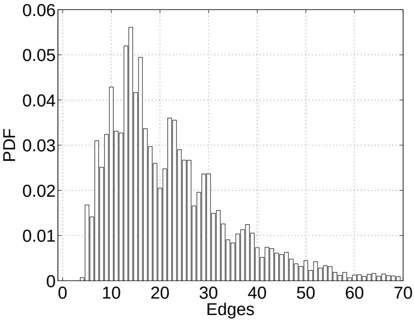
Property		IMOTE Mean	IMOTE Std. Dev.	IMOTE Corr. Time (s)
#Active links	$E(t)$	21.9	12.4	5200
#Connected vertices	$V(t)$	19.9	4.7	7400
Avg degree	$D(t)$	2.1	0.8	3600
#CC	$N_c(t)$	4.8	2.1	5600
#Triangles	$T(t)$	6.9	8.30	4700

Standard graph properties (cont)

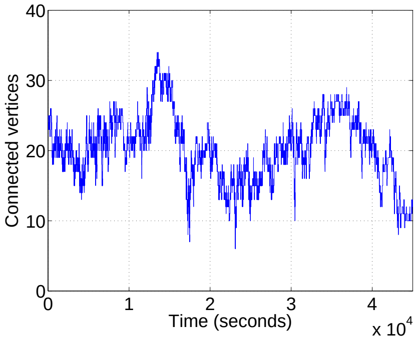
Probability distribution

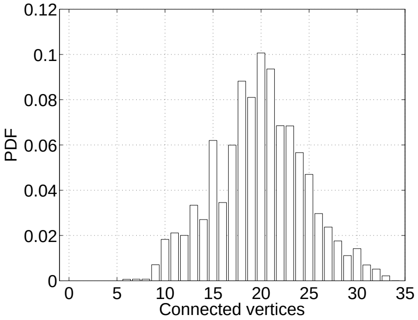
- time bin of $1s << $ period.
- PDF obtained are not heavy tailed
- variability is not very large (stdv is a good measurement of the variability)





Standard graph properties (cont)

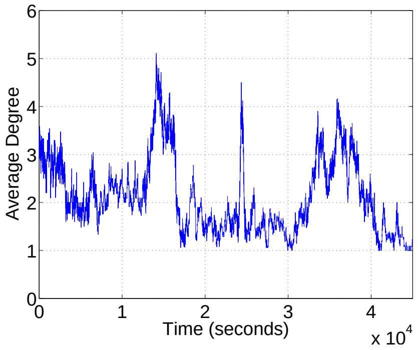


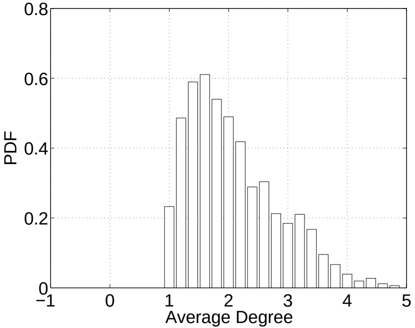


Network is sparse

- less than 10% of active links among the 820 possible links
- at no time the network is a single connected component.
- many nodes remain isolated during long times (around 50% on average for daytime and more than 90% for nighttime).

Standard graph properties (cont)



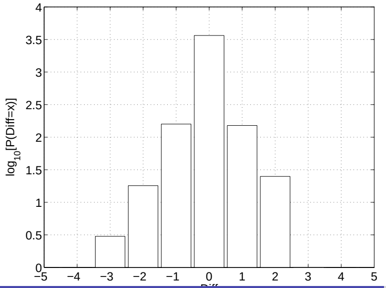


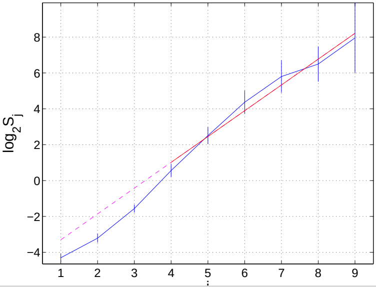
Standard graph properties (cont)

differential sequence: $DS[k] = D[k+1] - D[k]$

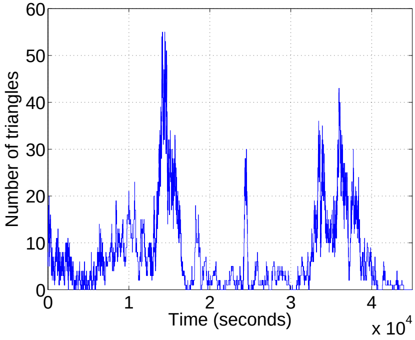
- log-log representation of the covariance in the wavelet domain[a]
- S_j is roughly the average of the wavelet coef. at scale j
- Hurts exponent is close to the special value 0.5.
- no long range $\longrightarrow$ Independent Identically Distributed (IID)

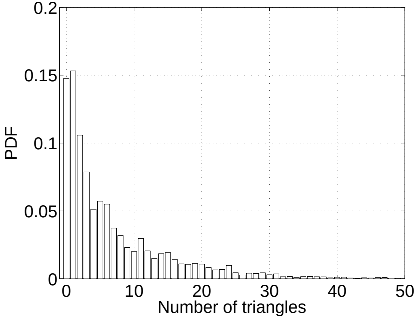
[a] P. Abry and D. Veitch, Wavelet analysis of long-range dependent traffic, TIT, 1998





Standard graph properties (cont)





Large number of triangles

- $\mathbb{E}(T(G(p,n))) = \binom{N}{3}3!p^3$ & $\mathbb{E}(E(G(p,n))) = p\frac{N(N-1)}{2}$
- When there is k links, $\mathbb{E}(T(G(n,k))) \sim \frac{8k^3(N-2)}{N^2(N-1)^2}$
- 70 links (max) $\longrightarrow$ 40(60)
- 22 links (avg) $\longrightarrow$ 1(7)

Dynamical characteristics

Correlation times

- temporal evolution ($X(t)$: univariate time-series)
- The autocorrelation function of $X(t)$:

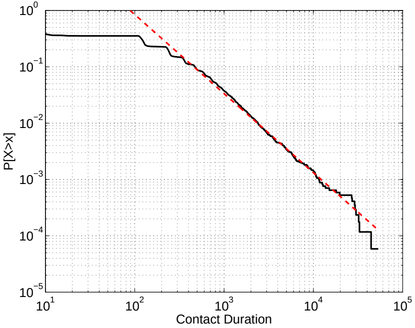
$$C_X(\tau) =< X(t+\tau)X(t) >_t - (< X(t) >_t)^2$$

- correlation time: **first time where the function $C_X(\tau)$ goes to zero**

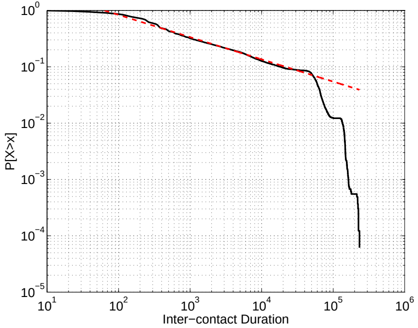
notes

- correlation times of E, V and N_c are rather large: $\sim$ 1h15.
- D and T have comparable correlation times.
- **This suggests that these properties evolve under a common cause.**

Dynamical characteristics (cont)



Mean : 140; $\alpha = 1.66$

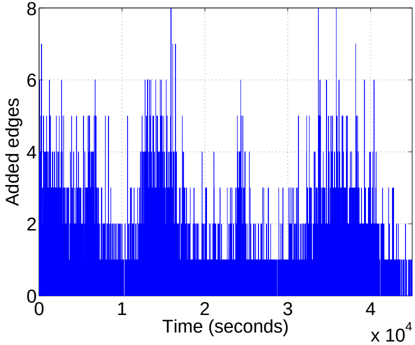


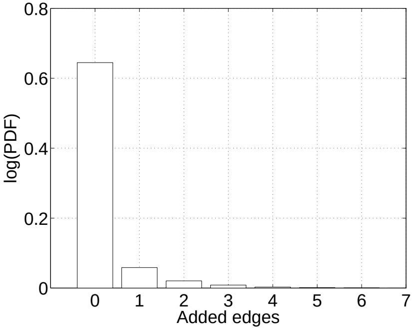
Mean : 3680; $\alpha = 0.60$

Contact and inter-contact durations

- $P[X > x] \underset{x \to \infty}{\sim} cx^{-\alpha}$.
- $\alpha > 2$: finite mean/variance; $\alpha < 2$, infinite variance (*heavy tailed*).
- $\alpha < 1$, infinite mean/variance.

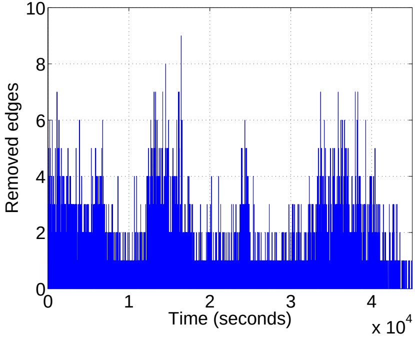
Dynamics of links creation and deletion

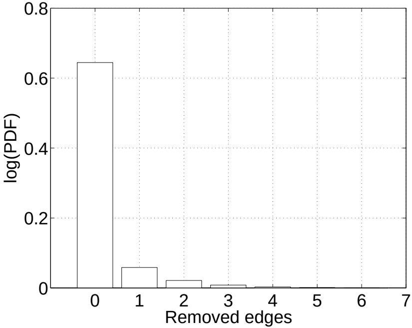




- $E_{\oplus}(t) = |\{e \in E_t, e \notin E_{t-1}\}|$, the number of links added at time t

Dynamics of links creation and deletion (cont)





- $E_{\ominus}(t) = |\{e \in E_{t-1}, e \notin E_t\}|$, the number of links removed at time t

		IMOTE		
Property		Mean	Std. Dev.	Corr. Time (s)
Edge creation	$E_{\oplus}(t)$	0.15	0.55	680 $\sim$ 12min
Edge delation	$E_{\ominus}(t)$	0.15	0.55	680$\sim$ 12min

Multivariate statistics of graph properties

Cross-correlations

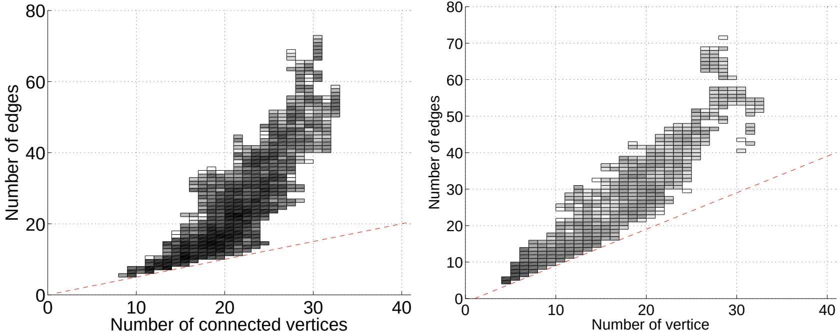
- Strong influence $E(t)$ over $V(t)$;
- $N_c(t)$ related to $E(t)$
- Less related: $N_c(t)$ and $V(t)$
- $E_\oplus(t)$ and $E_\ominus(t)$: mostly uncorrelated

	$E(t)$	$V(t)$	$N_c(t)$	$D(t)$	$T(t)$	$E_\oplus(t)$	$E_\ominus(t)$
$E(t)$	1	0.85	-0.56	0.95	0.90	**0.19**	**0.15**
$V(t)$	0.85	1	-0.20	0.70	0.66	**0.15**	**0.11**
$N_c(t)$	-0.56	-0.20	1	-0.70	-0.41	**-0.16**	**-0.15**
$D(t)$	0.95	0.69	-0.69	1	0.86	**0.19**	**0.15**
$T(t)$	0.90	0.66	-0.41	0.86	1	**0.15**	**0.11**
$E_\oplus(t)$	**0.19**	**0.15**	**-0.16**	**0.20**	**0.15**	1	**0.03**
$E_\ominus(t)$	**0.15**	**0.11**	**-0.15**	**0.16**	**0.10**	**0.03**	1

Multivariate statistics of graph properties

Joint distributions

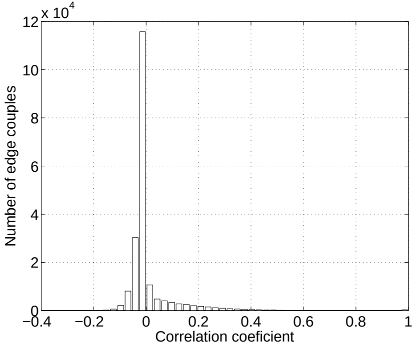
- $P_{XY}(x, y) = P[X = x \text{ and } Y = y] = P[X = x / Y = y]P[X = x]$
- variation of the # links is not constant over the # vertices



Multivariate statistics of graph properties

Link correlations

- Most pairs of links have a very low correlation coefficient.



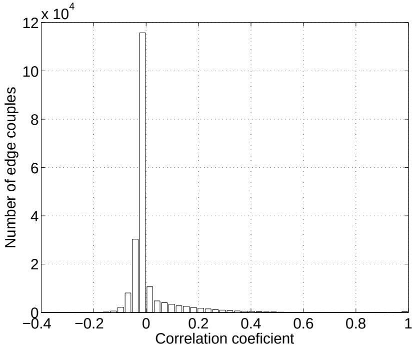
Markovian evolution

1. Correlation time link creation/deletion is small
2. Independent from the evolution of other graph properties
3. Links are independents

Multivariate statistics of graph properties

Link correlations

- Most pairs of links have a very low correlation coefficient.



Markovian evolution

1. Correlation time link creation/deletion is small
2. Independent from the evolution of other graph properties
3. Links are independents

Towards a global analysis of the dynamics

global properties

- not directly interpretable in the sequence of static graphs
- stability of connected components
- communities embedded in the network
- proportion of creation of triangles

Towards a global analysis of the dynamics

global properties

- not directly interpretable in the sequence of static graphs
- stability of connected components
- communities embedded in the network
- **proportion of creation of triangles**

Triangles in the graphs

	$P_{+/tri+}$	$P_{+/tri=}$	$f_{+/tri+}$	$f_{+/tri=}$
IMOTE	44 %	56 %	6 %	94 %
RANDOM	10 %	90 %	5 %	95 %

links / triangles

- $P_{+/tri+}$: link creation → triangle
- $f_{+/tri+}$: innactive link → triangle
- 40% of link creations increase the number of triangles
- proportion of inactive links that would create a triangle is very low
- More potential links doesn not imply higher $P_{+/tri+}$

Modeling of the dynamics

Simulation algorithm

- transition model with Markovian property
- links e are independent
- state of the network
- links e changes with $P_{tr}(e, G_t)$
- duration $\tau(e)$ since the link e has last changed its status

Ingredients

- contact / inter contact duration distribution
- elaborated graph properties $(E(t), V(t), N_C(t), D(t))$
- dynamical information (triangles)

Modeling of the dynamics

Input: Simulation time

Output: Random Dynamic Graph

```
foreach Simulation Time Step t do
    foreach link e do
        P_tr(e, G_t) = TransitionProbability(e) given the state G_t ;
        p_r = Uniform(0,1);
        if (p_r ≤ P_tr(e)) then
            ChangeState(e);
        end
    end
end
```

Ingredients I

Contact distribution

- heavy-tailed distributions for contact P_{ON} and inter-contact P_{OFF} durations
- $P_+(\tau)$: probability that one link that was OFF since τ ($\tau \geq 1$) is activated
- $P_{ON}(\tau) = P_-(\tau) \times \prod_{i=1}^{\tau-1}(1 - P_-(i))$

$$P_-(\tau) = \frac{P_{ON}(\tau)}{\prod_{i=1}^{\tau-1}(1 - P_-(i))}, \quad \tau \geq 2, \quad P_-(1) = P_{ON}(1) \tag{1}$$

$$P_+(\tau) = \frac{P_{OFF}(t)}{\prod_{i=1}^{\tau-1}(1 - P_+(i))}, \quad \tau \geq 2, \quad P_+(1) = P_{OFF}(1) \tag{2}$$

Ingredients II

Imposed graph property distribution

- Rejection Sampling based on a Metropolis-Hastings algorithm
- new state $G'_t = \{G_t + S_e(t) \text{ changed}\}$, is accepted with probability $P_{RS}(G_t, G'_t) = \min\left(1, \frac{F(x(G'_t))}{F(x(G_t))}\right)$
- F is the target PDF for the graph
- The total probability of transition of link e is then: $P_{tr}(e, G_t) = P_{-/+}(\tau(e)) \cdot P_{RS}(G_t, G'_t)$.

Ingredients III

Imposed dynamics of triangles

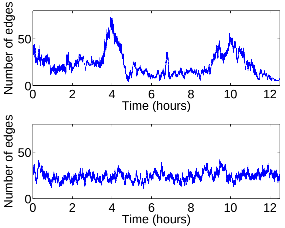
- reproduce the correct dynamical transition process concerning triangles
- do not want to change the mean probabilities of transition
- The weighted probabilities are then:

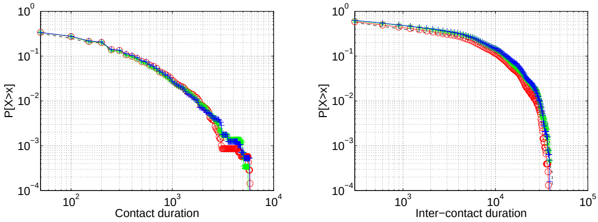
$$P_{tr}(e, G_t) = \begin{cases} P_+(\tau(e))\frac{P_{+/tri=}}{f_{+/tri=}} & \text{for link creation without new triangle,} \\ P_+(\tau(e))\frac{P_{+/tri+}}{f_{+/tri+}} & \text{for link creation with a new triangle.} \end{cases}$$

Simulation results

Investigated models

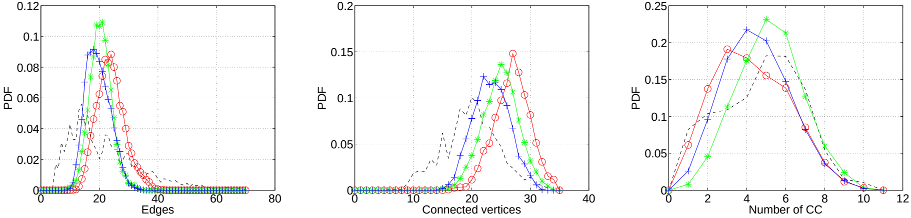
- $\mathcal{A}$: imposed empirical contact and inter-contact duration distribution only.
- $\mathcal{B}$: imposed distributions of contact / inter-contact durations , and of number of connected components.
- $\mathcal{C}$: distributions imposed contact / inter-contact durations and of number of connected vertices.





−− Imote / o Model $\mathcal{A}$ / $*$ Model $\mathcal{B}$ / + Model $\mathcal{C}$

Simulation results (cont)

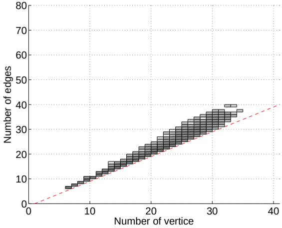


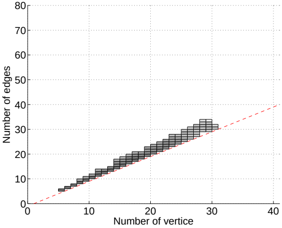
$--$ Imote / o Model $\mathcal{A}$ / $*$ Model $\mathcal{B}$ / + Model $\mathcal{C}$

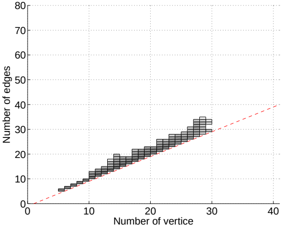
$\mathcal{A}$: sole contact and inter-contact duration fails

- the number of connected vertices is strongly over-estimated
- the number of connected components is under-estimated

Simulation results (cont)



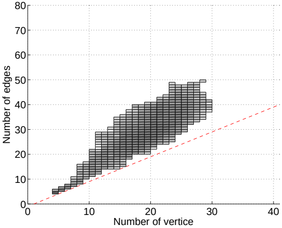


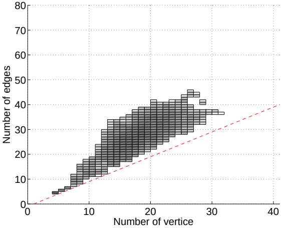


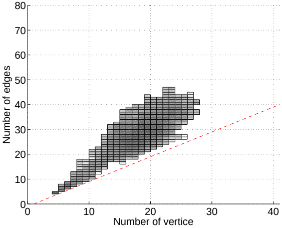
$\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$ fail!

- The density of the connected components (the groups) is underestimated
- Links are spread uniformly in the graph

Simulation results (cont)



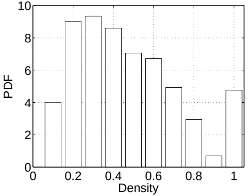




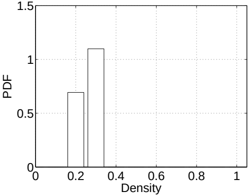
Weighted models

- does not have an impact on the contact and inter-contact duration distributions
- the density of connected components is comparable to the experimental data

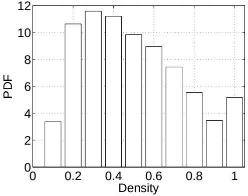
Simulation results (cont)



IMOTE$\mathcal{A}$



$\mathcal{B}_\omega$



Density of frequent connected components

- ($\tau = 7$ and $\sigma = 6$)
- classical models fail to create dense frequent connected components
- the number of frequent connected subgraphs is larger in the simulated data than in the original

Outline

- MOSAR Project
  - Project overview
- Dynamic Network Characterization
  - Motivation
  - Statistical analysis of snapshots of graphs
  - Towards a global analysis of the dynamics
  - Modeling of the dynamics
- **Conclusion**

Conclusion

contributions

- rigorous / coherent set of properties (basic / advanced)
- probability distribution of contacts and inter contacts is only one parameter
- global analyses to characterize the dynamics of the graph as a whole:
  - correlation between links
  - stability of the connected components
  - number of triangles
  - evolution of communities inside the interaction networks.
- simple / accurate models that generate random interaction graphs with satisfactory temporal properties.

Conclusion

Futur / On going works

- Introduce non-stationarity (piecewise stationary model)
- Dynamic community computation
- Overlapping community detection
- Trajectories of individuals as a *signature*
- Large in situ test beds to be deployed...

Some references

Dynamic networks

- Antoine Scherrer, Pierre Borgnat, Éric Fleury, Jean-Loup Guillaume and Céline Robardet, *Description and simulation of dynamic mobility networks*, in Computer Network 2008.
- Pierre Borgnat, Éric Fleury, Jean-Loup Guillaume, Céline Robardet and Antoine Scherrer, *Analysis of Dynamic Sensor Networks: Power Law Then What?*, in Comsware 2007.