CHIST-ERA - From data to new knowledge

COMPLEX NETWORKS EMERGING FROM LARGE DYNAMIC DATABASES

New mathematical and computational tools

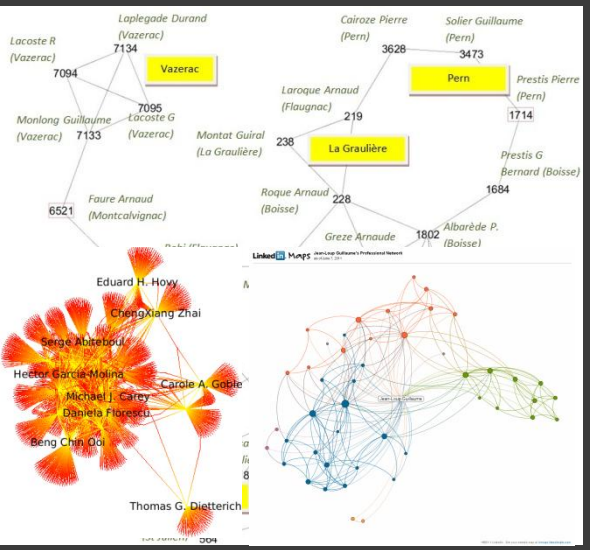
Eric Fleury
Professor ENS Lyon
IXXI

Bertrand Jouve
Professor Univ. Lyon
ERIC

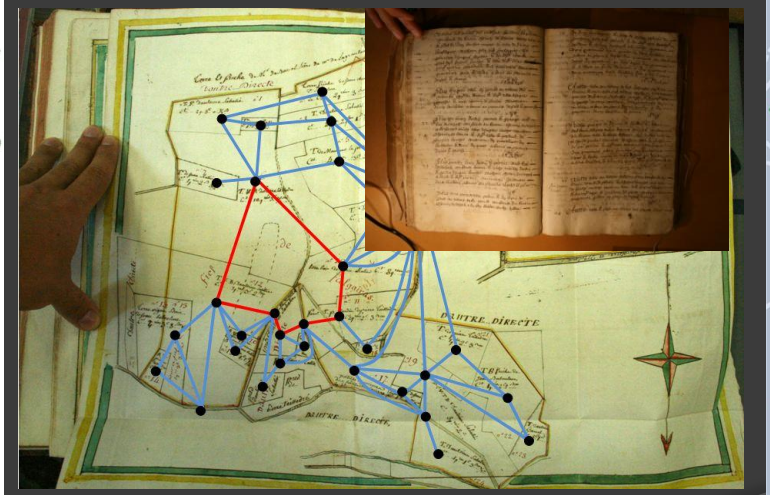
Matthieu Latapy
Senior Researcher
CNRS – LIP6 Paris

From large databases to complex networks

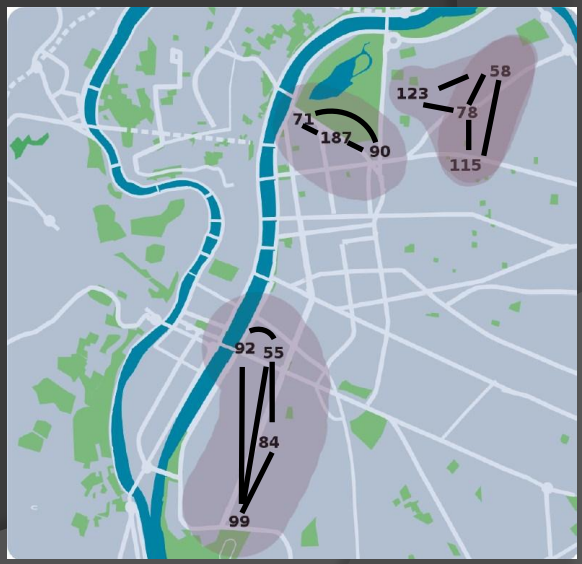
In many experimental data sets, some information can be modeled as interaction networks.



Social networks



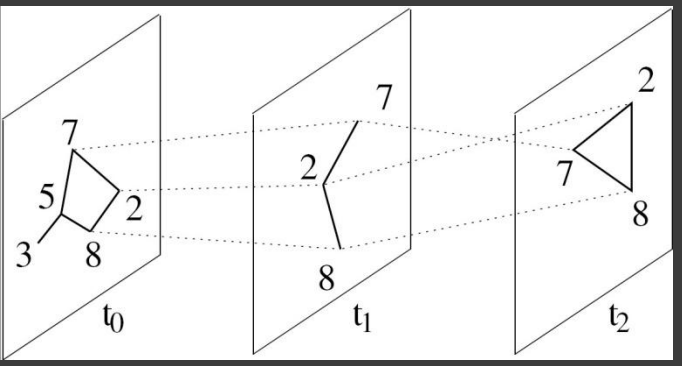
Landscape organisation



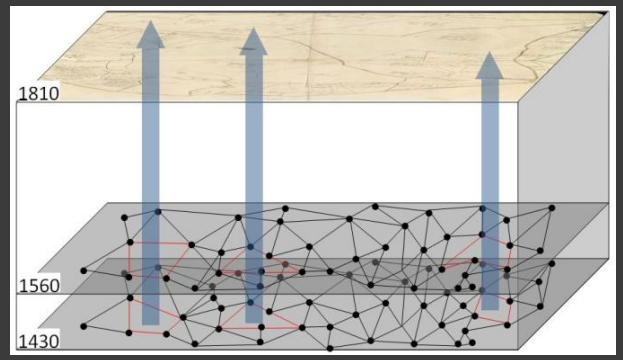
Velo'v

Dynamics properties of dated networks

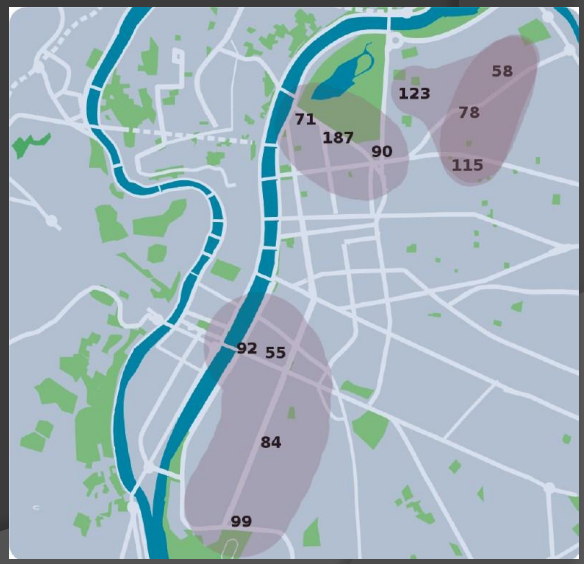
- The nodes are entities that may evolve over time (appearance, disappearance, change).
- The links are dated interactions



Social network evolution



dynamics of landscape fragmentation

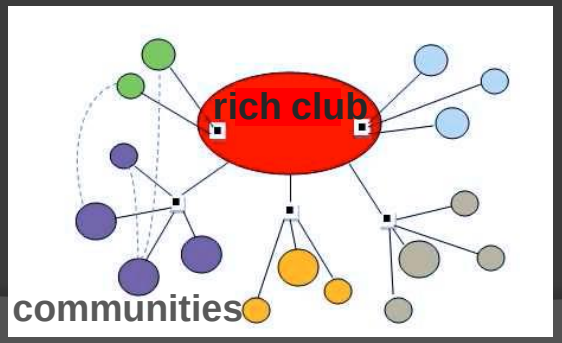


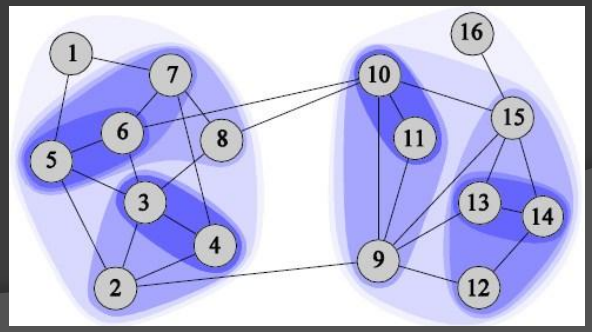
Velo'v

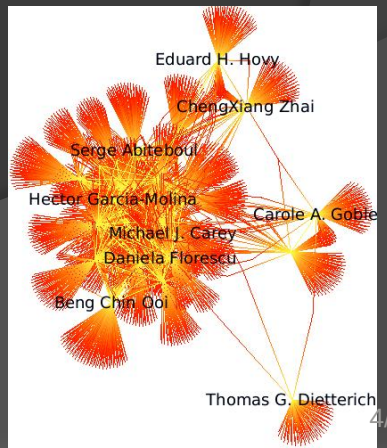
Multi-scale properties of networks : what is a ‘community’ ?

At a given time t_0 , a ‘community’ may be made of individuals (nodes) that have:

- similar neighbours (*i.e.* twins and complete subgraphs)
- similar neighbours outside the community (*i.e.* interval)
- the same graph distances with individuals outside the community (isometric nodes)
- …





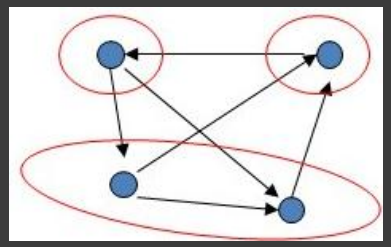


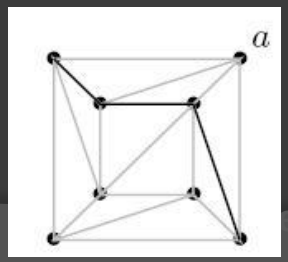
State of the art

Mathematical and computational tools provide a framework to explore and model the complex interaction networks emerging from large structured databases

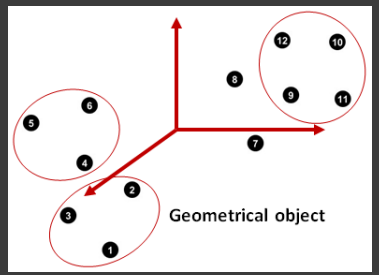
- *Static networks*

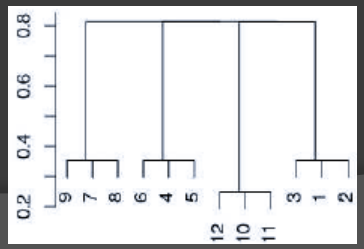
Graph theory



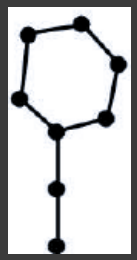


Data analysis





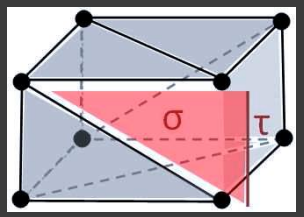
Algebra

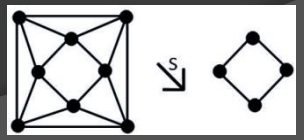


$$\begin{pmatrix} 0 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \cdots & 0 \end{pmatrix}$$

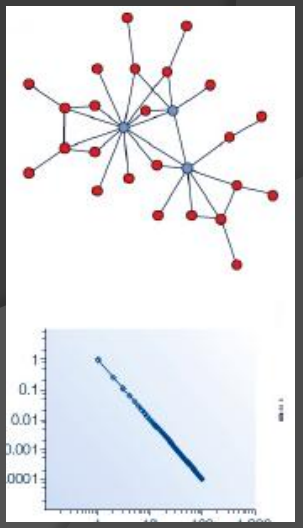
$$A \cdot u = \lambda u$$

Topology





Statistics & Probability

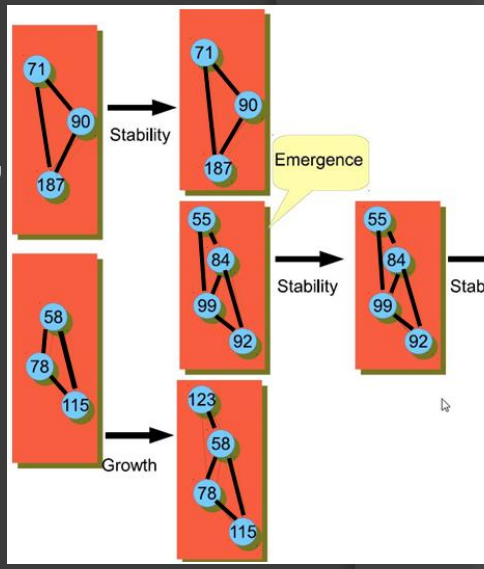


Scale free network

- *Dynamic networks*

Global approach: evolution of statistical indicators of the graph (degree, density, …).

Local approach: tracking patterns or communities, information spreading, detecting ‘events’.



Few established tools for the study and description of the dynamic :

- Most real-world networks share non-trivial static properties: is it the same for dynamic properties ?
- What is a community in a dynamic situation ? (multislice)
- Is it possible to adapt the mathematical tools to take account of time or should we create new ones ?

Scientific challenges

- Evaluate the limits of a disciplinary approach in the modeling of a complex network emerging from a large dated database (Human and Social Sciences).
- Define mathematical tools for modeling dynamics in multi-scale, hierarchical, and uncertain large networks emerging from real databases.
- Gather the scientific communities of data mining, graph theory and signal processing (statistical physics).

communities

Conclusion

- *Target outcome*

New mathematical and computational tools for modeling dynamics in real-world network, and, for instance, being able to define the notion of a community. They could be used to detect events (regular dynamics vs. anomalies) and predict evolution (link prediction, recommendation).

- *Expected impact*

- Enhancing interdisciplinary around the study of large databases.
- A better knowledge of large dynamic data in various domains.

communities