

Services et interactions en réseau (SIR)

Partie 2/3 Modèles

Luigi Lancieri

<http://www.lifl.fr/~lancieri/>

Plan

1. Introduction
2. Modèles statistiques et sociocognitifs
 - Self Similarité
 - Emergence et chaos
 - Structures en réseaux
3. *Mesure des interactions et Services adaptables*

Modélisation

- Réduction ou simplification de la réalité
- Objectif: comprendre et prévoir
- Mesurer (expérience) la réalité pour valider le modèle.
- Importance des statistiques et des probabilités

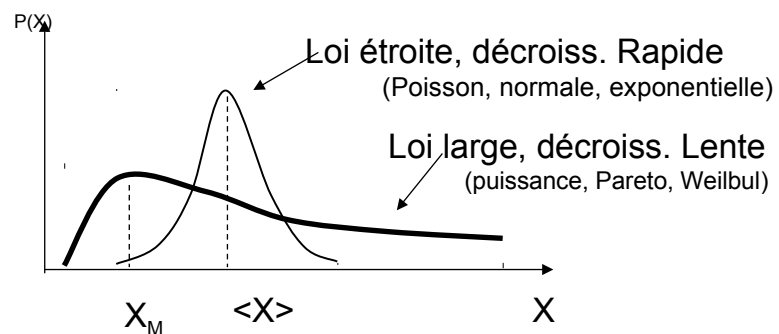
Self-similarité

Lois étroites et larges

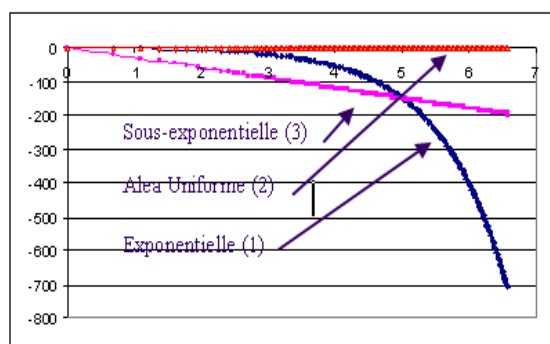
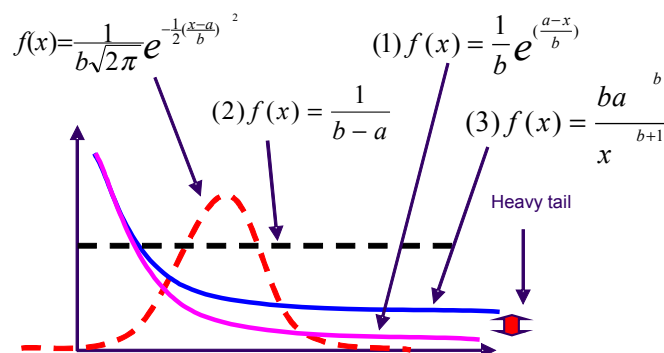
- Le choix du type de distribution est essentiel,
 - La moyenne et l'écart type n'ont pas toujours un sens.
 - Rôle du hasard plus ou moins important: relation avec les modèles chaotiques et fractals, (rôle de la mémoire)
- Type de décroissance et structure de répartition des événements.
 - Les événements rares ont plus ou moins d'importance

On découvre de plus en plus de phénomènes naturels suivants des lois larges.

- rythme cardiaque,
- distribution des populations,
- comportement des atomes, ..



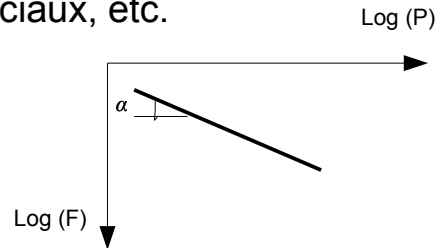
Rôle de la transformation logarithmique



Distribution sous exponentielle: La Loi de Zipf-Mandelbrot

- Log fréquence vs log rang est une « quasi » droite qui modélise la relation d'ordre entre les éléments
- Appliquée en linguistique, phénomènes sociaux, etc.

$$F = \frac{K}{P^{-\alpha}} \quad H = 1 + \frac{\alpha}{2}$$



- K=constante (= 1 dans un cas général)
- F=fréquence de l'élément (événement, objet, ..)
- P= Rang de l'élément
- α = pente de la droite caractéristique
- H=Paramètre de Hurst

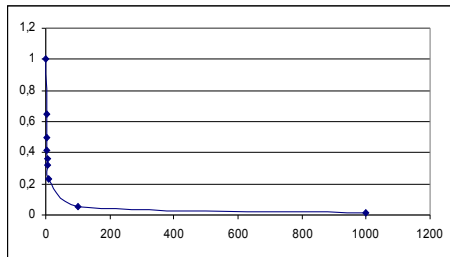
Étude de cas 1/2

- On considère qu'un groupe d'utilisateurs émet un flux de requêtes http distribué de manière sous-exponentielle depuis un intranet à destination du Web. Ce flux correspondant à un paramètre de Hurst de 0.75. On sait que l'utilisateur qui a été le plus actif a réalisé 1000 requêtes. On sait aussi que ce modèle nous permet de prévoir le nombre de requêtes des autres usagers en fonction de leur rang d'activité.
- Calculer:
 - 1) La pente de la droite caractéristique et la valeur de K
 - 2) Le nombre de requêtes réalisées par le second utilisateur le plus actif.
 - 3) Comparer l'écart d'activité (nb de requêtes) entre les 2 utilisateurs les plus actifs et les 5 et 6 em puis les 100 et 101 em utilisateurs les plus actifs.

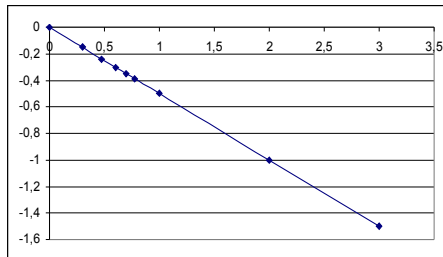
– Rappel :

$$F = \frac{K}{P^{-\alpha}} \quad H = 1 + \frac{\alpha}{2}$$

Étude de cas 2/2



1	1
2	0,70710678
3	0,57735027
4	0,5
5	0,4472136
6	0,40824829
10	0,31622777
100	0,1
1000	0,03162278

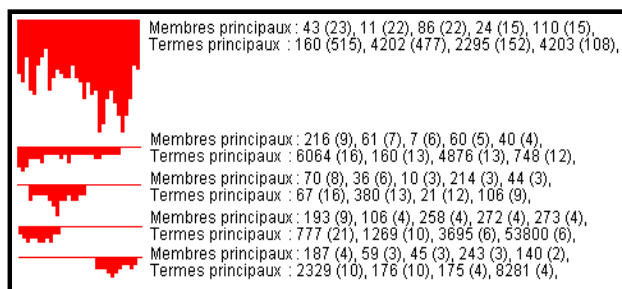
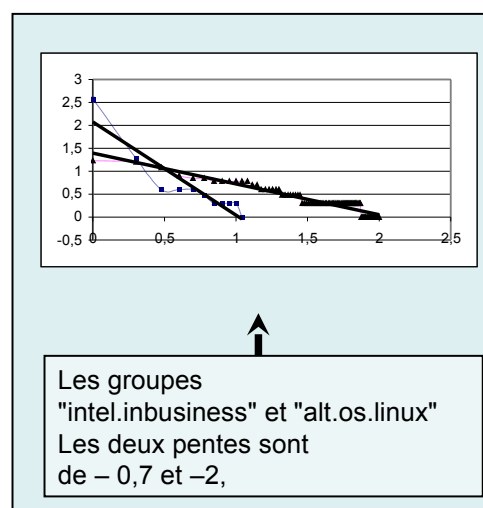
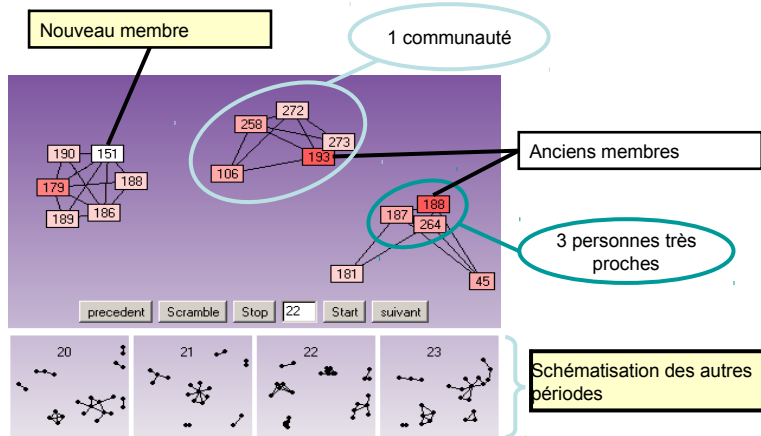


0	0
0,30103	-0,150515
0,47712125	-0,23856063
0,60205999	-0,30103
0,69897	-0,349485
0,77815125	-0,38907563
1	-0,5
2	-1
3	-1,5

Courbes
normalisées
pour K=1

- Alpha= **-0.5** (pente de la droite), $H=1.5/2=0.75$
- Si $U1=1000$ Req et $p=1$, $k=1000$; $U2=1000*(1/(2*0.5))=707$ Req; **delta $U1-U2=293$ Req**
- $U5=447$; $U6=408$; **delta $U5-U6=39$ Req**; $U100=100$; $U101=99$; **delta $U100-U101=1$ req**

Visualiser la structure des groupes et son évolution.

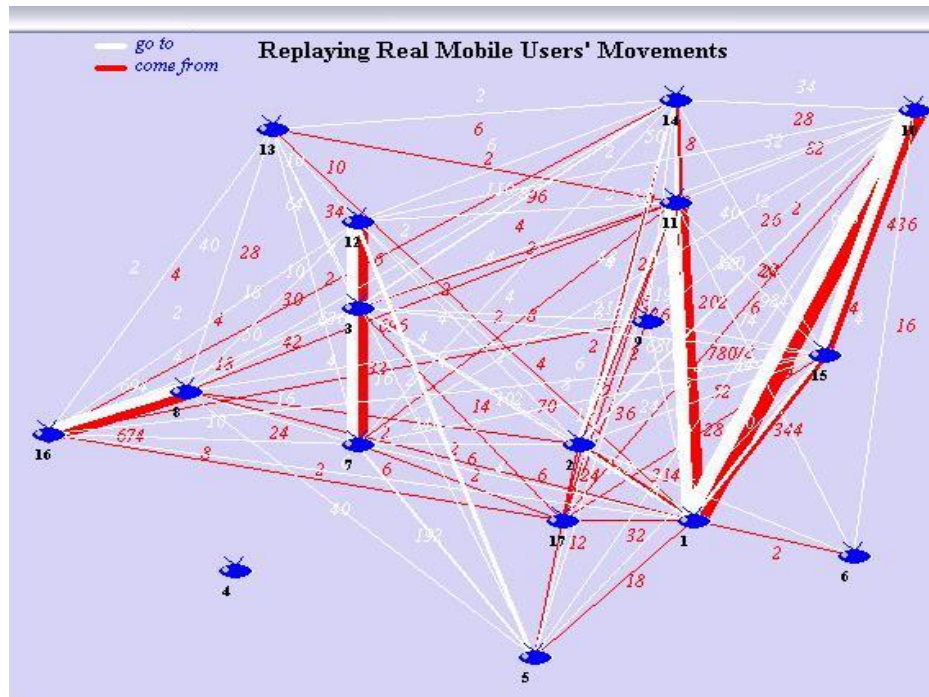


nombre de membres appartenant
à un groupe.

Membres les plus actifs et termes
les plus structurants.

Implication physique des comportements collectifs

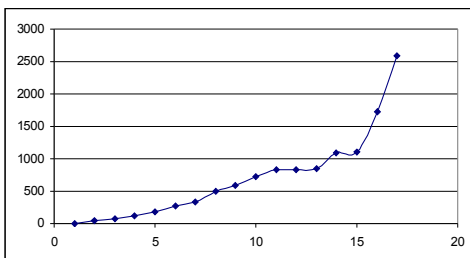
- 17 WIFI Spots
- 125 days period
- 240 mobile users.



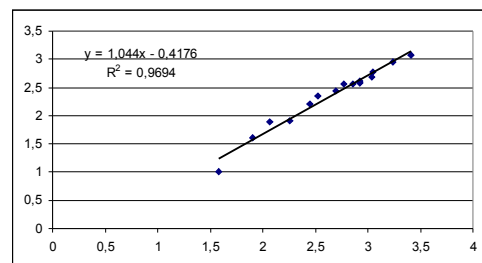
(Benayoune et Lancieri, 2006)

Structure d'occupation des espaces

- Les taux d'occupation des espaces classés par ordre d'importance suit une loi de puissance.



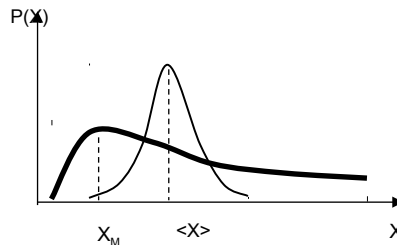
Les zones de l'espace par rang d'occupation (Lancieri 2007, InterSymp)



La pente de la transformée matérialise la structure d'occupation des espaces

Invariance d'échelle

- Même **caractéristique** à tous les niveaux d'échelles. (Scale-free, self similarité, fractale)
 - Accroissement (i.e comportement) **identique pour des événements rares et fréquents.**
- Perte de valeurs caractéristique (moyenne, écart type) d'un ensemble (relativité des échelles).
 - La taille des individus est de l'ordre de ????
 - Sur un segment important: répartition identique quelque soit la taille.



Invariance d'échelle et mémoire



- *IF fractal*
- *Fonctions itératives*
- $X_{n+1} = 0.2 X_n - 0.26 Y_n$
- $Y_{n+1} = 0.23 X_n + 0.22 Y_n + 1.6$

Où interviennent les phénomènes de mémoire ?

Transition de phase

Théorie du chaos

Constats

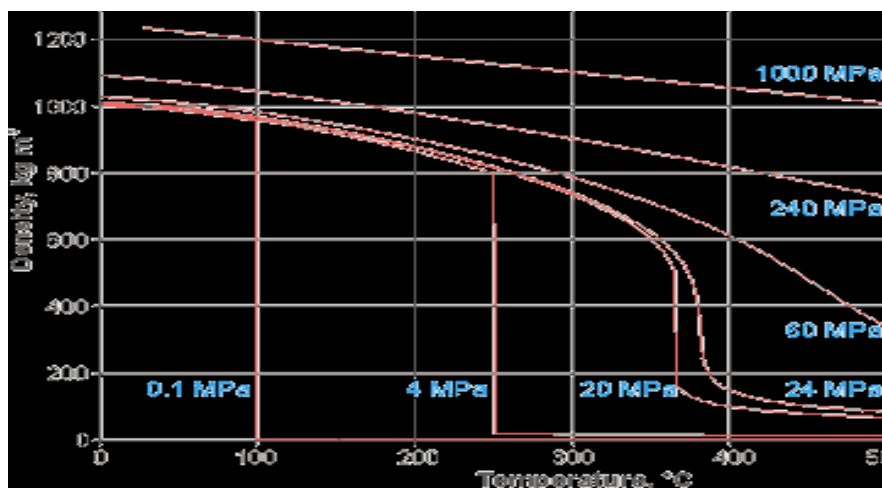
- En situation d'incertitude, les êtres vivants ont tendances à s'imiter (Herd behaviour)
- Les interactions de groupes et la technologie (le media) amplifient les conséquences de ce phénomène
- La perte d'information (oubli) dans les interactions aboutis à des erreurs et des excès liés aux interprétations approximatives ou à la rationalité limitée.
- La Self-similarité (symétrie) et le caractère émergent (le temps) des phénomènes sont des indices mesurables.

Chaos et Hasard

- Une faible variation dans les causes engendre de grandes variations dans les conséquences (e.g météorologie, cf travaux de Lorentz).
- A mi chemin entre phénomènes déterministes et purement aléatoires.
- Le hasard ne se définit plus comme binaire (présence ou absence) mais est continu (plus ou moins de hasard)

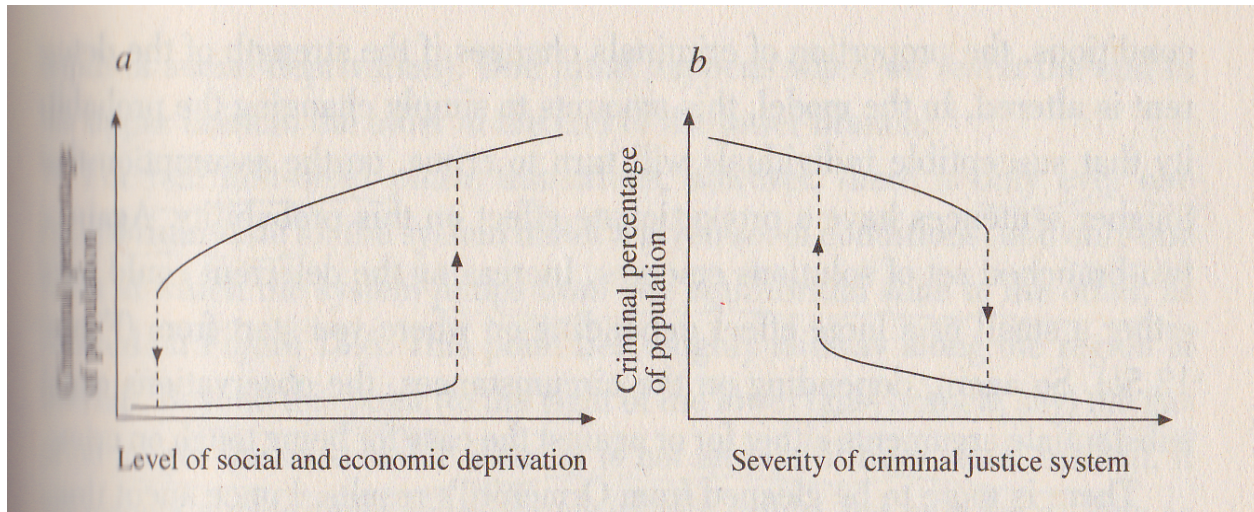
Transition de phase et phénomènes physiques

- Passage **abrupt** (point critique) de la phase liquide à la phase gazeuse, engendré par une très faible variation de température
- Densité vs température



Transition de phase et phénomènes sociaux

- Modèle de Campbell et Ormerod



Relation entre chaos et mémoire (voir équation logistique : état futur de la variable fonction de son état précédent)

Le rôle de l'interaction

- Les phénomènes de suivisme impliquent que les individus aient une perception mutuelle de leurs intentions.
- Le suivisme émerge par contagion par les interactions de proche en proche.
- Ce mode d'interactions distribuées peut être modélisé sous la forme d'un réseau.
- De nombreux systèmes s'apparentent à des réseaux : Langues, société, neurones, ...

Exemples

- 3 vidéos
 - Dans le règne animal
 - Occupation des espaces
 - Déplacement des individus

Dans le monde animal



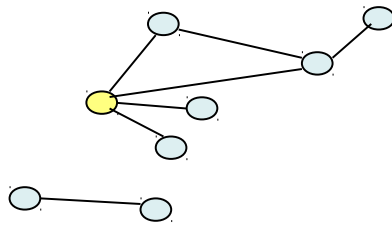
Occupation des espaces (Paris)



Déplacements (Paris)



Modélisation des réseaux



Les réseaux sociaux

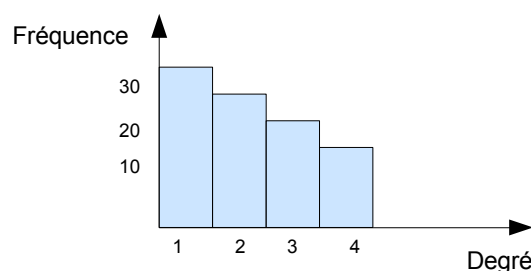
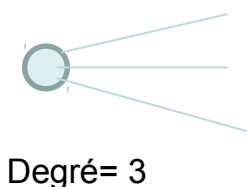
- Ensemble de relations entre des individus,
 - influencés par travaux de Simmel: *The Web of Group Affiliations* (1908), S. Burt, B. Wellman, Freeman, etc
 - Pluridisciplinarité: Sociologie, Psychologie sociale, anthropologie, sciences cognitive, Informatique.
 - Expérience de Stanley Milgram (1933-1984) Small worlds: La chaîne de contacts reliant 2 individus est en moyenne de 6.
- Métrologie, modélisation, visualisation
 - Notion de: cohésion, densité ; prestige ; intermédiarité ; cliques et n-cliques ; trous structuraux; équivalence structurale, popularité
- Applications
 - échanges de mails, P2P, systèmes de recommandations
 - Mise en relation entre individus, voir Node(www.dotnode.com), Friendset(www.friendset.com), www.agorami.com, small world project(ref)

De la théorie des graphes aux phénomènes sociaux

- Origine de la théorie des graphes au 18^{em} siècle Leonhard Euler
- Modèle des graphes aléatoire en 1960 par Paul Erdős et Alfred Rényi.
- De nombreuses propriétés des graphes aléatoires **émergent** relativement soudainement.
- Principales métriques : Degré, parcours, diamètre, coefficient de regroupement, ..

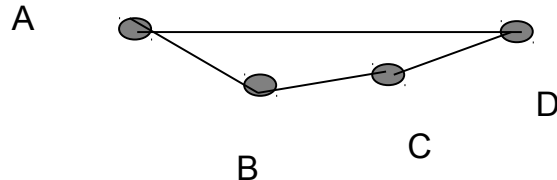
Degré d'un nœud ou d'un réseau

- Nombre de liens associés à un nœud.
- $P(k)$: Probabilité qu'un nœud ait **k** liens.
- Dans un réseau aléatoire P de la plupart des nœuds est le degré moyen: $\langle k \rangle$ (Moyenne des $P(k)$ de chaque nœud)
- Ce n'est pas le cas dans un réseau réel



Parcours et diamètre

- **Parcours (L):** Chemin moyen (nombre de liens) entre une paire quelconque de nœuds (Parcours AD=2).



- **Diamètre (D):** Chemin maximal entre une paire quelconque de nœuds.
- Dans un réseau aléatoire D et L sont proches.

Coefficient de regroupement

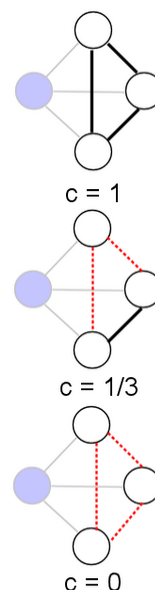
- $C_i = \text{Ratio } [0-1]$ entre :
 - nombre de relations existantes (**Re**)
 - nombre de relations possibles entre les n voisins d'un nœud i .

$$C_i = \text{Re} \cdot \frac{2}{n(n-1)} \quad \text{Graphes non orientés}$$

$$C_i = \frac{\text{Re}}{n(n-1)} \quad \text{Graphes orientés}$$

$$C = \frac{1}{n} \sum_{i=1}^n C_i \quad \text{Pour tout un réseau : C est la moyenne des } C_i$$

$C(k)$: C Moyen pour un réseau restreint à k nœuds

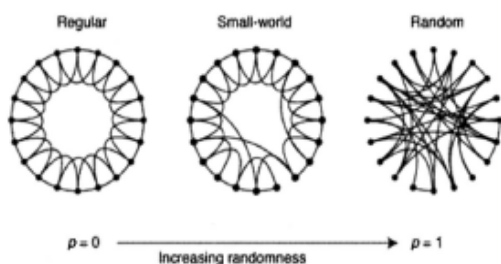


Les réseaux réels

Entre organisation aléatoire et structurée

Le modèle de Watts et Strogatz

- Modèle simple (1998)
 - issu du principe des interactions sociales (i.e. dominantes locales)
 - Permet une interpolation progressive entre un graphe aléatoire et une structure ordonnée.



Principe:

-**P=0**: Chaque nœud d'un réseau de N nœuds est connecté à ses k plus proches voisins.

-**P croissant**: Progressivement, on reconnecte aléatoirement une fraction des nœuds.

-**P=1**: les nœuds se retrouvent aléatoirement connectés

-**P** est la probabilité de reconnections

Le modèle de Watts et Strogatz

- On étudie l'évolution du coefficient de regroupement (C) et du chemin moyen (L) en fonction de p .

Conclusions

Quand le réseau est fortement organisé

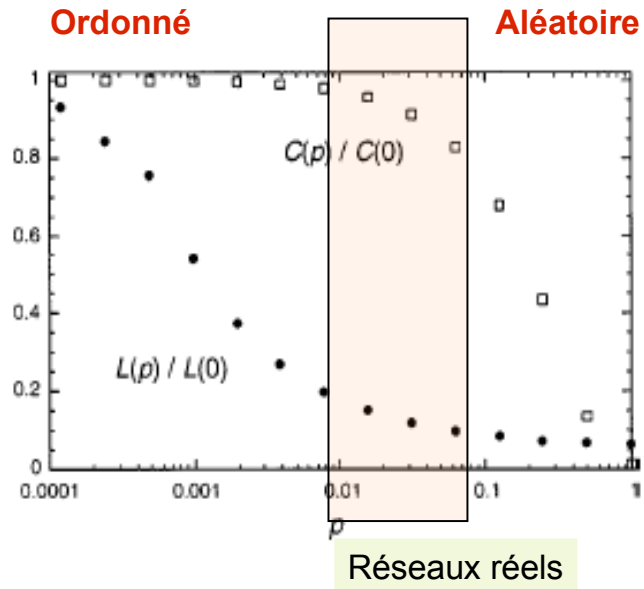
- L tends à décroître
- C est plutôt stable et élevé

Quand le réseau est plutôt aléatoire ($p > 1$)

- L et plutôt stable et faible
- C chute rapidement

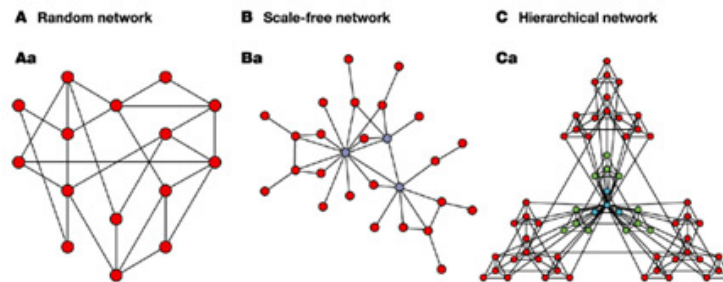
Les réseaux réels

- Coexistence de L faible et C élevé

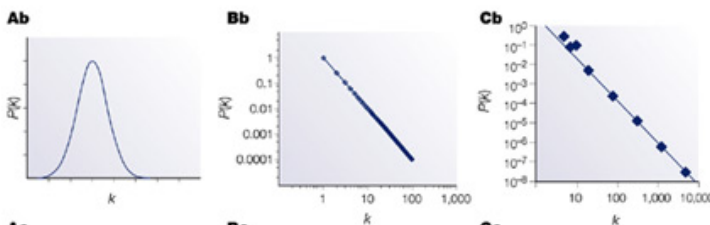


Propriétés des réseaux

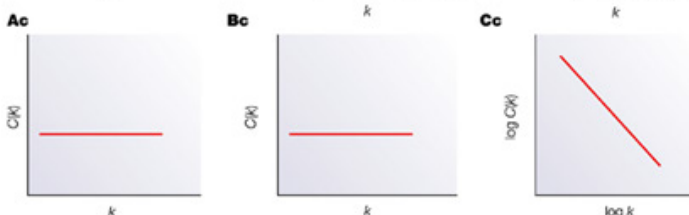
Structure



Degré



Coef. regroupmt



Selon Barabási et al, Nature (2004)

Exemples de réseaux réels

Ramener le cas à étudier à un réseau

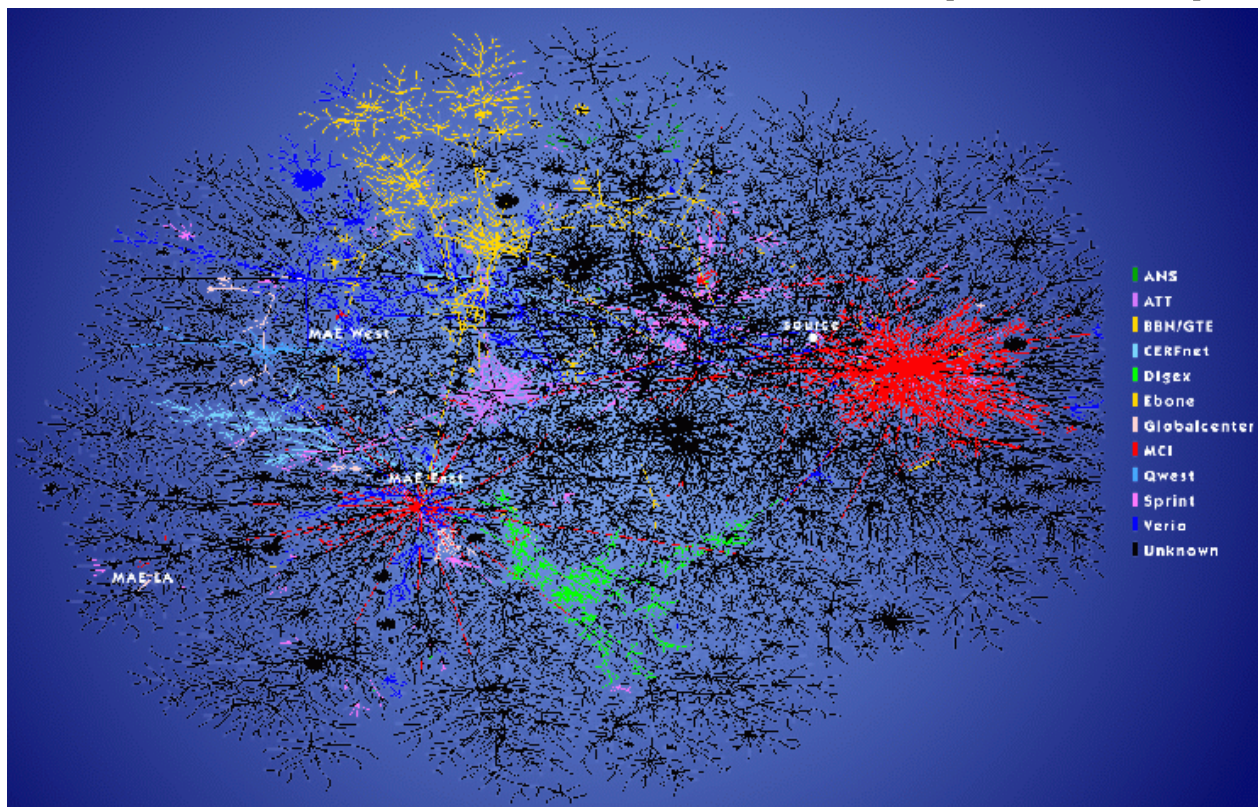
- **Web**: Réseau de pages web interconnectées par des hyperliens.
- **Réseau social** : l'évolution des contacts entre individus.
- **Réseau d'acteurs** : Les acteurs sont connectés s'ils ont joué ensemble dans un même film (cf Internet Movie Database)
- **Réseau de citations scientifiques** : Les articles se faisant référence
- **Réseau de coauteurs** : Deux auteurs sont connectés s'ils ont écrit un article ensemble
- **Réseau linguistique** : 2 mots sont liés s'ils sont proches
- **Réseau cellulaire** : nœuds représentent les substrats moléculaires communs eg (H_2O) avec leurs réactions chimiques dominantes.
- **Réseau écologique** : Relations proies prédateurs,
- **Réseau d'appel téléphonique** : les liens d'appels entre numéros de téléphones
- **Réseau électrique** : Générateurs ou transformateurs reliés par des lignes à haute et moyenne tension

Comparaison des caractéristiques

Réseau	Alpha	C	Grand
Web	-2,4	0,2	10^{-4}
Social	-3,4		
Acteurs	-2,3	0,8	$3 \cdot 10^{-4}$
Citations	-3		
Coauteurs	-1,9	0,4	10^{-4}
Linguistique	-2,75	0,43	10^{-4}
Cellulaire	-2,2	0,4	0,05
Écologique	-1,1		
Téléphone	-2,1		

(Ref Barabasi et al)

Connectivité entre ISPs (Skitter)



Remarques

- La structure des phénomènes (e.g loi d'occupation des espaces) est souvent self similaire.
- La dynamique et la propagation des phénomènes fait souvent intervenir des transitions abruptes de phase.
- Le niveau de mimétisme peut être anticipé par la mesure de α .