

Théorie des réseaux

Chapitre 6 : Densité, clustering et assortativité



La densité

Le clustering

L'assortativité

Ce qu'il faut retenir

La densité

Part des liens réalisés

$$\rho = \frac{m}{\binom{n}{2}} = \frac{2m}{n(n-1)}$$

Elle vaut 1 pour un graphe complet, 0 pour un graphe sans arête.

Sur notre réseau

$$\rho = \frac{2 \times 8}{7 \times 6} = \frac{16}{42} \approx 0,381$$

Un peu plus d'un tiers des liaisons possibles existent.

Pourquoi elle est trompeuse sur les grands réseaux

Une décroissance mécanique

Le dénominateur croît en n^2 , le numérateur rarement plus vite que n .

Tout grand réseau réel a une densité proche de zéro : Facebook compte quelques centaines d'amis pour trois milliards d'inscrits, soit $\rho \approx 10^{-7}$.

La conséquence

Comparer les densités de deux réseaux de tailles différentes n'a **aucun sens**.

C'est le degré moyen $\bar{k} = 2m/n$ qu'il faut comparer : il ne dépend pas mécaniquement de la taille.

Le clustering

Deux définitions à ne pas confondre

Local, par nœud

$$C_i = \frac{\text{aretes entre les voisins de } i}{\binom{k_i}{2}}$$

C'est la probabilité que deux voisins de i soient eux-mêmes voisins.

Global, ou transitivité

$$T = \frac{3 \times \text{nombre de triangles}}{\text{nombre de triples connectes}}$$

La différence

Le clustering **moyen** est la moyenne des C_i : chaque nœud pèse pareil, donc les petits degrés dominant.

La **transitivité** pèse chaque nœud par $\binom{k_i}{2}$: les hubs dominant. **Les deux nombres diffèrent souvent beaucoup.**

Les sept clusterings locaux

Noëud	Voisins	Arêtes entre eux	$\binom{k_i}{2}$	C_i
1	2, 3	1	1	1
2	1, 3	1	1	1
3	1, 2, 4	1	3	1/3
4	3, 5	0	1	0
5	4, 6, 7	1	3	1/3
6	5, 7	1	1	1
7	5, 6	1	1	1

Les deux résumés

$$\bar{C} = \frac{1 + 1 + \frac{1}{3} + 0 + \frac{1}{3} + 1 + 1}{7} = \frac{14/3}{7} \approx \mathbf{0,667}$$

Clustering nul, intermédiarité maximale

$C_4 = 0$: ses deux voisins ne se connaissent pas.

C'est exactement la définition du **trou structural**. Le **clustering local d'un nœud et son intermédiarité varient en sens inverse** — un pont ne peut pas être dans un triangle.

Une mesure vaut par ce qu'elle sépare

Trois mesures différentes — intermédiarité maximale, clustering nul, degré minimal — désignent **le même trait** du nœud 4 sous trois angles.

C'est ainsi qu'on valide un **diagnostic** : par la convergence de mesures indépendantes.

L'assortativité

Corrélation des degrés

$$r = \text{corr}(k_i, k_j) \quad \text{sur l'ensemble des aretes } ij$$

$$r \in [-1, 1].$$

Deux régimes

$r > 0$: réseau **assortatif** — les gros degrés se lient aux gros degrés.

$r < 0$: réseau **dissortatif** — les gros se lient aux petits, structure en étoile.

Les régularités empiriques

Type de réseau	r typique
Réseaux sociaux	+0,1 à +0,3
Internet, physique	−0,19
Web	−0,07
Interbancaire	−0,2 à −0,3

La régularité, et son interprétation

Les réseaux **sociaux** sont assortatifs : on se lie à ses semblables — c'est l'homophilie.

Les réseaux **techniques et financiers** sont dissortatifs : les petits se branchent sur les gros. **Cela produit une structure cœur-périphérie**, dont le chapitre 14 tirera le risque systémique.

Deux conséquences opposées

Un réseau **assortatif** a un cœur dense de hubs : une épidémie y démarre vite, mais reste longtemps confinée au cœur.

Un réseau **dissortatif** propage plus vite vers la périphérie, et se fragmente plus facilement si l'on retire les hubs.

La conséquence pour la régulation

Sur un réseau dissortatif, **protéger quelques nœuds protège tout le système.**

Sur un réseau assortatif, il n'y a pas de cible privilégiée : il faut une politique uniforme. **La structure dicte l'instrument.**

Ce qu'il faut retenir

Six points

1. La **densité** $\rho = 2m/n(n-1)$ vaut 0,381 ici, mais est incomparable entre tailles.
2. C'est le **degré moyen** qu'on compare, pas la densité.
3. **Clustering moyen** 0,667 et **transitivité** 0,545 ne mesurent pas la même chose.
4. $C_4 = 0$: un **pont n'est jamais dans un triangle**.
5. L'**assortativité** est positive dans les réseaux sociaux, négative dans les réseaux financiers.
6. Un réseau assortatif a un **cœur** : c'est là que porte la régulation.

La suite

Le chapitre 7 revient aux distances, et pose le fait empirique le plus célèbre du domaine : six degrés de séparation entre deux humains quelconques.