La science des réseaux
Une approche pluri-disciplinaire

Fabien Tarissan

CNRS – ENS Paris Saclay – ISP

Plan

Introduction au réseaux

Les réseaux

Définition : Ensemble d'entités reliées entre elles par des interactions.

Un cadre d'étude très générique :

- Informatique : Internet, pair-à-pair, web, ...
- Biologie : cerveau, gènes, protéines, éco-systèmes, ...
- Sciences sociales : collaboration, amitié, échanges, contacts, ...
- et bien d'autres : réseaux économiques, linguistiques, de transports, . . .

Ces réseaux partagent des caractéristiques :

- Faible densité
- Distribution des relations très hétérogènes
- Phénomène dit du ≪petit monde≫
- ...

Amène à reconsidérer les approches traditionnelles

Problématiques communes

Mesure : comment acquérir de l'information ?

- Graphe réel $\mapsto$ vue partielle
- Représentativité de l'échantillon ? Biais ?

Analyse : comment décrire le réseau ?

- Représentation des données
- Métriques pertinentes
- Propriétés partagées par d'autres réseau ?

Modélisation : comment se forme un réseau ?

- Génération aléatoire de structures similaires (avec propriétés observées)
- Mécanismes sous-jacents
- Support pour des simulations
- Explication des propriétés observées

... et plein d'autres : dynamique des réseaux, problèmes algorithmiques, ...

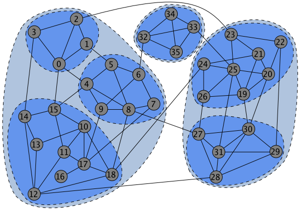
Problématique – Analyse

Décrire, extraire de l'information pertinent

statistiques

densité
degrés
densité locale
corrélations
...

structure



Graphe

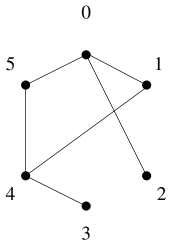
Graphe

Un graphe $G = (V, E)$ est un couple d'ensembles.

- V est l'ensemble des *sommets* (ou *nœuds*)
- $E \subseteq (V \times V)$ est l'ensemble des *arêtes* (ou *liens*).

Exemple :

- $V = \{0, 1, 2, 3, 4, 5\}$
- $E = \{(0, 1), (0, 2), (3, 4), (4, 5), (5, 0), (1, 4)\}$



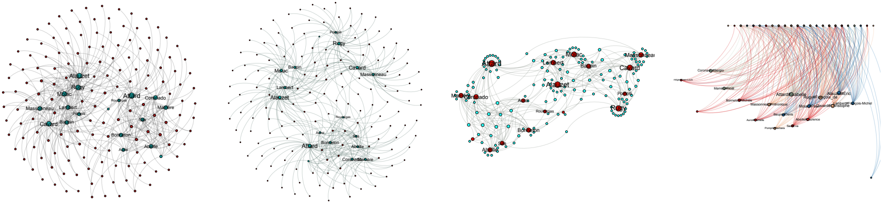
Graphe

Graphe

Un graphe $G = (V, E)$ est un couple d'ensembles.

- V est l'ensemble des *sommets* (ou *nœuds*)
- $E \subseteq (V \times V)$ est l'ensemble des *arêtes* (ou *liens*).

Attention à ne pas confondre le graphe avec sa représentation !



Plusieurs types de réseaux

On peut classer les réseaux en grands groupes (qui se recoupent) :

Graphes orientés *vs.* non orientés

Y a-t-il une asymétrie dans la relation ?

- Graphe non-orienté : $(u, v) = (v, u)$
- Graphe orienté : $(u, v) \neq (v, u)$

Graphe pondérés *vs.* non pondérés

Y a-t-il des liens plus importants que d'autres ?

- Graphe non pondérés : $G = (V, E)$
- Graphe pondérés : $G = (V, E, w : E \mapsto \mathbb{R}^+)$ où $w(u, v)$ décrit le poids du lien (u, v)

Plusieurs types de réseaux

Réseaux complets *vs* égo-centrés

- Réseaux complets : on connaît l'ensemble des nœuds et l'ensemble des relations entre ces nœuds.
- Réseaux égo-centrés : on s'intéresse à un nœud en particulier et à son voisinage.

Le *mode* du réseau

Combien de type distincts d'entités dans le réseau ?

- Réseaux one-mode : graphes uni-partis (1 seul ensemble)
- Réseaux two-mode : graphes bipartis (2 ensembles)
- ... etc. graphes multi-partis : ...
- Réseaux multiplexes : 1 ensemble de nœuds mais plusieurs types de liens

Représentation des données

Matrice d'adjacence

$$
\begin{array}{c c}
 & \begin{array}{cccc} 0 & 1 & \dots & n-1 \end{array} \\
\begin{array}{c} 0 \\ 1 \\ \dots \\ n-1 \end{array} & \begin{pmatrix} & & & \\ & & \dots & \\ & & & \\ & & & \end{pmatrix}
\end{array}
$$

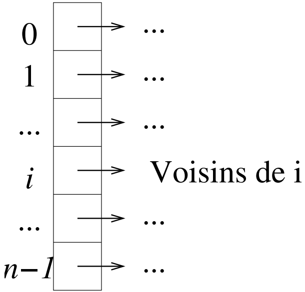
Case (i, j) :

- 1 s'il y a une arête entre i et j
- 0 sinon

Variante avec poids ...

Représentation des données

Listes d'adjacence



Taille d'un graphe

Notations

- n = le nombre de sommets
- m = le nombre d'arêtes

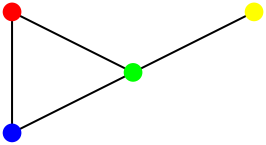
u et v sont voisins s'il y a une arête entre eux.

Définitions

- Degré : $d^\circ(v)$: nombre de voisins de v
- degré moyen du graphe $d^\circ(G)$: moyenne des degrés de tous les sommets
- densité du graphe, δ : probabilité d'existence de tout lien

$$\delta = \frac{\text{nombre de liens}}{\text{nombre de liens possibles}} = \frac{m}{\frac{n(n-1)}{2}}$$

Exemple



$n = 4,\ m = 4,$

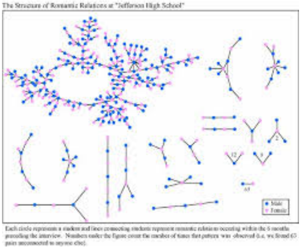
degrés : **2, 2, 3, 1** ; degré moyen 2

$\delta = \frac{8}{12} = 0.66..$

Quelques exemples de densité

- Réseaux de neurones : $n \sim 100 * 10^9$, $m \sim 50\,000 * 10^9$, $\delta \sim 10^{-7}$
- Réseau web : $n \sim 10 * 10^9$, $m \sim 4 * 10^9$, $\delta \sim 10^{-9}$

Notion de chemin



Chemin

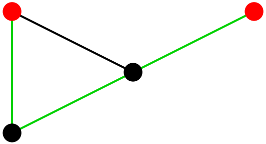
Chemin de u à v : suite d'arêtes $(u, v_1), (v_1, v_2), \ldots, (v_{k-1}, v)$
Longueur $= k$ (nombre d'arêtes)

Connexité

Composante connexe : ensemble maximal de sommets t. q. il existe un chemin entre toutes les paires de sommets.
Graphe connexe : une seule composante connexe

Distance

chemin de u à v = suite d'arêtes u...v

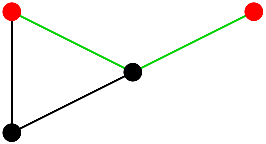


un chemin de longueur 3

Distance

chemin de u à v = suite d'arêtes u...v

distance $d(u, v)$ = longueur d'*un* plus court chemin



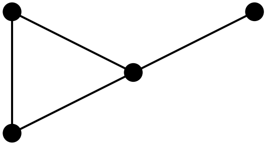
un plus court chemin ; longueur 2 $\Rightarrow$ distance $= 2$

Distance

chemin de u à v = suite d'arêtes u...v

distance $d(u, v)$ = longueur d'*un* plus court chemin

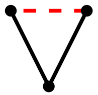
diamètre Δ = plus grande distance



diamètre = 2

Densités locales

Densité autour des nœuds : probabilité que deux voisins d'un nœud soient reliés



Coefficient de clustering

Défini pour un nœud v :

$$\mathrm{cc}(v) = \frac{\Delta(v)}{\vee(v)}$$

Le coefficient de clustering du graphe est alors la moyenne de cette valeur

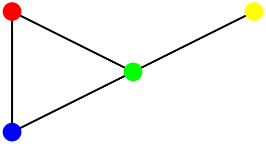
Ratio de transitivité

Défini pour le graphe (calcul global) :

$$\mathrm{tr}(G) = \frac{3\Delta(G)}{\vee(G)}$$

Coefficient de clustering

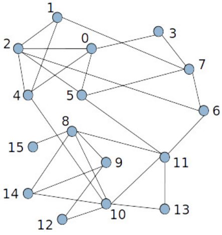
Exemple :



Coefficients de clustering : 1, 1, $\frac{1}{3}$, indéfini

Coefficient de clustering du graphe : $\mathrm{cc}(G) \sim 0.78$
Ratio de transitivité du graphe : $\mathrm{tr}(G) = 0.6$

Définir les communautés



But : Identifier (automatiquement) des *groupes pertinents*.

Défi :

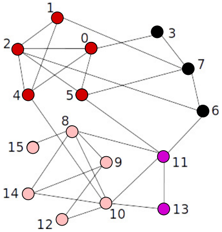
- On ne connaît pas le nombre
- On ne connaît pas la taille
- (doit passer à l'échelle. Exemple sur un graphe à 300 sommets ...)

Algorithmes

- Beaucoup d'approches possibles : percolation, marche aléatoire, k-core, ...
- Algorithme de Louvain : très efficace, passe à l'échelle. Basé sur la notion de modularité :

$$Q = \frac{1}{2m} \sum_{C} e_c - \frac{a_C^2}{2m}$$

Définir les communautés



But : Identifier (automatiquement) des *groupes pertinents*.

Défi :

- On ne connaît pas le nombre
- On ne connaît pas la taille
- (doit passer à l'échelle. Exemple sur un graphe à 300 sommets ...)

Algorithmes

- Beaucoup d'approches possibles : percolation, marche aléatoire, k-core, ...
- Algorithme de Louvain : très efficace, passe à l'échelle. Basé sur la notion de modularité :

$$Q = \frac{1}{2m} \sum_C e_c - \frac{a_C^2}{2m}$$

Distributions

Moyenne de valeurs : vision très (trop ?) agrégée de l'ensemble des valeurs.

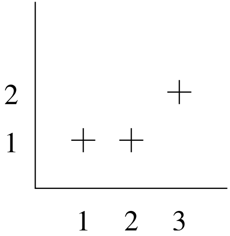
Distribution : manière synthétique de représenter une série de valeur.
$\rightarrow$ combien de fois la valeur x apparaît dans la série ?

Distributions

Moyenne de valeurs : vision très (trop ?) agrégée de l'ensemble des valeurs.

Distribution : manière synthétique de représenter une série de valeur.
→ combien de fois la valeur x apparaît dans la série ?

Exemple/rappel avec la distribution des degrés :
4 nœuds, degrés : 2 3 3 1



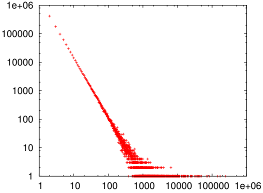
Caractérisation de la distribution

L'un des intérêt : permet de caractériser qualitativement la série de valeurs.

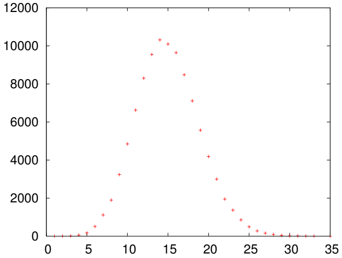
Loi de puissance

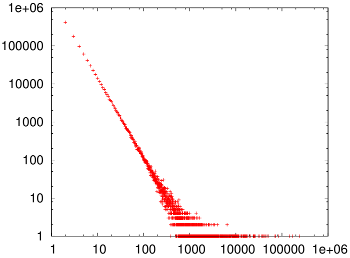
- $N_k \sim k^{-\alpha}$
- droite en échelle log-log

Distribution hétérogène : non homogène (comportements variés), en pratique, souvent proche d'une loi de puissance



hétérogène *vs* homogène





Homogène

Notion de normalité (et d'exceptions)

Hétérogène

Tous les comportements existent

→ pas de notion de normalité

Distributions normalisées

Distribution des valeurs, deux choix :

- N_k : nombre d'occurrences de la valeur k
- p_k : proportion de la valeur k dans la série

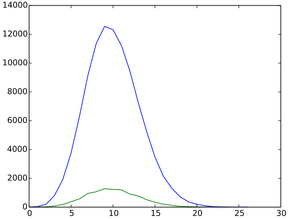
$\rightarrow$ Distribution normalisée

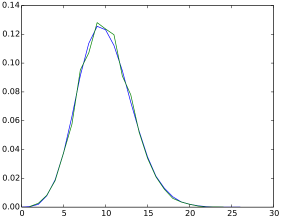
$$p_k = \frac{N_k}{n}$$

A priori, simplement une modification de la valeur en y.

Distributions normalisées

Permet de comparer des graphes de tailles différentes :

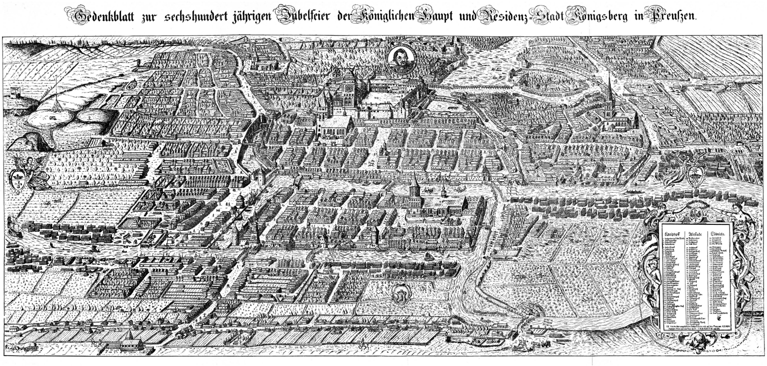




Les réseaux dans les sciences

Plan

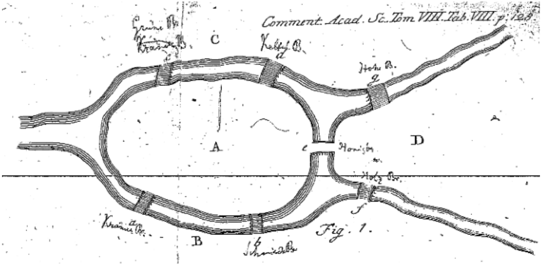
Les sept ponts de Königsberg



Un casse-tête (milieu XVIIIè siècle)

Est-il possible de parcourir la ville et revenir à son point de départ en empruntant tous les ponts une seule fois ?

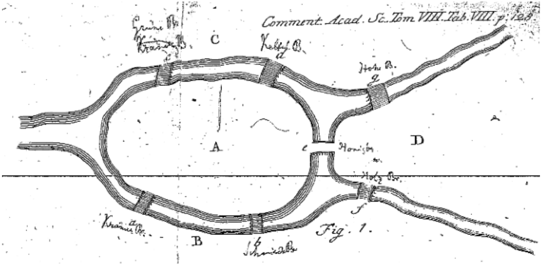
Les sept ponts de Königsberg



Approche de Leonhard Euler (mathématicien suisse)

- Considère la ville comme un graphe

Les sept ponts de Königsberg



Approche de Leonhard Euler (mathématicien suisse)

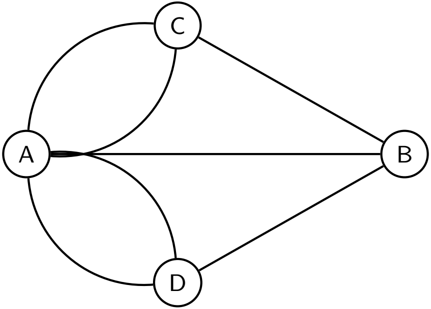
- Considère la ville comme un graphe non orienté et non pondéré

  $\Longrightarrow$ V : les parties de la ville ; E : les ponts
- La question devient :

  Existe-t-il un chemin partant et arrivant à un sommet (ie. un cycle) tel que chacune des arêtes est utilisée une et une seule fois ?

$\Longrightarrow$ Existe-t-il un cycle eulérien ?

Les sept ponts de Königsberg



Réponse de Leonhard Euler en 1735 : seul la géométrie des connexions compte !

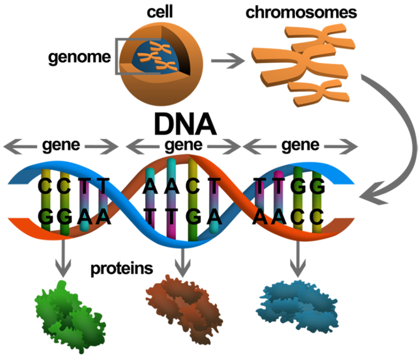
- Généralise le problème à tout graphe
- Étudie les conditions pour qu'un cycle eulérien existe :
  - condition nécessaire et suffisante : tous les degrés doivent être pairs

$\Longrightarrow$ il n'existe pas de promenade possible dans Königsberg !

Biologie moléculaire

Vocabulaire

- ADN (acide désoxyribo nucléique) : A / T / G / C
- Gène (simplifié) : séquence de nucléotide codant pour une protéine
- Protéine : molécule, dont la fonction dépend de la séquence, de la cellule, ...



Biologie moléculaire

Quelques éléments

- très peu de séquences sont reconnues
  ($\sim$ 30 000 gènes pour 3 milliards de nucléotides chez l'être humain)
- tous les gènes ne sont pas actifs ensemble $\Longrightarrow$ *profil d'activité*
- le profil d'activité est lié au rôle de la cellule : cellule sanguine (globule rouge) $\neq$ cellule du système nerveux (neurone)
- mais toutes les cellules possèdent exactement le même patrimoine génétique (ie. les mêmes gènes)

Ce qui soulève des questions :

Biologie moléculaire

Quelques éléments

- très peu de séquences sont reconnues ($\sim$ 30 000 gènes pour 3 milliards de nucléotides chez l'être humain)
- tous les gènes ne sont pas actifs ensemble $\Longrightarrow$ *profil d'activité*
- le profil d'activité est lié au rôle de la cellule : cellule sanguine (globule rouge) $\neq$ cellule du système nerveux (neurone)
- mais toutes les cellules possèdent exactement le même patrimoine génétique (ie. les mêmes gènes)

Ce qui soulève des questions :

Phénomène de différenciation

Comment expliquer que des cellules qui ont le même matériel génétique arrivent à avoir des profils d'activité différents ?

Phénomène de régulation

Comment stopper (réguler) la production d'une protéine lorsqu'un gène est actif ?

Réseau génétique

Les protéines régulatrices

Certaines protéines ont un effet sur l'activité des gènes

- activatrice : active la production de la protéine codée par un gène
- inhibitrice : arrête la production de la protéine codée par un gène

Réseaux de régulation génétique

Graphe

Réseau génétique

Les protéines régulatrices

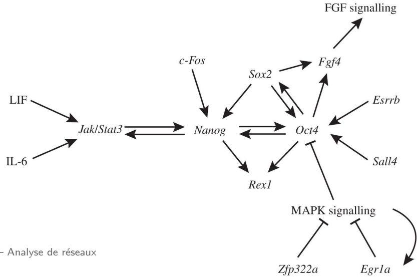
Certaines protéines ont un effet sur l'activité des gènes

- activatrice : active la production de la protéine codée par un gène
- inhibitrice : arrête la production de la protéine codée par un gène

Réseaux de régulation génétique

Graphe orienté non pondéré et signé :

- V : l'ensemble des gènes
- $(u, v, [+/-]) \in E$: l'activité de u a un effet $(+/-)$ sur celle de v.



Réseau génétique

Travaux de René Thomas (biologiste belge) dans les années 60

Étudier les cycles (comme Euler !) dans les réseaux de régulations. En particulier :

- cycles négatifs : nombre impair de liens négatifs
- cycles positifs : nombre pair de liens négatifs

Distinction rudimentaire ... mais suffisante !

Réseau génétique

Travaux de René Thomas (biologiste belge) dans les années 60

Étudier les cycles (comme Euler !) dans les réseaux de régulations. En particulier :

- cycles négatifs : nombre impair de liens négatifs
- cycles positifs : nombre pair de liens négatifs

Distinction rudimentaire ... mais suffisante !

Différenciation

Plusieurs profils d'activités $\Longrightarrow$ existence de circuits positifs

Régulation

Oscillations autour d'une valeur $\Longrightarrow$ existence de circuits négatifs

René Thomas, "On the relation between the logical structure of systems and their ability to generate multiple steady states or sustained oscillations", in *Numerical Methods in the Study of Critical Phenomena*, Springer Series in Synergetics, vol. 9, 1981, p. 180-193
René Thomas et **Denis Thieffry**, "Les boucles de rétroaction, rouages des réseaux de régulation biologiques", in *Médecine/Sciences*, vol. 11, n° 2, 1995, p. 189-198.

Les réseaux en sociologie

Origine des réseaux en sociologie

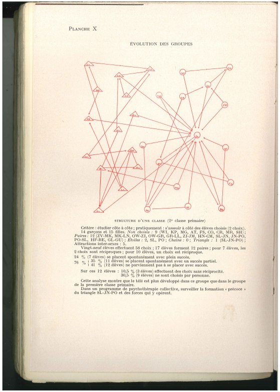
Parenté multiple ⟶ voir le cours 2 de Baptiste Coulmont

- individus, groupe, organismes, ... : au centre de la sociologie
- Georges Simmel : met en évidence l'importance des relations deux à deux
- Idée exploitée par la sociométrie (milieu XXè siècle)

Jacob Moreno (médecin/psychiatre américain)

Méthode d'enquête : le test sociométrique

- une même question posée aux enquêtés : *avec qui aimeriez vous ... ?*
- exprime les phénomènes d'attraction/répulsion
- visualisation des réseaux : les sociogrammes



Jacob Levy Moreno, *Who Shall Survive ? A New Approach to the Problem of Human Interrelations*, 1934.
Traduction française dans *Fondements de la sociométrie*, PUF, 1970.

La force des liens faibles

Mark Granovetter (sociologue américain)

Les relations entretenues entre 2 individus peuvent revêtir :

- une forte charge émotionnelle : famille, amour, ...
  $\longrightarrow$ rencontres fréquentes
- une faible charge émotionnelle : collègues, amis d'enfance, ...
  $\longrightarrow$ rencontres peu fréquentes

La force des liens faibles

- Intuition que les liens faibles sont moins importants

La force des liens faibles

Mark Granovetter (sociologue américain)

Les relations entretenues entre 2 individus peuvent revêtir :

- une forte charge émotionnelle : famille, amour, ...
  $\longrightarrow$ rencontres fréquentes
- une faible charge émotionnelle : collègues, amis d'enfance, ...
  $\longrightarrow$ rencontres peu fréquentes

La force des liens faibles

- Intuition que les liens faibles sont moins importants
- Or ce sont eux qui élargissent le cercle social des individus : c'est leur force
- Étude des liens portant connaissance d'un emploi (Boston) :
  $\longrightarrow$16.7% seulement par des liens forts.

Mark S. Granovetter, "The strength of weak ties", *American Journal of Sociology*, vol. 78, n$^\circ$ 6, 1973, p. 1360-1380
Traduction française dans l'ouvrage *Sociologie économique*, Seuil, 2008.

La force des liens faibles

The Strength of Weak Ties

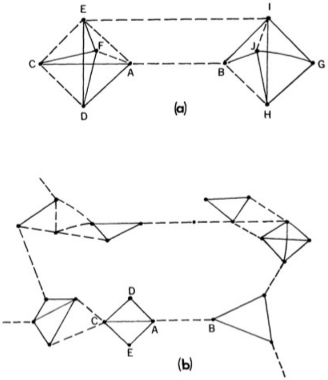


FIG. 2.—Local bridges. *a*, Degree 3; *b*, Degree 13. ——— = strong tie; – – – = weak tie.

Conclusion

Que montrent ces exemples ?

Conclusion

Que montrent ces exemples ?

Réseau = liens

- Euler $\Longrightarrow$ degré des sommets
- Thomas $\Longrightarrow$ circuits positifs/négatifs
- Granovetter $\Longrightarrow$ force des liens

Conclusion

Que montrent ces exemples ?

Réseau = liens

- Euler $\Longrightarrow$ degré des sommets
- Thomas $\Longrightarrow$ circuits positifs/négatifs
- Granovetter $\Longrightarrow$ force des liens

Réseau = cadre d'analyse

- Ne pas minimiser le rôle des sommets
- Focus sur le rôle des liens dans l'émergence d'un phénomène
- Abstraction de la réalité mais cadre qui suffit à expliquer un phénomène

$\Longrightarrow$ Besoin d'une science des réseaux

La science des réseaux

Une démarche scientifique nouvelle

Des propriétés en communs

- Réseaux de nature différentes partagent des propriétés structurelles
- Les raisons de la formation des réseaux n'est pas disciplinaire.
- Besoin de chercher des explications indépendamment de la nature réelle des réseaux

$\Longrightarrow$ De nouvelles questions

Mise en évidence de propriétés

- distances courtes
- faible densité
- forte densité locale
- ...

Recherche de modèles explicatifs

- Comment les réseaux se forment ?
- Pourquoi s'organisent-ils sous cette forme ?

Contraste entre des propriétés globales et des interaction locales !
$\Longrightarrow$ Propriétés émergentes

2 exemples :

1. Réseaux petit-monde (*Nature* 1998)
2. Réseaux sans échelle (*Science* 1999, *Nature* 2000)

Les six degrés de séparation

Diffusion rapide de l'information

- existence de chemin courts entre les nœuds d'un réseau ...
- ... et les nœuds sont capable de les découvrir

Expérience de Stanlay Milgram (psychologue américain)

- 300 lettres qui doivent parvenir à 1 personne
- le destinataire : agent de change de Boston
- 300 participants (100 de Boston / 100 du Nebraska / 100 courtiers du Nebraska)
- ne doit transmettre la lettre qu'à une connaissance directe.

Les six degrés de séparation

29% de lettres qui arrivent à destination : longueur moyenne de 5.2 intermédiaires
$\Longrightarrow$ 6 degrés de séparations

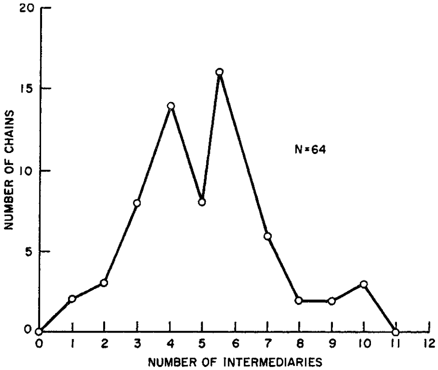


FIGURE 1

Jeffrey Travers and **Stanlay Milgram**, "An experimental study of the small world problem", in *Sociometry*, vol. 32(4), 1969, p. 425-443.

Les réseaux petit-monde

Étude de Duncan Watts (sociologue) et Steven Strogatz (mathématicien)

Étude empirique de 3 réseaux de nature différentes

- réseau biologique : réseau de neurones (ver *C. Elegans*)
- infrastructure : réseau électrique (d'une partie) des États-Unis
- réseau social : réseau de collaboration entre acteurs de films

	n	m	L_{actual}	L_{random}	C_{actual}	C_{random}
Film actors	225 226	6 869 393	3.65	2.99	0.79	0.00027
Power grid	4 941	6 596	18.7	12.4	0.080	0.005
C. elegans	282	1974	2.65	2.25	0.28	0.05

Les réseaux petit-monde

Étude de Duncan Watts (sociologue) et Steven Strogatz (mathématicien)

Étude empirique de 3 réseaux de nature différentes

- réseau biologique : réseau de neurones (ver *C. Elegans*)
- infrastructure : réseau électrique (d'une partie) des États-Unis
- réseau social : réseau de collaboration entre acteurs de films

	n	m	L_{actual}	L_{random}	C_{actual}	C_{random}
Film actors	225 226	6 869 393	3.65	2.99	0.79	0.00027
Power grid	4 941	6 596	18.7	12.4	0.080	0.005
C. elegans	282	1974	2.65	2.25	0.28	0.05

Les réseaux présentent des distances courtes et une densité locale forte
$\Longrightarrow$ réseaux "petit-monde"

Quels mécanismes ?

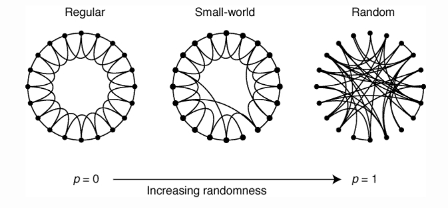
Or ces deux propriétés sont *a priori* incompatibles !

Modèles classiques

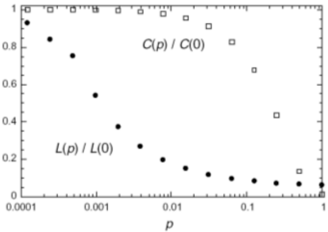
- Graphes Erdös-Rényi (aléatoires) : distances courtes mais densité locale faible
- Graphes k-réguliers : densité locale forte mais distances élevées

Modèles de Watts-Strogatz

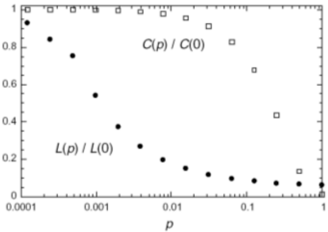
On part d'un graphe k-régulier et on réarrange aléatoirement les liens avec probabilité p (paramètre du modèle).



Le modèle de Watts-Strogatz



Le modèle de Watts-Strogatz



Résultats

Avec une très faible valeur de $p \in [0.001 : 0.01]$ (ie. peu de liens aléatoires) on obtient des graphes ayant les deux propriétés (graphes petit-monde).

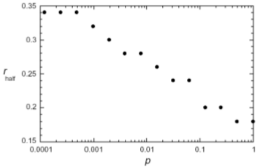
Interprétation

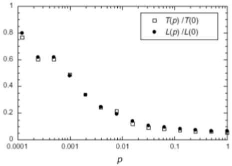
- les liens s'organisent avant tout de manière locale ($\mapsto$ densité locale forte)
- l'aléatoire a la faculté de générer des ponts entre région distantes ($\mapsto$ distances moyennes faibles)

Quels bénéfices ?

Modèle de diffusion épidémiologie (modèle SIR)

La propagation se fait de proche en proche en fonction de la viralité de l'épidémie.
$\longrightarrow$ Étude de l'impact des propriétés structurelles sur la diffusion

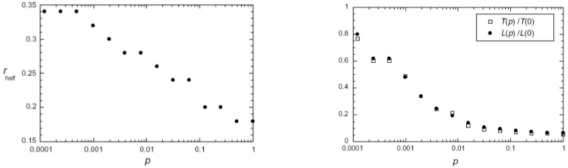




Quels bénéfices ?

Modèle de diffusion épidémiologie (modèle SIR)

La propagation se fait de proche en proche en fonction de la viralité de l'épidémie.
$\longrightarrow$ Étude de l'impact des propriétés structurelles sur la diffusion



Résultats

1. Plus il y a de liens aléatoire, moins la viralité doit être forte
2. À viralité constante, la diffusion est plus efficace en présence de liens aléatoires

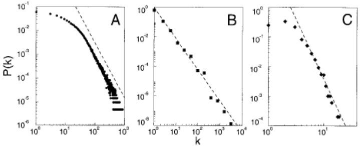
Duncan J. Watts et **Steven H. Strogatz**, "Collective dynamics of 'small-world' networks", *Nature*, vol. 393, n°6634, 1998, p. 440-442.

Les réseaux sans échelle

Étude des physicien.nes Albert-Lázló Barabási et Réka Albert

Tous les nœuds ont le même degré. Est-ce réaliste ?

- réseau de collaboration d'acteurs
- réseau du web
- réseau électrique américain



Résultats

1. Distribution des degrés hétérogènes (proche d'une loi de Puissance).
2. Majorité de nœuds de faible degré
3. Existence de hubs

$\Longrightarrow$ Propriété d'invariance d'échelle $\Longrightarrow$ Réseaux sans échelle

Modèle d'attachement préférentiel

Un modèle pour expliquer l'invariance ?

Défaut majeur des modèles existants : il sont statiques !

- graphes aléatoires
- graphes k-réguliers
- modèle de Watts-Strogatz

Or les réseaux ont une dynamique de croissance (web, collaborations scientifiques, ...)
$\Longrightarrow$ Comment s'insèrent les nouveaux nœuds ?

Modèle de Barabási-Albert

Hypothèse simple mais réaliste : les nœuds se lient au nœuds existants proportionnellement à leur degré.

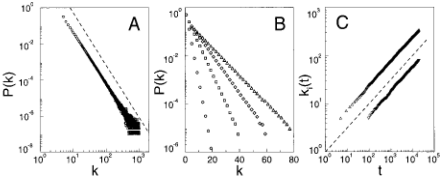
$$\prod(k_i) = \frac{k_i}{\sum_j k_j}$$

L'attachement préférentiel

Effet vertueux du modèle

Plus un nœud a un degré élevé, plus il va attirer de nouveaux nœuds
... et donc augmenter son degré !

$\longrightarrow$ Matthew effect, "rich get richer", ...



Résultat

Le modèle de Barab'asi-Albert génère des graphes sans échelle

Albert-László Barabási et **Réka Albert**, "Emergence of scaling in random networks", *Science*, vol. 286, n° 5439, 1999, p. 509-512.

Quels rôles jouent les *hubs* ?

Modèle de pannes

- aléatoires : on retire les nœuds aléatoirement
- ciblées : on retire les nœuds de plus fort degré

Modèle de graphe

- graphes aléatoires
- graphes sans échelle

$\Longrightarrow$ Quel impact sur la distance moyenne dans les réseaux ?

Quels rôles jouent les *hubs* ?

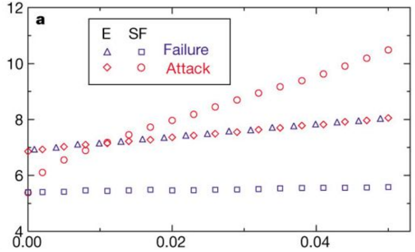
Modèle de pannes

- aléatoires : on retire les nœuds aléatoirement
- ciblées : on retire les nœuds de plus fort degré

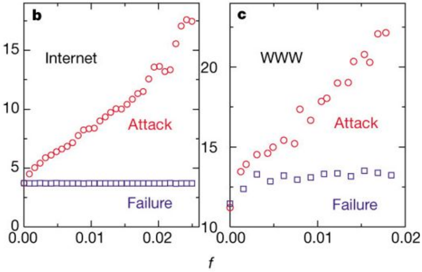
Modèle de graphe

- graphes aléatoires
- graphes sans échelle

$\Longrightarrow$ Quel impact sur la distance moyenne dans les réseaux ?



Quels rôles jouent les *hubs* ?



Résultat

L'existence de *hubs* rend les réseaux :

- robustes vis à vis des pannes aléatoires
- vulnérables à des attaques ciblées

Réka Albert, Hawoong Jeong et **Albert-László Barabási**, "Error and attack tolerance of complex networks", *Nature*, vol. 406, n° 6794, 2000, p. 378-382.

Conclusions

Propriétés des réseaux

Un langage pour les décrire les réseaux : les graphes

- nœuds, liens
- degré, densité
- chemin, longueur, composante connexe, diamètre
- densité locale, coefficient de clustering, ratio de transivité
- communauté

Propriétés des réseaux

La plupart des réseaux ont des propriétés communes.
On dispose de modèles pour expliquer l'émergence de ces propriétés.

	réseaux	aléatoire	k-régulier	WS	AB
densité	faible	faible	faible	faible	faible
connexité	comp géante	comp géante	comp géante	comp géante	comp géante
distances	courtes	courtes	longues	courtes	courtes
degrés	hétérogènes	homogènes	homogènes	homogènes	hétérogènes
clustering	fort	faible	forte	forte	faible

école normale supérieure paris—saclay

Science des réseaux

Une nouvelle démarche scientifique

1. Recherche de propriétés communes aux réseaux
2. Identification des mécanismes permettant l'émergence de ces propriétés
3. Mise en évidence de l'intérêt de ces propriétés pour les réseaux

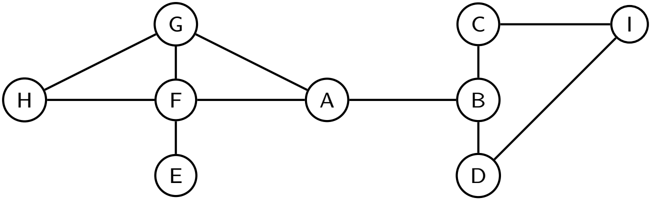
Réseaux petit-monde

- Propriété de petit-monde : distances courtes et densité locale forte
- Modèle de Watts-Strogatz : quelques liens aléatoires dans un graphe k-régulier
- Bénéfice : diffusion rapide de l'information

Réseaux sans échelle

- Propriété d'invariance d'échelle : distribution des degrés hétérogènes
- Modèle de Barab'asi-Albert : phénomène d'attachement préférentiel
- Bénéfice : robustesse face au pannes aléatoires

Mesures de centralité (suite)



Différente mesure de la *centralité* d'un nœud dans un réseau. Centralité de :

degré : $C_{deg}(u) = d(u)$

proximité : $C_{prox}(u) = \frac{1}{\sum_{v \neq u} dist(u,v)}$

intermédiarité : $C_{inter}(u) = \sum_{s \neq t \neq u} \frac{\sigma_{st}(u)}{\sigma_{st}}$ où

- σ_{st} : nb de plus courts chemins entre s à t
- $\sigma_{st}(u)$: nb de plus courts chemins entre s and t **passant par** u

...

Questions ?

http://tarissan.complexnetworks.fr/