

Que le monde est petit !

Nous sommes tous à quatre poignées de main du Président de la République !
Une catégorie particulière de graphes aléatoires était passée inaperçue :
les « petits mondes ».

Les gens que vous connaissez ont plus de chances de se connaître entre eux que des gens pris au hasard. De plus, alors que chaque Français ne connaît qu'un petit nombre de personnes au total, chacun est à quatre poignées de main ou moins du Président de la République : chacun a serré la main de quelqu'un qui a serré la main de quelqu'un qui a serré la main de quelqu'un qui a serré la main du Président de la République. Ce sont là des remarques que les sociologues qui étudient les réseaux de relations sociales ont formulées depuis longtemps. Certains soupçonnent même, sans disposer de preuve, que chaque humain est à six poignées de main au plus de chaque autre humain.

Les réseaux de relations sociales sont extrêmement complexes et irréguliers, mais ils n'ont pas les caractéristiques des graphes aléatoires classiques tels que les mathématiciens les ont étudiés, ni bien sûr celles des réseaux géométriques réguliers. Les réseaux des relations sociales sont des graphes de type « petit monde ». Ce nom a pour origine l'exclamation que vous proférez lorsque vous découvrez que votre voisin dans l'avion connaît les parents de votre beau-frère, ou que vous êtes à trois poignées de main d'une célébrité qui vous semblait lointaine : « Mon Dieu, que le monde est petit ! »

Le développement de bases de données volumineuses et d'algorithmes pour les explorer systématiquement a permis de découvrir que presque tous les gros graphes (c'est-à-dire comportant plusieurs milliers de nœuds) tirés du monde réel possèdent une série de caractéristiques remarquables et inattendues semblables à celles des réseaux de relations sociales. En particulier, chacun de ces graphes a un petit diamètre – on va rapidement d'un point à un autre – et est « aggloméré » – si B et C sont reliés à A alors il est assez probable que B et C sont reliés l'un à l'autre.

Dans un article paru en 1998, Duncan Watts et Steven Strogatz attirèrent l'attention sur ce phénomène étrange qui a donné lieu à des centaines de travaux scientifiques concernant toutes sortes de domaines différents. D. Watts et S. Strogatz avaient pris pour exemples trois gros graphes du monde réel dont la connaissance détaillée était devenue possible (ce qui n'est toujours pas le cas du

réseau complet correspondant aux poignées de mains) : le premier de ces graphes concernait les acteurs de cinéma ayant tourné des films ensemble, le second représentait le système de distribution électrique en Amérique du Nord et le troisième était celui des interconnexions entre les neurones du ver nématode *Caenorhabditis elegans*.

Assez étrangement, la structure de ces trois graphes présente des similarités qui ont depuis été retrouvées dans un grand nombre d'autres situations du monde réel : réseaux des collaborations scientifiques, réseau des échanges de e-mail, réseau physique d'Internet, réseau des liens entre pages web, réseaux des contacts sexuels, réseaux des interactions entre protéines dans une cellule, etc. Il semble que, pour des raisons encore un peu floues, les gros graphes du monde réel sont régis par des règles communes que, depuis cinq ans, la communauté scientifique tente de préciser. Toutes ces recherches sont importantes entre autres choses pour le contrôle des épidémies (humaines ou de virus logiciels), la conception de systèmes robustes de communications, la compréhension des réseaux de neurones ou des interactions entre les cellules et composants des cellules et la mise au point des moteurs de recherche d'Internet.

Nœuds, arcs et degrés

Pour décrire ces structures et définir ce qu'est un graphe de type « petit monde », nous allons considérer les paramètres associés à un graphe et qui permettent d'en parler globalement sans en connaître le détail.

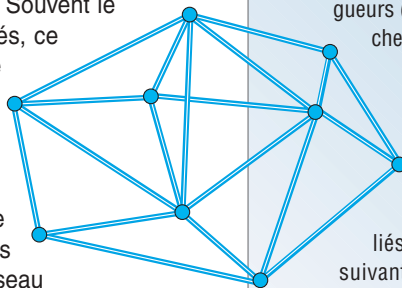
Puisqu'un graphe (ou réseau) est un ensemble de nœuds (ou sommets) liés par des arcs (ou liens, ou arêtes), les deux premiers paramètres importants d'un graphe sont le nombre de nœuds n et le nombre d'arcs m . Sauf mention contraire, les arcs que nous considérerons seront non orientés : l'arc AB est identique à l'arc BA.

Pour les graphes qui nous concernent, n est supérieur à 100. Remarquons que dans un graphe s'il y a m arcs et n nœuds alors cela fait en moyenne $2m/n$ arcs attachés à chaque nœud (chaque arc possède deux extrémités). Le paramètre $2m/n$ sera noté k , et est nommé degré moyen du graphe. Sur la figure 1 est dessiné un graphe de $n=9$ nœuds, $m=19$ arcs et ayant donc un degré moyen $k=4,22$.



©Corbis/L. Langevin

S'il y a n nœuds et que chaque nœud est lié à chaque autre (sauf lui-même) cela fait un total de $m = n(n - 1)/2$ arcs. Dans un tel graphe, nommé graphe complet, le degré moyen k vaut $n - 1$, ce qui est beaucoup. S'il y a 1 000 nœuds, le graphe complet comporte 499 500 arcs. Les graphes du monde réel sont très rarement des graphes complets et ont un nombre bien plus petit d'arcs. Souvent le degré moyen k est de quelques unités, ce qui signifie que chaque nœud est relié à un nombre limité d'autres nœuds. Ainsi, le degré moyen vaut 3,92 pour le graphe des coopérations entre mathématiciens (un arc relie deux mathématiciens, s'ils ont écrit un article ensemble), il vaut environ 8 pour les réseaux de neurones et 5,98 pour le réseau physique Internet.



Graphes réguliers et aléatoires

On distingue trois sortes de graphes : les graphes réguliers, les graphes aléatoires classiques et les graphes de type « petit monde » qui sont, comme on va le voir, intermédiaires.

1. Les paramètres d'un graphe

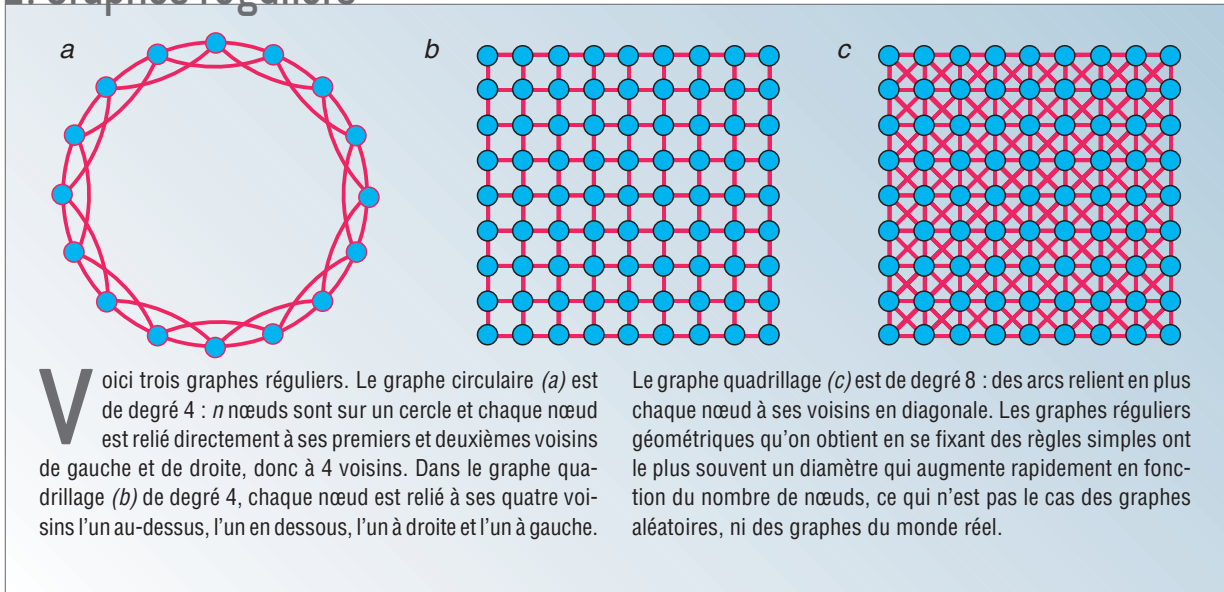
Sur ce graphe ayant $n = 9$ nœuds, et $m = 19$ arcs, on définit certaines caractéristiques. Le *diamètre* de ce graphe est 3, car deux nœuds peuvent toujours être joints l'un à l'autre par un chemin de trois arcs ou moins.

La *longueur caractéristique* L est la moyenne des longueurs (mesurées en nombre d'arcs) des plus courts chemins entre nœuds différents, ici $L = 1,555$.

Le *degré moyen* k des nœuds est le nombre moyen d'arcs attachés à un nœud, c'est-à-dire $k = 2m/n$. Ici k est égal à 4,22.

Le *coefficient d'agglomération* C mesure si des nœuds directement reliés à un même nœud sont souvent directement liés l'un à l'autre. On le calcule de la manière suivante. Pour chaque nœud N , on considère les nœuds qui lui sont directement connectés. Supposons qu'il y en ait c . Parmi les $c(c - 1)/2$ arcs possibles reliant ces c nœuds, certains existent dans le graphe, d'autres non. On calcule la proportion de ceux qui existent, soit $n(N)$. Le coefficient C est par définition la moyenne de tous les $n(N)$ lorsqu'on fait varier N . Ici $C = 62,2\%$.

2. Graphes réguliers



Voici trois graphes réguliers. Le graphe circulaire (a) est de degré 4 : n nœuds sont sur un cercle et chaque nœud est relié directement à ses premiers et deuxièmes voisins de gauche et de droite, donc à 4 voisins. Dans le graphe quadrillage (b) de degré 4, chaque nœud est relié à ses quatre voisins l'un au-dessus, l'un en dessous, l'un à droite et l'un à gauche.

Le graphe quadrillage (c) est de degré 8 : des arcs relient en plus chaque nœud à ses voisins en diagonale. Les graphes réguliers géométriques qu'on obtient en se fixant des règles simples ont le plus souvent un diamètre qui augmente rapidement en fonction du nombre de nœuds, ce qui n'est pas le cas des graphes aléatoires, ni des graphes du monde réel.

Les graphes réguliers sont ceux que l'on obtient en précisant des règles simples et systématiques pour disposer les nœuds et les arcs. Par exemple, le graphe régulier circulaire de degré 4 est obtenu en plaçant n nœuds N_1, N_2, \dots, N_n sur un cercle et en les reliant par des arcs de telle façon que chacun soit relié à ses deux voisins de droite et à ses deux voisins de gauche (le degré de chaque nœud est donc bien $k = 4$, et le nombre d'arc est $m = 2n$).

À l'opposé des graphes réguliers, les graphes aléatoires classiques sont obtenus par un processus de tirage au sort. On choisit d'abord n le nombre de nœuds et m le nombre d'arcs. Les nœuds sont nommés N_1, N_2, \dots, N_n et on dispose les m arcs au hasard : m fois de suite on tire au sort deux nœuds et on les relie (on ne retient pas un arc déjà obtenu ou un arc qui relierait un nœud à lui-même).

On sait que dès que le degré moyen k d'un graphe aléatoire classique est de l'ordre de $\log(n)$ alors le graphe a toutes les chances d'être connexe, c'est-à-dire d'un seul tenant : de n'importe quel nœud, on peut en suivant les arcs rejoindre n'importe quel autre nœud. Il s'agit d'un résultat classique de la théorie des graphes aléatoires créée par les mathématiciens A. Rényi et P. Erdős en 1959. Une multitude de théorèmes est disponible concernant ces graphes aléatoires classiques dont on pensait qu'ils pouvaient servir de modèle aux graphes du monde réel, jusqu'à l'article de D. Watts et S. Strogatz de 1998 attirant l'attention sur les « petits mondes ».

Les paramètres qui dénotent la distinction entre les graphes réguliers, les graphes aléatoires classiques et les graphes de type « petit monde » sont la longueur caractéristique d'un chemin L , le diamètre D et le coefficient d'agglomération C .

La longueur caractéristique L est la longueur moyenne des plus courts chemins entre nœuds différents. On calcule L de la manière suivante : pour tous les couples A et B, on relie A et B par le plus court chemin possible utilisant les arcs du graphe. Il y a $n(n-1)/2$ choix possibles pour A et B. On évalue la moyenne de toutes ces longueurs (mesurées en nombre d'arcs du chemin). Cette moyenne est L . Lorsque le graphe n'est pas connexe, L vaut $+\infty$ puisqu'on fait la moyenne entre des quantités dont une au

moins vaut $+\infty$. Le diamètre D est la plus grande longueur possible entre deux nœuds. Là encore si le graphe n'est pas connexe, D vaut $+\infty$. Remarquons que ce diamètre D n'est pas toujours facile à calculer : par exemple, on ne connaît pas le diamètre du graphe connexe constitué par toutes les configurations du cube de Rübick.

Les deux coefficients L et D permettent d'évaluer la facilité avec laquelle on passe d'un point à un autre dans le graphe et c'est bien sûr une caractéristique importante d'un graphe. Si le réseau est utilisé pour transporter de l'information, elle circulera bien dans un graphe ayant de petits coefficients L et D . Si le réseau représente des contacts susceptibles de transmettre des organismes infectieux, alors des valeurs faibles pour L et D signifient une forte contagiosité. Notons que Réka Albert, Hawoong Jeong et Albert Barabasi ont évalué le diamètre du graphe des liens entre pages web. Leur étude conclut que celui-ci vaut environ 19 : quelle que soit la page sur laquelle vous êtes, vous pouvez passer en 19 clics de souris à n'importe quelle autre page du web (trouver ce chemin de 19 clics, n'est pas nécessairement facile).

Pour le graphe régulier circulaire de degré 4 à n nœuds, le diamètre est $n/4$. En effet, deux nœuds sont le plus éloignés l'un de l'autre s'ils sont diamétralement opposés sur le cercle servant de support ; or pour joindre deux tels nœuds par le plus court chemin, on saute de deux en deux (c'est possible puisque chaque nœud est lié à ses deux voisins de gauche et à ses deux voisins de droite) et donc on emprunte $n/4$ arcs environ. La longueur caractéristique de ce même graphe vaut environ $n/8$.

Considérons un autre exemple de graphe géométrique régulier, le quadrillage de degré 4 (voir la figure 2b). Ses nœuds sont tous les points de coordonnées entières du plan (x, y) pour x et y compris entre 1 et p où p est un entier fixé (il y a donc $n = p^2$ nœuds). Chaque nœud est relié à ses quatre voisins. Pour ce graphe, le diamètre est $2(p-1)$ et la longueur caractéristique est environ $p/2$. On le voit pour les graphes géométriques réguliers, le diamètre et la longueur caractéristique augmentent rapidement avec n : certains nœuds sont réellement éloignés les uns des autres.

En revanche, la longueur caractéristique et le diamètre d'un graphe aléatoire classique augmentent lentement en fonction de n . On démontre en effet que L et D sont de l'ordre de $\log(n)/\log(k)$. Cela signifie que, dans un graphe aléatoire, dès que le degré moyen est supérieur à 1, tout nœud peut être joint à tout autre nœud par un court chemin.

Les graphes du monde réel ont le plus souvent un petit diamètre et une petite longueur caractéristique (ce qui les rapproche des graphes aléatoires classiques) et c'est pourquoi vous êtes à quatre poignées de mains du Président de la République. Les paramètres longueur caractéristique L et diamètre D permettent donc de distinguer les graphes du monde réel des graphes réguliers, mais pas des graphes aléatoires.

Le coefficient d'agglomération

Le coefficient d'agglomération différencie les graphes aléatoires classiques des graphes « petit monde ». Ce coefficient C mesure si deux nœuds proches dans le graphe sont fréquemment reliés entre eux. On le calcule de la manière suivante. Pour chaque nœud N , on considère tous les nœuds qui lui sont directement connectés. Supposons qu'il y en ait c (c'est dans notre analogie première, le nombre des gens que vous connaissez). On examine maintenant les $c(c-1)/2$ arcs possibles reliant ces c nœuds ; certains existent dans le graphe, d'autres non. On calcule la proportion de ceux qui existent, soit $n(N)$. Le coefficient C est par définition la moyenne de tous les $n(N)$ lorsque l'on fait varier N . Ce coefficient est une grandeur sans dimension qui varie entre 0 (faible agglomération) à 100 % (forte agglomération). Lorsque ce coefficient n'est pas trop petit cela

signifie que le graphe est robuste : en supprimant quelques-uns des nœuds du graphe, on ne risque pas de faire varier beaucoup le diamètre et la longueur caractéristique, car il existe des « chemins de secours » de petite taille qui suppléent les arcs supprimés.

Pour le graphe régulier circulaire de degré 4, le coefficient vaut C et est égal à $1/6$, car chaque nœud est lié à quatre autres et que parmi les 6 arcs possibles entre ces quatre voisins, un seul est présent. Dans le cas du graphe régulier du quadrillage de degré 4, le coefficient C vaut 0, car parmi les quatre points qui sont reliés au point de coordonnées (x, y) aucun n'est rattaché à un autre. Dans le cas d'un quadrillage plus dense de degré 8 (on relie chaque point non plus seulement avec des voisins horizontalement et verticalement, mais aussi avec ses quatre voisins en diagonale, voir la figure 2c), le coefficient C vaut $12/28$: il y a 8 voisins pour un nœud donné, et parmi les 28 arcs possibles entre eux, 12 sont effectivement présents.

Dans le cas d'un graphe aléatoire classique, le coefficient d'agglomération vaut environ $C = 2m/n^2$, ce qui signifie que si m est de l'ordre de n (comme on le constate dans les graphes réels) alors le coefficient C est d'autant plus petit que le graphe est grand et est donc presque nul pour les graphes ayant un grand nombre de nœuds. On ne constate pas du tout cela pour les graphes du monde réel. Par exemple, C vaut 0,34 pour le graphe des coopérations entre mathématiciens, il vaut environ 0,28 pour les réseaux de neurones, il vaut 0,39 pour le réseau Internet.

En résumé, les réseaux trouvés dans le monde réel possèdent le plus souvent les caractéristiques suivantes qui définissent ce que nous appellerons les graphes « petit monde » : ils ont beaucoup de nœuds, assez peu d'arcs,

3. Paramètres de différents graphes de « petits mondes »

Acteurs de cinéma ayant tourné ensemble dans un film

$n = 449\,913$; $m = 25\,516\,482$; $k = 113,4$; $L = 3,48$; $C = 0,78$

Mathématiciens ayant cosigné un article

$n = 253\,339$; $m = 496\,489$; $k = 3,92$; $L = 7,57$; $C = 0,34$

Biologistes ayant cosigné un article

$n = 1\,520\,251$; $m = 11\,803\,064$; $k = 15,53$; $L = 4,92$; $C = 0,60$

Personnes ayant échangé des e-mail

$n = 59\,912$; $m = 86\,300$; $k = 1,44$; $L = 4,95$; $C = 0,16$

Graphe physique d'Internet

$n = 10\,697$; $m = 31\,992$; $k = 5,98$; $L = 3,31$; $C = 0,39$

Graphe de distribution électrique

$n = 4\,941$; $m = 6\,594$; $k = 2,67$; $L = 18,99$; $C = 0,08$

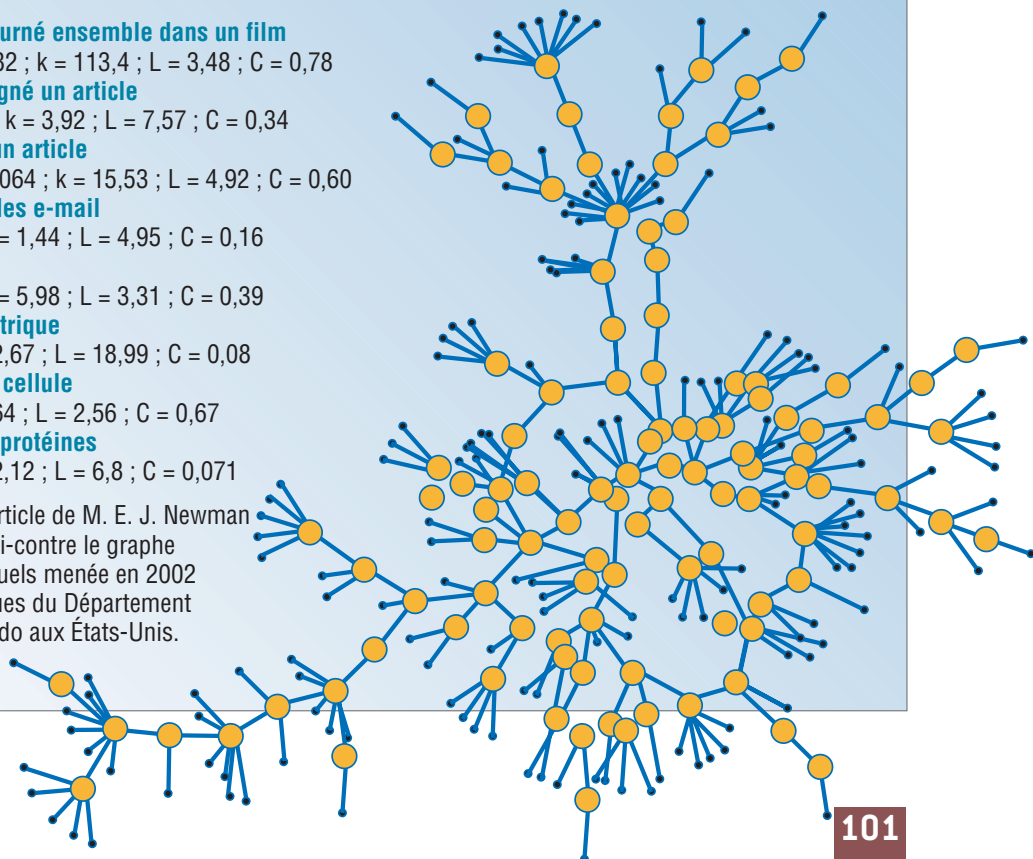
Réseau métabolique d'une cellule

$n = 765$; $m = 3\,686$; $k = 9,64$; $L = 2,56$; $C = 0,67$

Réseau d'interaction entre protéines

$n = 2\,115$; $m = 2\,240$; $k = 2,12$; $L = 6,8$; $C = 0,071$

Pour plus de détails, voir l'article de M. E. J. Newman indiqué à la fin de l'article. Ci-contre le graphe d'une étude de contacts sexuels menée en 2002 par J. Poterat et ses collègues du Département de santé publique du Colorado aux États-Unis.



possèdent un diamètre et une longueur caractéristique faibles (contrairement aux réseaux géométriques réguliers) et ont un coefficient d'agglomération bien supérieur à celui des réseaux aléatoires classiques.

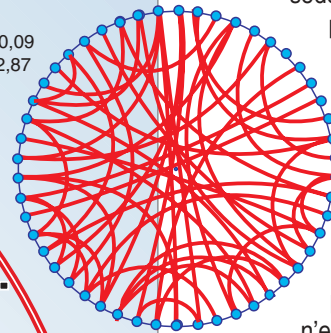
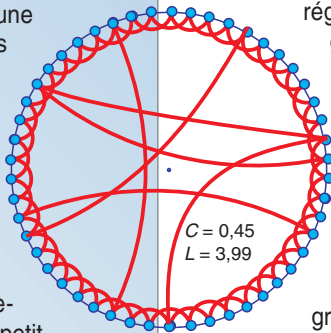
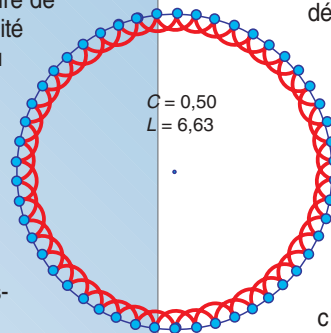
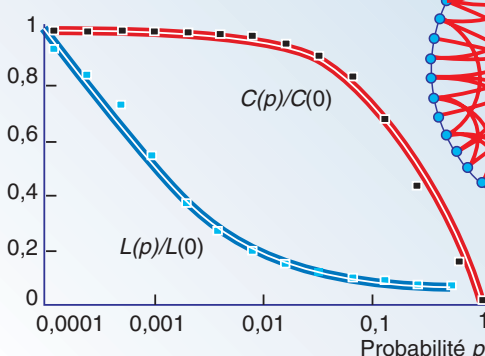
Les caractéristiques des graphes « petit monde » ne correspondent donc à aucun type de réseaux connus jusqu'alors et semblent occuper une place intermédiaire entre réseaux réguliers et aléatoires. Cette place intermédiaire des graphes « petit monde » entre ordre et désordre a été confirmée par une expérience remarquable réalisée par D. Watts et S. Strogatz. Cette simulation montre qu'un processus simple de « recâblage » fait passer d'un graphe régulier à un graphe aléatoire et que les graphes qu'on trouve dans une large zone d'évolution intermédiaire sont justement des graphes de type « petit monde ».

4. Le recâblage engendre des « petits mondes »

On part d'un graphe régulier circulaire de degré 4. On se donne une probabilité p (entre 0 à 1). Pour chaque arc du graphe, on tire au sort avec la probabilité p pour savoir si on le modifie. Lorsqu'on le modifie, on garde l'une de ses extrémités et on change l'autre qui est alors tirée au hasard. Pour chaque p , on obtient donc un graphe recâblé qui, si p est petit, ressemble beaucoup au graphe régulier initial, et qui si p est proche de 1 ressemble fortement à un graphe aléatoire.

On note que lorsque p possède une valeur intermédiaire, alors le graphe possède une faible longueur caractéristique (tous les nœuds peuvent être joints facilement), et un coefficient d'agglomération élevé.

Dériver quelques arcs crée des raccourcis (d'où la faible longueur caractéristique) sans pour autant nuire sensiblement à la bonne agglomération du graphe. Le processus de recâblage fait passer de l'ordre au désordre en passant par des graphes qui possèdent les caractéristiques des graphes « petit monde », c'est-à-dire les caractéristiques des graphes qu'on trouve dans le monde réel.



L'expérience proposée consiste à prendre le graphe circulaire de degré 4 et à lui faire subir le processus de recâblage partiel suivant. On se donne une probabilité p (qu'on fera varier de 0 à 1) et pour chaque arc du graphe on tire au sort avec la probabilité p pour savoir si on le modifie. Lorsqu'on doit le modifier, on garde l'une de ses extrémités et on change la seconde qui est alors tirée au hasard.

De régulier à aléatoire

Lorsque p est égal à 0, le graphe circulaire n'est pas modifié. Lorsque p est égal à 1, le graphe obtenu est très semblable à un graphe aléatoire. Plus intéressant, lorsque p possède une valeur intermédiaire, les paramètres du graphe recâblé sont ceux d'un graphe « petit monde » : faible diamètre, faible longueur caractéristique, et coefficient d'agglomération nettement positif. Le seul fait de dériver quelques arcs crée des raccourcis dans le graphe (d'où la faible longueur caractéristique) sans pour autant nuire à son agglomération. Les courbes de la figure 4 montrent la

décroissance lente du coefficient C lorsque p varie de 0 à 1, alors que la longueur caractéristique diminue vite.

Bien sûr, des variantes de cette simulation ont été réalisées en prenant comme point de départ des graphes réguliers à deux dimensions (graphe du quadrillage de degré 8 par exemple) ou plus. À chaque fois, on voit apparaître des graphes « petit monde ». Des résultats mathématiques précisent les courbes d'évolution des coefficients L et C et confirment qu'entre graphes réguliers et graphes aléatoires classiques surgissent spontanément les graphes « petit monde » qu'il n'y a donc rien d'étonnant à rencontrer dans la réalité.

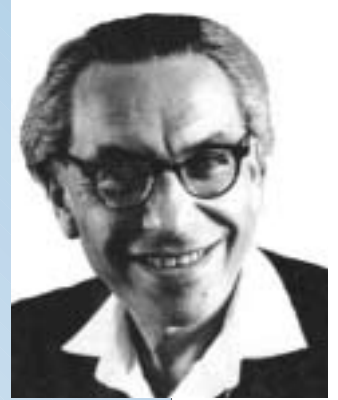
La probabilité qu'un nœud soit de degré k (c'est-à-dire possède k arcs) dans les graphes du monde réel a été étudiée avec soin et a conduit à affiner le modèle des graphes « petit monde » de D. Watts et S. Strogatz. Bien sûr, plus le nombre k est grand, moins il est probable qu'un nœud possède k arcs. Cependant la façon dont décroît la probabilité p_k en fonction de k présente des régularités remarquables. La plupart du temps p_k possède une expression approchée du type $p_k = bk^{-a}$ où b est positif et a compris entre 1 et 5. Pour le graphe des acteurs ayant tourné des films en commun a vaut 2,3 de même que pour le graphe des collaborations entre mathématiciens. Une telle décroissance de type polynomial pour p_k s'appelle une loi de puissance et n'est pas présente dans le cas des graphes aléatoires classiques qui eux sont tels que p_k décroît selon une loi exponentielle $p_k = ba^{-k}$. Dans le cas des graphes réguliers, il n'y a pas non plus de loi de puissance puisque p_k est nul à l'exception d'une valeur de k .

5. Le nombre de Erdős

Paul Erdős (1913-1996) est le plus prolifique mathématicien du xx^e siècle. Il a écrit 1 475 articles et a cosigné des articles avec presque 500 personnes différentes. Un petit jeu consiste à mesurer à quelle distance se trouvent de lui d'autres auteurs en calculant leur « nombre de Erdős ». Par définition, le nombre de Erdős de Erdős est 0, et toutes les personnes ayant cosigné un article avec lui ont un nombre de Erdős valant 1. Les gens ayant cosigné un article avec quelqu'un ayant un nombre de Erdős de 1 ont un nombre de Erdős de 2, etc. Bien sûr, quelqu'un qui n'a jamais publié d'articles ou qui n'en a publié qu'en les signant tout seul a un nombre de Erdős de $+\infty$. Cependant un très grand nombre de scientifiques ont un nombre de Erdős fini et assez petit. En explorant les banques de données bibliographiques, Rodrigo de Castro et Gerold Grossman ont pu établir les choses suivantes qui suggèrent que le graphe des collaborations entre mathématiciens est un « petit monde ». 485 personnes ont un nombre de Erdős de 1, c'est-à-dire ont écrit un article ou un livre en collaboration avec lui ; 5 000 personnes au moins ont un nombre de Erdős valant 2 ; plus de 60 prix Nobel ont un nombre de Erdős inférieur ou égal à 9 ; tous les médaillés

Field (le plus prestigieux des prix en mathématiques) jusqu'en 1998 ont un nombre de Erdős inférieur ou égal à 6 ; Albert Einstein possède le nombre de Erdős 2 (il a cosigné deux articles avec son assistant Ernst G. Straus qui lui en a cosigné une vingtaine avec Erdős). Einstein ayant écrit des articles avec les prix Nobel Otto Stern et Wolfgang Pauli, ceux-ci ont 3 pour nombre de Erdős. Comme Einstein publia un fascicule avec Sigmund Freud en 1933 intitulé *Why War ?* Freud a donc un nombre de Erdős de 3 (voir <http://www.oakland.edu/~grossman/erdoshp.html>).

Voici le nombre de Erdős de quelques personnalités scientifiques : Max Born 3 ; Jean Bourgain 2 ; Louis de Broglie 5 ; Gregory Chaitin 3 ; Alain Connes 3 ; Gérard Debreu 3 ; Pierre Deligne 3 ; Kurt Gödel 3 ; Mikhael Gromov 3 ; Alexandre Grothendieck 5 ; Serge Lang 3 ; John Nash 4 ; Roger Penrose 3 ; Jean Piaget 3 ; Laurent Schwartz 4 ; Jean-Pierre Serre 3 ; Claude Shannon 3 ; René Thom 4 ; John Von Neumann 3 ; André Weil 4 ; Andrew Wiles 3.



Lorsque les degrés des nœuds d'un graphe suivent une loi de puissance, le nombre de nœuds ayant un grand nombre de liens avec d'autres diminue assez lentement, ce qui fait qu'on trouve réellement des nœuds ayant un degré important. En revanche, dans le cas d'une loi exponentielle, les probabilités p_k deviennent vite infinitésimales, ce qui se traduit concrètement par le fait qu'on ne trouve aucun nœud fortement connecté. Les nœuds très fortement connectés des graphes obéissant à une loi de puissance sont des sortes de points centraux du graphe et leur présence joue un rôle important. Dans Internet, ils permettent au « surfeur » de se repérer et de mener sa recherche plus facilement.

Les graphes « petit monde » engendrés par les simulations de D. Watts et S. Strogatz ne présentent pas une loi de puissance pour les degrés, ce qui montre qu'ils constituent une trop grossière approximation des graphes réels. D'autres modèles ont donc été recherchés. Le plus simple d'entre eux est séduisant, car il semble se fonder sur une modélisation de la façon même dont Internet s'accroît jour après jour. L'idée de ce modèle a été proposée par Derek de Solla Price en 1965 dans un travail presque oublié sur les citations entre articles scientifiques. L'idée est d'ajouter des nœuds progressivement et de les rattacher au hasard aux nœuds déjà présents en considérant qu'un nœud ayant déjà beaucoup de liens sera plus facilement choisi qu'un nœud en ayant peu (probabilité de rattachement proportionnelle aux degrés des nœuds).

Ce principe est la transcription pour les graphes du principe de l'enrichissement : « l'argent va à l'argent ». Dans le cas d'Internet, un tel principe est naturel : plus un lien est déjà référencé, plus il a de chances lors de la création d'une nouvelle page d'être référencé à nouveau. Une telle règle appliquée systématiquement pour engendrer de gros

graphes produit des graphes dont les paramètres p_k suivent une loi de puissance.

Les sciences de la complexité

Ce modèle de « petit monde » avec loi de puissance semble rendre compte des propriétés des graphes trouvés dans la réalité avec plus d'exactitude que les graphes de D. Watts et S. Strogatz qui étaient déjà bien meilleurs que les graphes aléatoires ou réguliers.

Des variantes du modèle de D. Price et d'autres modèles complémentaires ont été récemment proposées pour encore affiner la modélisation des réseaux réels. C'est finalement tout un domaine de recherche qui est né où l'objectif est la compréhension des caractéristiques des gros graphes innombrables du monde réel. Entre physique, informatique et mathématiques, une discipline nouvelle a vu le jour qui enrichit ce qu'on dénomme parfois la famille des sciences de la complexité.

Jean-Paul DELAHAYE est professeur d'informatique à l'univ. de Lille.

WATTS, D. J., *Six Degrees : The Science of a Connected Age*, Norton, New York, 2003.

NEWMAN, M. E. J., *The Structure and Function of Complex Networks*, 2003, <http://www.santafe.edu/~ffmark/pubs.html>

BARABASI, A., *Linked, the science of networks*, Perseus Publishing Cambridge, Massachusetts, 2002.

STROGATZ, S. H., *Exploring Complex Networks*, in *Nature*, 410, 268–276, 2001.

HAYES, B., *Graph Theory in Practice* : Part I, in *American Scientist* 88 (1), 9–13, 2000. Part II, 88 (2), 104–109, 2000.

HOFFMAN, Paul, *Erdős. L'homme qui n'aimait que les nombres*, collection « Un savant, une époque », éditions Belin, 2000.