



HAL
open science

L'analyse de réseaux en sciences sociales et en histoire

Laurent Beauguitte

► **To cite this version:**

Laurent Beauguitte. L'analyse de réseaux en sciences sociales et en histoire : Vocabulaire, principes et limites. Le réseau. Usages d'une notion polysémique en sciences humaines et sociales, Presses Universitaires de Louvain, pp.9-24, 2016, 978-2-87558-500-4. halshs-01476090v2

HAL Id: halshs-01476090

<https://shs.hal.science/halshs-01476090v2>

Submitted on 25 Mar 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

L'analyse de réseaux en sciences sociales et en histoire : vocabulaire, principes et limites

Laurent Beauguitte, CNRS, UMR Géographie-cités
laurent<dot>beauguitte<at>cnrs.fr

2016

Chapitre d'introduction de l'ouvrage dirigé par Rosemonde Letricot, Mario Cuxac, Maria Utcategui et Andrea Cavaletto, 2016, *Le réseau. Usages d'une notion polysémique en sciences humaines et sociales*, UCL Presses Universitaires de Louvain.

Les nombres en gras entre crochets correspondent à la pagination de la version éditeur.

[9]

Introduction

L'analyse de réseaux a connu ces dernières années un essor marqué dans la plupart des sciences sociales, notamment en histoire. Dans la sphère francophone, la multiplication récente des rencontres liées à ces problématiques est un indicateur fort de cette tendance¹. Le terme de réseau étant d'une polysémie et d'une plasticité fortes², il est sans doute nécessaire de commencer par expliciter les termes et formules utilisés.

Une distinction essentielle concerne la différence entre analyse de réseau et analyse d'un réseau³. L'analyse de réseau désigne un ensemble de méthodes, de notions et de concepts fondés sur la théorie des graphes pour étudier un phénomène relationnel donné, et des méthodes d'analyse de réseau ont été développées dans différentes disciplines au moins depuis les années 1930. L'analyse d'un réseau désigne l'étude d'un réseau donné, matériel ou immatériel. Elle peut être réalisée à l'aide d'outils très divers, qualitatifs (ex. étude sur l'implantation d'un tronçon autoroutier faite à l'aide d'entretiens

1. On citera notamment, sans prétention à l'exhaustivité, les rencontres *Res-Hist* de Nice (2013), Toulouse (2014) et Paris (2015), l'école thématique *Réseaux* de Ferney-Voltaire (2014) ou encore le colloque *Réseaux et histoire* de Reims (2015).

2. P. MUSSO, *Critique des réseaux*, Paris, PUF, 2003.

3. G. GARROTE, « Analyse de réseau ou analyse des réseaux? », 2013, <http://reshist.hypotheses.org/287> (visité le 27 mars 2015).

avec des acteurs) ou quantitatifs (étude des capacités de charge de tel ou tel réseau). Analyser un réseau ne suppose pas nécessairement de faire appel à des méthodes d'analyse de réseau et, inversement, on peut mobiliser les méthodes d'analyse de réseau pour étudier des œuvres littéraires, des systèmes écologiques, etc. Le terme de réseau utilisé seul est une notion floue, polysémique et d'une utilité parfois relative : pourtant chaque discipline tend à privilégier l'étude d'un type particulier de réseaux (ex. réseaux personnels en sociologie, réseaux techniques en géographie humaine, réseaux hydrologiques en géographie physique) et à développer des outils, mesures et concepts dédiés à l'analyse de ce type de réseaux.

Après avoir rappelé les définitions d'un certain nombre de termes clés en analyse de réseau, nous évoquerons les différentes approches méthodologiques existant en sociologie et en géographie. Il s'agira moins de dresser un tableau exhaustif que de présenter de manière synthétique et nécessairement réductrice les convergences et [10] les divergences d'approche, tant en matière conceptuelle que méthodologique. Une troisième partie s'intéresse au cas particulier de l'analyse de réseaux en histoire, tant certaines questions apparaissent spécifiques à cette discipline (problème des sources, anachronisme du vocabulaire, problèmes liés à la représentation des données, etc.).

1 Précisions terminologiques

L'analyse de réseau existant dans différentes disciplines depuis une cinquantaine d'années, un vocabulaire spécifique existe et il est sans doute préférable de se l'approprier avant de souhaiter mener à bien des analyses de réseau en histoire. Il faut cependant noter que le même terme peut avoir un sens différent dans deux disciplines et, qu'inversement, une même notion peut être nommée différemment en fonction des champs académiques⁴. Il est devenu impossible d'être exhaustif dans la définition des termes, des mesures et des notions utilisés dans l'ensemble des disciplines mobilisant l'analyse de réseaux⁵. Cette section s'intéresse à trois aspects complémentaires : la qualification du réseau observé ; l'étude des liaisons à l'intérieur de ce réseau ; et enfin deux types particuliers de réseaux fréquemment étudiés.

1.1 Définir les réseaux : quelques oppositions de base

Un réseau est constitué d'un ensemble fini et non vide de points, symbolisant des acteurs (individus, groupes, institutions, textes, etc.) et d'un ensemble fini et éventuellement vide de lignes symbolisant les relations entre

4. La densité d'un réseau (nombre de liens présents divisé par le nombre de liens possibles) sera ainsi nommée densité en sociologie, indice gamma en géographie et *connectance* en écologie.

5. Pour une présentation quasi exhaustive, voir M.E.J. NEWMAN, *Networks : An introduction*, Oxford, Oxford University Press, 2010.

ces acteurs. Le nombre de sommets est appelé ordre du graphe ; le nombre du liens taille du graphe. Dans la pratique, on parle cependant de grands graphes quand le nombre de sommets et/ou de liens est élevé.

Les termes de graphe et de réseau sont la plupart du temps employés comme des synonymes même si, pour certains, le terme graphe désigne l'objet mathématique constitué de points et de lignes, l'ajout d'attributs sur les liens et/ou les sommets permettant la transformation en réseau. Dans la pratique, on étudie toujours des réseaux (sauf en théorie de graphes *stricto sensu*), et une série d'oppositions permet de définir les différents types de réseaux ⁶.

Une première opposition distingue les réseaux non orientés, dans lesquels toute relation d'un sommet i vers un sommet j implique une relation de j vers i et les réseaux orientés où le sens de la relation importe et où l'existence d'une relation de i vers j n'implique pas nécessairement de relation de j vers i . Certains réseaux dits mixtes [11] incluent ces deux types de liens : l'exemple le plus courant concerne les arbres généalogiques qui associent des liens non orientés (être le conjoint/la conjointe de) et des liens orientés (être l'ascendant.e de).

Les graphes dits simples comportent un seul type de relations entre les sommets (à la différence d'un multigraphe) et il n'existe pas pas de boucle (lien d'un sommet vers lui-même et, dans ce cas, on parlera de pseudo graphe). Les boucles ne sont quasiment jamais considérées en analyse de réseau, or elles peuvent avoir un intérêt thématique évident (cas de l'auto-citation en sciences ou de l'étude des flux et des migrations en géographie). Le graphe complexe - terme récent, apparu à la fin des années 1990 - est très souvent un graphe simple mais comprenant plusieurs (dizaines de) milliers de sommets et/ou de liens, ce qui nécessite des approches méthodologiques différentes.

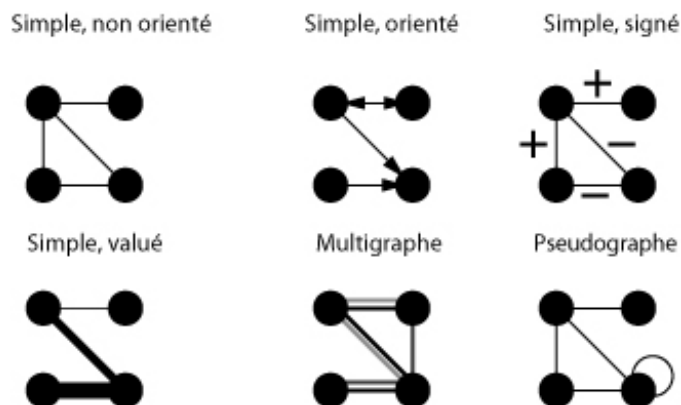
Enfin, la caractérisation de liens permet de différencier les graphes booléens (une relation a deux états possibles, présence et absence), les graphes valués (les liens sont porteurs d'une intensité) et les graphes signés (les liens présents peuvent traduire un lien positif ou un lien négatif). Dans la sociométrie, qu'il est possible de considérer comme l'ancêtre de l'analyse de réseaux sociaux ⁷, les liens négatifs et positifs étaient étudiés simultanément, ce qui n'est plus le cas aujourd'hui. La figure 1 illustre ces différences de types de réseaux, différences fondamentales dans la mesure où elles déterminent en partie le type d'analyse qu'il sera possible de réaliser ensuite.

Une autre typologie possible s'intéresse à des formes remarquables de réseaux, formes qui servent de point de repère conceptuel et/ou méthodo-

6. Voir également L. BEAUGUITTE, « Graphes, réseaux, réseaux sociaux : vocabulaire et notation », Working paper du groupe fmr , 2010, <https://halshs.archives-ouvertes.fr/FMR> (visité le 27 mars 2015).

7. L. FREEMAN, *The development of social network analysis*, Charleston, Booksurge Publishing, 2004.

FIGURE 1 – Principaux types de réseaux



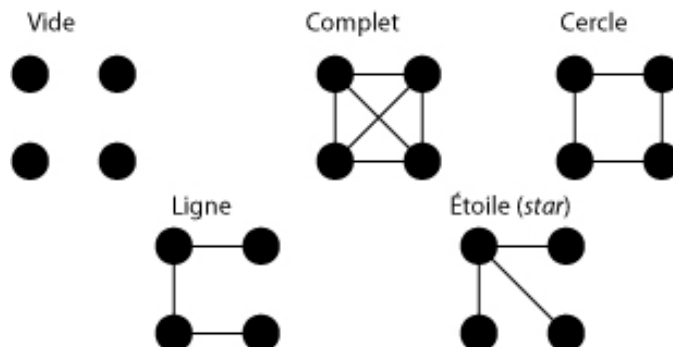
logique (voir [12] figure 2). On pourra ainsi distinguer les graphes vides ne comprenant aucun lien et les graphes complets où tous les liens possibles sont présents. Cette opposition met en regard les deux extrêmes entre lesquels nous qualifions une propriété d'ensemble de nos graphes empiriques *via* le calcul de la densité. D'autres figures idéal-typiques sont d'un intérêt thématique spécifique : les cercles, les lignes et les étoiles. Dans un cercle, chaque sommet occupe vis-à-vis des autres une position équivalente, tous ont le même nombre de liens adjacents (on parle de graphe k -régulier, tous les sommets ayant le même degré k) et il s'agit donc d'une situation parfaitement égalitaire si l'on se situe dans des problématiques d'accès à l'information dans un réseau personnel par exemple. Inversement, le graphe étoilé représente une hiérarchie extrême entre un acteur central, point de passage obligé pour toute circulation (de biens, d'information, de ressources au sens large), et des acteurs périphériques et dépendants du bon vouloir de l'acteur central. La ligne forme une situation intermédiaire dans la mesure où plus l'on est proche des extrémités, plus l'accès à l'information devient difficile⁸.

1.2 Chemins, chaînes et connexité

Une suite de liens permettant de rejoindre deux sommets est appelée chemin (prise en compte de l'orientation des liens) ou chaîne (non prise en compte de l'orientation des liens). On s'intéresse notamment aux plus courts chemins entre sommets ou distance géodésique, mesurés en nombre de liens (moyen le plus rapide de faire circuler un bien, une information, etc.), ainsi qu'à l'existence de différents chemins entre deux [13] sommets (robustesse

8. S. WASSERMAN et K. FAUST, *Social Network Analysis. Methods and Applications*, Cambridge Mass., Cambridge University Press, 1994.

FIGURE 2 – Quelques idéaux-types de réseaux



du réseau). Si le sommet de départ et celui d'arrivée sont les mêmes, on parle de circuit (graphe orienté) ou de cycle (graphe non orienté).

Un graphe est dit connexe lorsqu'un chemin existe entre toute paire de sommets. Si des sommets ne sont adjacents à aucun lien, ils sont nommés isolés et le graphe n'est donc pas connexe. Les sous-graphes non connexes d'un graphe sont appelés des composantes connexes. Il peut exister dans un graphe des points ou des liens possédant une propriété liée à la connexité : leur suppression entraîne une augmentation du nombre de composantes connexes. On parle pour les sommets de points d'articulation et, pour les liens, d'isthmes - la présence d'un isthme s'accompagne nécessairement de la présence de deux points d'articulation, on peut par contre avoir des points d'articulation sans présence d'isthme. Ces points et liens peuvent selon les cas révéler une vulnérabilité ou une position stratégique dans un réseau. Un arbre est un graphe connexe acyclique et formé uniquement d'isthmes (si l'arbre possède V sommets, il comprend $V - 1$ liens).

1.3 Ego-networks et réseaux bipartis

Un type particulier de réseau est le réseau personnel ou *ego-network* : il est constitué par un sommet central (*ego*), l'ensemble de ses relations (*alters*) et les relations entre ces *alters*. L'*ego-network* peut être extrait d'un réseau plus large (ex. réseau d'un élève extrait du réseau d'une classe) ou être le résultat des méthodes utilisées pour le recueil de données⁹.

Les réseaux bipartis ou *2-mode* forment un autre type particulier de réseaux : ils représentent les liens entre deux ensembles différents de sommets. Il peut s'agir, pour reprendre des exemples fréquemment étudiés dans la littérature, des liens entre un ensemble d'acteurs et un ensemble de films (un lien signale que tel acteur joue dans tel film), de liens entre des scientifiques

9. J. SCOTT, *Social network analysis. A handbook*, New York, Sage, 2000 (2^e éd.).

et des revues (un lien signale que tel chercheur a publié dans tel revue). Un champ d'étude spécifique, l'étude des *interlocks*, désigne les analyses de réseaux bipartis entre actionnaires et conseils d'administration. Le terme de réseau d'affiliation est également employé. La matrice d'adjacence¹⁰ correspondante est de type $V1 \times V2$ où $V1$ est la population de l'ensemble de sommets 1, $V2$ la population de l'ensemble de sommets 2, et les intersections signalent la présence (1) ou l'absence de relations (0) entre ces deux populations.

L'utilisation de ces réseaux est très fréquente sur internet (voir par exemple sur Facebook, « Vos amis aiment ça » ou sur Amazon, « Les livres recommandés pour vous »). Les réseaux bipartis sont souvent transformés en réseau *one-mode* valué en multipliant la matrice d'adjacence par sa transposée ou l'inverse. [14]

2 Des approches disciplinaires plurielles

Si la théorie des graphes existe comme branche des mathématiques depuis le fameux problème des ponts de Königsberg résolu par Euler en 1759, il faut attendre les années 30 pour déceler son utilisation dans les sciences sociales avec la sociométrie de J. Moreno (1934). Le développement de l'analyse de réseaux en sociologie est plus tardive et suit deux orientations distinctes, tant au niveau géographique (Manchester *vs* Harvard) que méthodologique (approche plus qualitative en Angleterre, plus formaliste aux États-Unis). D'autres disciplines élaborent au même moment (années 1960) des paradigmes spécifiques d'analyse de réseaux et, pour des raisons de compétence de l'auteur et de place, seul le paradigme de l'analyse géographique des réseaux techniques sera évoqué ici. Si le paysage scientifique sera durablement marqué par l'irruption des physiciens dans l'analyse de réseau à la fin des années 90, cet aspect ne sera pas développé ici dans la mesure où son impact dans les différentes sciences sociales a été très divers et, à ce jour, peu décisif dans les études historiques. Il est nécessaire de rappeler, sans pour autant tomber dans le déterminisme technologique, que l'analyse de réseau repose sur du calcul matriciel et nécessite une puissance de calcul importante. Si de nombreuses pistes suggérées par Moreno ne sont pas poursuivies avant les années 60, c'est en grande partie en raison du manque de matériel informatique adapté. Et si les sociologues étudient dans les années 1970 de petits réseaux quand les physiciens analysent aujourd'hui des réseaux de plusieurs (centaines de) milliers de sommets et de liens, la puissance informatique actuellement disponible est sans aucun doute un facteur explicatif majeur.

10. La matrice d'adjacence donne exactement les mêmes informations que la représentation graphique : les sommets sont en ligne et en colonnes, les cases sont remplies par des 0 (absence de liens), des 1 (présence d'un lien) ou une intensité dans le cas de graphe valué. La forme matricielle autorise le calcul d'indicateurs sur le réseau étudié.

2.1 La sociométrie, ancêtre de l'analyse des réseaux sociaux ?

Le rôle de Moreno dans l'histoire de l'analyse de réseaux sociaux est jugé très différemment selon les auteurs : Linton Freeman le considère comme un précurseur majeur¹¹ quand Scott (*op. cit.*) l'évoque à peine et Rainie et Wellman ne le citent pas du tout¹². Les avis sont tout aussi divergents en France et on peut s'amuser à comparer le sort respectif fait à ses travaux dans les manuels francophones de référence¹³. Or, la lecture de l'ouvrage co-écrit avec Helen Jennings en 1934¹⁴ montre que la plupart des pistes [15] suivies à partir des années 1960, notamment par l'école de Harvard, y trouvent leur origine. Comment expliquer cette vision très contrastée ? Il semble que des éléments tant biographiques que structurels fournissent des éléments de réponse pertinents.

Jacob L. Moreno (1889-1974), né à Bucarest, arrive à Vienne avec ses parents à 6 ans et fait des études de médecine (psychiatrie), de mathématiques et de philosophie. Il s'installe comme médecin en Autriche mais sa carrière stagne et il migre alors pour New York en 1925. Il rencontre en 1927 une étudiante en psychologie, Helen Hall Jennings qui a une formation en recueil des données et en statistique. Ils enquêtent notamment à SingSing et à l'*Hudson School for Girls*. Moreno met au point ce qu'il appelle la géographie psychologique puis, à partir de 1934, la sociométrie ainsi définie : *an « experimental technique... obtained by application of quantitative methods... which inquire into the evolution and organization of groups and the position of individuals within them »*. Dans *Who shall survive ?*, l'utilisation massive des sociogrammes - dessins des réseaux relationnels des personnes enquêtées¹⁵ - est une innovation forte. Et contrairement à de mauvaises habitudes qui se répandront plus tard, tous ses graphes sont accompagnés d'une légende complète et explicite. Si l'ouvrage propose de très nombreuses pistes, les moyens mathématiques de l'époque ne permettent pas de calculer tous

11. « There is practically no idea or practice in contemporary SNA that is not present in some form in that book. », FREEMAN, *op. cit.*, p. 160.

12. L. RAINIE et B. WELLMAN, *Networked. The new social operating system*, Cambridge Mass, MIT Press, 2012.

13. A. DEGENNE et M. FORSÉ, *Les réseaux sociaux*, Paris, Armand Colin, 2004 (2e édition) ; E. LAZEGA, *Réseaux sociaux et structures relationnelles*, Paris, PUF, 2007 ; P. MERCKLÉ, *Sociologie des réseaux sociaux*, Paris, La Découverte, 2004.

14. J.L. MORENO, *Who shall survive ? A New Approach to the Problem of Human Interrelations*, Washington, Nervous and Mental Disease Publishing Co. Si tous les auteurs s'accordent à louer le rôle essentiel joué par Helen Jennings dans la genèse de l'ouvrage, elle n'est pas créditée comme auteur. La comparaison avec la ré-édition de 1950, supervisée par le seul Moreno et encombrée de longues considérations théologico-personnelles parfois embarrassantes, permet pourtant de mettre en évidence le rôle clé joué par cette étudiante en psychologie.

15. L. BEAUGUITTE, 2013, « (re)lire les classiques : Who shall survive? de J.L. Moreno, 1934 », Groupe fmr, <http://groupefmr.hypotheses.org/1069>, visité le 27 mars 2015.

les indicateurs proposés. Moreno crée le journal *Sociometric review* en 1936 renommé *Sociometry* en 1937 et l'élite de la sociologie américaine (Lazarfeld, Becker, Boas, Newcomb, Stouffer, etc.) y publie.

Puis, en quelques années, la sociométrie disparaît... Selon ses biographes, Moreno a un caractère personnel... particulier : mégalomane, paranoïaque (tout le monde vole ses idées), il entend des voix, se prend pour Dieu à l'occasion, et s'attribue le mérite de toutes les découvertes importantes en science sociale. Il fait rapidement fuir les chercheurs académiques et discrédite la sociométrie comme champ de recherche. De plus, il oriente de plus en plus la sociométrie vers la thérapie (jeux de rôle ayant pour but soigner les problèmes psychiques). Et, si la revue continue, il n'y a pas d'institutionnalisation de la sociométrie : Moreno est en dehors du champ académique, ne donne pas de cours, ne dirige pas de thèses et ne peut donc pas assurer la reproduction universitaire de son approche. [16]

2.2 École de Manchester et école de Harvard : deux traditions d'analyse des réseaux sociaux

16

Dans les années 1950-60, deux foyers de développement d'analyse de réseaux se développent : l'un à Harvard autour d'Harrison White, l'autre à Manchester autour du Département d'anthropologie.

Le premier foyer a été et reste sans nul doute le plus influent aujourd'hui. Autant que ses talents de chercheur (auteur d'une thèse en physique théorique au MIT en 1955, donc capable d'un haut niveau de formalisation mathématique, puis en sociologie à Princeton en 1960), ce sont sans doute ses talents d'enseignant qui ont permis à White de créer une école pérenne. La liste des thèses dirigées dans les années 1970 constitue un *Who's who* de l'analyse de réseau anglophone avec des auteurs devenus aussi fameux que Bonacich, Breiger, Granovetter, Fisher, Levine ou Wellman. Les talents d'organisateur de White et de ses élèves (création de l'INSNA¹⁷ et du bulletin *Connections* en 1977), leur capacité à diffuser leurs résultats via les *SunBelt* (colloque nord-américain annuel à partir de 1981) et la revue *Social Networks* (créée en 1978), la maîtrise du développement informatique permettant de normaliser tant les mesures utilisées que les formats de données (logiciel *Structure* créé par R. Burt en 1975, logiciel *Ucinet* développé à partir de 1980 par L. Freeman) ont assuré la pérennité du paradigme devenu classique de la SNA (*Social network analysis*). Les mesures et techniques

16. Le titre de ce paragraphe reprend celui de l'article de Michael Eve : M. EVE, « Deux traditions d'analyse des réseaux sociaux », *Réseaux*, 2002, 5, p. 183-212.

17. *International Network for Social Network Analysis* : l'association est toujours active aujourd'hui et continue à organiser les *Sunbelt* nord-américaines tous les ans (<http://www.insna.org/>, visité le 27 mars 2015).

proposées restent largement utilisées aujourd’hui, y compris dans d’autres sciences sociales dont l’histoire ou l’archéologie.

La tradition de l’école de Manchester, qui regroupe des auteurs comme John Barnes, Clyde Mitchell ou Elizabeth Bott, n’a pas su diffuser ses méthodes ni surtout les pérenniser avec autant de succès. S’appuyant sur des démarches moins formalisées au niveau mathématique, basé sur des enquêtes ethnographiques de longue durée, ce paradigme a en partie échoué à proposer un ensemble de méthodes reproductibles et mobilisables sur des terrains autres que ceux initialement étudiés (essentiellement des milieux urbains de pays dits alors en voie de développement). Si l’on doit à J. Barnes la première définition conceptualisée du réseau social¹⁸ (« *the whole of social life’ could be seen as ‘a set of points some of which are joined by lines’ to form a ‘total network’* »), la lecture de l’ouvrage collectif coordonné par Mitchell¹⁹ permet de comprendre pourquoi le paradigme échoue à s’imposer : d’une contribution à l’autre, et [17] malgré des pistes tout à fait prometteuses (méthodes de normalisation, étude de la spatialité des réseaux sociaux notamment), les définitions et les termes employés varient considérablement d’un chapitre à l’autre. Quand au même moment outre-Atlantique, un vocabulaire normalisé se met en place, avec notamment l’apparition des premiers manuels²⁰, le flou terminologique ne pouvait que compromettre la pérennité de l’entreprise²¹.

2.3 L’analyse des réseaux techniques en géographie

Si la sociologie étudie de manière préférentielle voire exclusive des réseaux non planaires, l’étude de réseaux techniques, territoriaux et planaires a très longtemps caractérisé les approches géographiques des réseaux. Si certains géographes s’intéressent depuis le milieu des années 1990 à d’autres types de réseaux et ont, de fait, renouvelé leurs approches méthodologiques, important notamment des mesures proposées par les physiciens²², la grande majorité des travaux géographiques contemporains reste marquée par le caractère planaire des réseaux étudiés. Dans l’un des ouvrages de référence sur l’étude géographique des réseaux²³, ces derniers sont ainsi définis : « toute

18. J. BARNES, « Classes sociales et réseaux dans une île de Norvège », *Réseaux*, 6/2013 (182), p. 209-237. Traduction de l’article « Class and Committees in a Norwegian Island Parish » paru en 1954 dans la revue *Human Relations* (7, p. 39-59).

19. C. MITCHELL (éd.), *Social networks in urban situations : Analyses of personal relationships in Central African towns*, Manchester, Manchester University Press, 1969.

20. F. HARARY, D. CARTWRIGHT et R. NORMAN, *Structural models : An introduction to the theory of directed graphs*, New York, John Wiley & Sons, 1965.

21. Pour une relecture contemporaine et anthropologique des apports de l’école de Manchester, voir T. EVENS et D. HANDELMAN (éd.), *The Manchester School : practice and ethnographic praxis in anthropology*, New York, Berghahn Books, 2006.

22. C. DUCRUET et L. BEAUGUITTE, « Spatial science and network science : Review and outcomes of a complex relationship », *Networks and Spatial Economics*, 14(3-4), 2014, p. 297-316.

23. G. DUPUY et Groupe Réseaux, *Réseaux territoriaux*, Caen, Paradigme, 1988.

infrastructure permettant le transport de matière, d'énergie ou d'information, qui s'inscrit sur un territoire où elle est caractérisée par la topologie de ses points d'accès ou points terminaux, de ses arcs de transmission, de ses nœuds de bifurcation ou de communication. »

Les réseaux techniques présentent au moins deux caractéristiques distinctes par rapport aux réseaux sociaux : ils ont nécessairement une inscription spatiale (coordonnées précises pour les sommets, tracé donné pour les liens) et ils servent à transporter une quantité donnée (de personnes, de biens, d'énergie, etc). On aura donc affaire à des réseaux le plus souvent valués et spatialement contraints. Leur étude et leur représentation différeront donc - tout comme les logiciels utilisés (SIG - Systèmes d'information géographique). Une autre différence majeure concerne les questions de recherche : si la recherche d'acteurs centraux devient la recherche des lieux les plus accessibles, les questions de performance du réseau, de vulnérabilité et surtout des modifications à entreprendre pour améliorer ses capacités font de l'analyse des réseaux techniques une discipline aux aspects applicatifs essentiels. Contrairement à la sociologie, il est en effet fréquent que l'analyse des réseaux techniques serve à préconiser la modification du réseau étudié, ce qui suppose la prise en compte des coûts [18] associés (créer un tronçon de TGV a un coût largement supérieur à l'ajout d'un.e ami.e sur Facebook, même si le maintien de liens sociaux a un coût - en temps, en énergie -, non négligeable. . .). À noter également que les réseaux techniques sont généralement redondants (existence de plusieurs chemins entre sommets voire multiplicité des liens) afin de gérer le mieux possible les aléas. Cette redondance rend les réseaux techniques plus complexes, plus coûteux mais permet un fonctionnement correct même en cas de perturbations récurrentes (cf le cas du métro parisien qui est dans les parties les plus denses un multigraphe non planaire).

Deux ouvrages ont exercé une influence majeure sur l'étude géographique des réseaux : une thèse, éditée en 1963²⁴ et un manuel édité en 1969²⁵. L'ouvrage de Kansky s'intéresse uniquement aux réseaux de transports dont les sommets et les lignes s'inscrivent sur le territoire - en excluant donc le transport aérien et maritime (« *The essential elements of transportation networks are thus man-made features located on the earth's surface in geometric patterns* »). Le spectre couvert par le manuel de Chorley et Haggett est plus large mais la focalisation sur les graphes planaires reste prédominante, tout comme la prise en compte de la valuation des liens (distance, intensité, etc.). Si le graphe est utilisé pour présenter les différents indicateurs proposés, les études empiriques sont illustrées très majoritairement par le marqueur dis-

24. K.J. KANSKY, *The structure of transportation networks*, Chicago, University of Chicago Press, 1963.

25. R.J. CHORLEY et P. HAGGETT, *Network analysis in geography*, Londres, Edward Arnold, 1969.

ciplinaire numéro un en géographie : la carte. Kansky propose pour l'étude des réseaux une série d'indicateurs inspirés par les travaux du mathématicien Claude Berge : simples à calculer, prenant en compte des indicateurs basiques (nombre de liens, nombre de sommets, distance, intensité des relations), ils restent aujourd'hui encore très largement utilisés. Très pédagogique, largement illustré par des petits graphes permettant de détailler le calcul (et l'interprétation) des indices présentés, et proposant une étude empirique solide sur le réseau ferré de Sardaigne, l'ouvrage reste une référence aujourd'hui.

Les méthodes classiques d'analyse sont le calcul d'indices avant et après un changement (réel ou projeté), la simulation de trafic et l'analyse multivariée des indicateurs calculés. Les études plus récentes sur la vulnérabilité²⁶ utilisent des méthodes proches dans la mesure où il s'agit de mesurer l'efficacité d'un réseau avant et après la suppression d'un sommet ou d'un lien. Trois tendances récentes peuvent être signalées : la prise en compte du multiviveau²⁷ pour calculer des accessibilités plus proches des pratiques réelles et prenant en compte l'intermodalité - utilisation de plusieurs réseaux lors des déplacements ; la différenciation des effets spatiaux et des effets [19] réseaux²⁸ (Gleyze, 2007) ; l'utilisation des outils de modélisation et de simulation (déplacements, gestion des risques). Ce dernier aspect était déjà présent dans Kansky, mais il a été grandement facilité par le développement des logiciels et des capacités informatiques.

Si deux approches disciplinaires ont été présentées ici, beaucoup d'autres existent et, sans prétention à l'exhaustivité, on citera notamment les approches développées en écologie depuis les années 1950, les approches développées en économie et mêlant souvent analyse de réseau et théorie des jeux ou encore les approches développées depuis la fin des années 1990 par les physiciens pour l'étude des grands réseaux. Cependant, l'importance des méthodes développées en analyse de réseaux sociaux pour les études historiques et la solidité du couple scolaire histoire-géographie en France justifiait à nos yeux cette restriction à deux disciplines seulement.

26. S. LHOMME, *Les réseaux techniques comme vecteur de propagation des risques en milieu urbain. Une contribution théorique et pratique à l'analyse de la résilience urbaine*, Thèse de géographie, Université Paris-Diderot-Paris 7, 2012.

27. L. CHAPELON, « Modélisation multi-échelles des réseaux de transport. Vers une plus grande précision de l'accessibilité », *Mappemonde*, 3, p. 28-36, 1996. <http://www.mgm.fr/PUB/Mappemonde/M396/Chapelon.pdf> (visité le 30 mars 2015) ; P. MATHIS (dir.), *Graphes et réseaux. Modélisation multiviveau*, Paris, Lavoisier, 2003.

28. J.F. GLEYZE, « Effets spatiaux et effets réseau dans l'évaluation d'indicateurs sur les nœuds d'un réseau d'infrastructure », *Cybergeo : European Journal of Geography*, 2007, <http://cybergeo.revues.org/5532> (visité le 15 mars 2015).

3 Quand et comment utiliser l'analyse de réseaux en histoire ?

Comme dans toute autre discipline, l'analyse de réseaux en histoire suppose qu'un phénomène de nature relationnelle est étudié, et que ce phénomène peut être symbolisé par des points symbolisant des acteurs et des lignes symbolisant des liens entre ces acteurs. L'une des utilisations les plus anciennes concerne l'étude des réseaux généalogiques qui sont d'un type particulier dans la mesure où ils mêlent des liens orientés (être le parent de) et des liens non orientés (être marié.e à). Si l'étude des lignages ne pose pas de souci méthodologique particulier, l'utilisation des méthodes génériques d'analyse de réseau pose un certain nombre de problèmes plus ou moins délicats. De manière schématique, on peut opposer un certain nombre de questions pertinentes à des problèmes plus apparents que réels ²⁹.

3.1 Quelques vraies questions...

Comme rappelé en introduction, l'analyse de réseau est devenue une méthode de plus en plus utilisée en histoire, et certains chercheurs s'inquiètent d'une mode qui risquerait de faire perdre toute pertinence à cet outil d'analyse et de représentation ³⁰. L'adoption par une science sociale de nouveaux outils entraîne fréquemment ce type de comportement : dans notre cas, tout deviendrait alors réticulaire et tout pourrait donc se prêter à une analyse de réseau. Cette attitude n'est pas totalement absurde [20] dans la mesure où il est rare qu'un phénomène donné ne puisse être en partie considéré comme relationnel... Et si l'on voit aujourd'hui dans les colloques de nombreuses communications où le graphe s'invite, peu de ces communications donnent ensuite lieu à des articles dans des revues à comité de lecture. S'il est en effet peu coûteux de produire une image réticulaire, parvenir ensuite à justifier son utilisation dans le cadre d'une problématique solide est en effet plus délicat - et ce problème n'est nullement le privilège de la discipline historique.

De très nombreuses méthodes ont été développées ces dernières décennies. S'il est toujours possible de mesurer tel ou tel indicateur sur n'importe quel graphe, cela ne signifie pourtant pas que toutes les mesures ont une pertinence thématique. Une démarche naïve consisterait à calculer tous les indicateurs possibles pour garder les résultats qui paraissent faire sens et chercher *a posteriori* à justifier leur utilisation. Or, l'utilisation d'outils quantitatifs ne dispense pas de commencer par formuler des hypothèses précises

29. Cette partie, écrite par un géographe, ne prétend pas adopter une position de surplomb vis-à-vis des pratiques historiennes en analyse de réseau ; il s'agit plus simplement de remarques inspirées par une fréquentation amicale et continue de mes collègues historien.ne.s.

30. G. GARROTE, « Réseaux : de la notion à l'analyse. Heurs et malheurs d'un outil », 2014, <http://reshist.hypotheses.org/464>, (visité le 10 mars 2015).

et de lier les indicateurs sélectionnés avec des enjeux thématiques³¹. Hormis peut-être la centralité de degré (et encore, si les liens sont valués, son calcul suppose des arbitrages) dont l'interprétation est relativement passe-partout, aucune mesure n'est en soi universelle.

Comme pour toute méthode de traitement quantitatif, le format des données doit également être réfléchi. Si la plupart des logiciels exigent en entrée une liste de liens et une liste de sommets, chacune de ces listes pouvant être complétée par des attributs, la mise en forme systématique n'est pas toujours aisée tant elle force à appliquer une logique unique et nécessairement réductrice à des données dont la nature est parfois différente (sources hétérogènes, plus ou moins grande certitude sur le degré de précision des données, etc.). Il est sans doute nécessaire alors d'adopter une attitude pragmatique : un niveau minimal d'homogénéité des données doit être atteint pour prétendre entreprendre la moindre démarche quantitative, ce qui n'empêche pas ensuite de nuancer les résultats obtenus en rappelant les éventuelles limites et restrictions interprétatives.

Au risque de rappeler des évidences, les trois aspects complémentaires rappelés ici (pertinence de la démarche, choix des indicateurs et formatage des données) ne sont pas propres à la démarche historique et ces questions se posent avec la même acuité en géographie, en sociologie ou en anthropologie. Pourtant, certaines spécificités propres à l'histoire peuvent à première vue apparaître comme des obstacles à l'utilisation de l'analyse de réseau.

3.2 ...et quelques et faux problèmes

Trois grands obstacles sont fréquemment évoqués par les chercheurs en histoire débutant en analyse de réseau : l'incomplétude des données ; la nature non réticulaire [21] du phénomène observé (avec l'argument d'autorité « Ceci n'est pas un vrai réseau ») et enfin la complexité propre à l'analyse de réseau.

L'incomplétude des données est fréquemment invoquée pour justifier la réticence à toute analyse quantifiée d'un phénomène, le réseau étant alors uniquement employé pour proposer une représentation graphique du phénomène étudié. Ce scrupule honore à première vue la communauté historique mais peut-être est-il utile de rappeler que, dans les autres disciplines (de la sociologie à la géographie en passant par l'informatique ou la physique), ces scrupules n'empêchent nullement les analyses quantitatives. Qu'il s'agisse d'étude sur les réseaux de l'internet, la circulation sur les réseaux routiers ou ferroviaires ou les réseaux personnels dans une collectivité donnée, les données analysées n'ont jamais la prétention à être exhaustives. . . Quel que soit le protocole de recueil de données utilisé, y compris pour des phénomènes

31. C. LEMERCIER et C. ZALC, *Méthodes quantitatives pour l'historien*. Paris, La Découverte, 2010.

contemporains, la saisie totale des relations entre telle ou telle communauté d'acteurs est illusoire. Il serait aisé de reprendre l'ensemble des articles classiques en analyse des réseaux sociaux par exemple pour montrer que l'échantillon finalement analysé n'est toujours qu'un échantillon dont la représentativité n'est pas assurée (taux de non-réponse distribué de manière non aléatoire, définition différente de termes de l'enquête selon les enquêtés, etc.³²). Cela n'a pourtant pas été considéré comme un obstacle majeur...

L'argument lié à la nature non réticulaire de l'objet étudié est fréquemment utilisé et ce quelle que soit la discipline. Ces dernières années, l'auteur de ces lignes aura par exemple entendu un géographe spécialiste des réseaux techniques asséner doctement que Facebook n'était pas un vrai réseau³³ ou un informaticien expliquer qu'en informatique, ils étudiaient des « vrais réseaux » donc l'approche réticulaire était pleinement justifiée. Ces arguments n'ont de sens que pour ceux qui les utilisent et prétendent ainsi borner un champ d'études de manière stricte - champ où ces personnes occupent par ailleurs le plus souvent des positions dominantes. Il est pourtant possible de considérer l'analyse de réseaux comme une boîte à outils méthodologiques susceptible d'éclairer de nombreuses thématiques de recherche, que celles-ci portent ou non sur de « vrais réseaux ». Cette notion de « vrai réseau » se heurte par ailleurs aux limites disciplinaires ; un vrai réseau pour un géographe orthodoxe n'a rien d'un réseau pour un sociologue (également orthodoxe) ou un informaticien. Une attitude heuristiquement féconde serait plutôt de considérer que, pour telle ou telle [22] question, l'analyse de réseau peut apporter des éclairages utiles et donc de l'utiliser sans complexe superflu.

Le dernier argument régulièrement entendu concerne la complexité des analyses, des mesures et des outils. Le recul ces dernières décennies de la maîtrise des outils quantitatifs au sens large, qu'il s'agisse de la maîtrise des méthodes statistiques ou de la gestion des bases de données par exemple, dans bon nombre de sciences sociales³⁴, recul qui se situe dans un contexte

32. On trouvera un rappel utile des limites liées au recueil des données en analyse des réseaux sociaux dans M. HENNIG, U. BRANDES, J. PFEFFER et I. MERGEL, *Studying Social Networks. A Guide to Empirical Research*, Campus Verlag, 2012 ainsi que la présentation des moyens utilisés pour limiter les biais dans C. BIDART, A. DEGENNE et M. GROSSETTI, *La vie en réseaux. Dynamique des relations sociales*, Paris, PUF, 2011.

33. Les sociologues préciseraient eux qu'il s'agit d'un réseau technique pouvant être le support de réseaux sociaux personnels. Mais utiliser des méthodes d'analyse de réseaux pour étudier Facebook ou Twitter est tout à fait pertinent.

34. En histoire, faut-il rappeler le constat déjà ancien de Bernard Lepetit : « L'histoire quantitative n'est plus à la mode », B. LEPETIT, « L'histoire quantitative : deux ou trois choses que je sais d'elle », *Histoire & Mesures*, 4(3-4), 1989, p. 191-199. Il rappelait d'ailleurs dans le même article que « la formation des historiens inclut rarement un apprentissage de la statistique qui dépasse le niveau le plus élémentaire ». Il n'est pas certain que la situation se soit beaucoup améliorée depuis.

épistémologique empreint d'un post-modernisme souvent hostile à toute formalisation stricte, explique sans doute ce réflexe défensif fréquent. Or, la grande majorité des méthodes disponibles (degré, densité, cliques notamment) se caractérise au contraire par leur grande simplicité dans la mesure où un phénomène donné est représenté par une matrice remplie de 0 et de 1... Et c'est d'ailleurs cette simplification des phénomènes observés (passage de relations souvent valués, multiplexes et dynamiques en un graphe booléen statique) qui explique sans doute la popularité pluri-disciplinaire de l'analyse de réseau aujourd'hui. Si certaines méthodes se caractérisent par un niveau d'exigence mathématique élevé³⁵, elles restent largement minoritaires. Et si la pléthore des outils disponibles peut déconcerter (plus de 80 logiciels dédiés à l'analyse de réseaux existent sur le marché³⁶), l'abondance de documentation disponible aujourd'hui permet de s'appropriier aisément les différents logiciels.

Il existe enfin un dernier aspect pouvant être problématique, celui relatif à la visualisation des données. Cet aspect jouant un rôle non négligeable dans la popularité actuelle de l'analyse de réseau, il nous a semblé nécessiter un développement spécifique.

3.3 Visualiser pour explorer, visualiser pour communiquer

L'analyse de réseau ne suppose pas plus que toute autre méthode quantitative une représentation graphique des données étudiées. Pourtant, très tôt (cf Moreno, 1934, *op. cit.*), l'habitude a été prise de montrer graphiquement la structure du réseau étudié³⁷. La représentation graphique de données quantitatives suppose le respect d'un certain de règles visant à favoriser la lisibilité de la figure produite. La représentation des grands réseaux peut poser des problèmes et nécessiter le recours à d'autres formes [23] de représentation. Reste qu'une première et nécessaire distinction doit être opérée entre visualisation exploratoire et visualisation communicationnelle.

La visualisation des données quantitatives fait partie intégrante de l'exploration des données et l'un des premiers réflexes lors de l'étude d'un réseau est d'étudier le graphe proprement dit (projection sur un plan des sommets et des liens). Dès les premiers sociogrammes dessinés manuellement par Moreno dans les années 1930, un certain nombre de conventions vont être utilisées pour représenter les graphes - elles vont ensuite s'imposer et être intégrées dans la plupart des algorithmes existants : 1) les sommets les plus connectés sont placés au centre de la figure, 2) les sommets les moins connectés sont

35. P.J. CARRINGTON, J. SCOTT et S. WASSERMAN, *Models and methods in social network analysis*, Cambridge Mass., Cambridge University Press, 2005.

36. <http://www.gmw.rug.nl/huisman/sna/software.html>, visité le 30 mars 2015.

37. F. BAHOKEN, L. BEAUGUITTE et S. LHOMME, « La visualisation des réseaux. Principes, enjeux et perspectives », Working paper du groupe fmr, 2013, <https://halshs.archives-ouvertes.fr/FMR> (visité le 17 mars 2015).

placés en périphérie de la figure 3) il faut limiter, autant que possible, le chevauchement des liens.

La visualisation comme outil exploratoire des données permet une grande souplesse : choix des algorithmes de visualisation justifié par la seule lisibilité interprétative des figures produites, inutilité d'une légende (l'auteur est le seul lecteur des figures), exhaustivité des données étudiées. Il ne faut pourtant pas surestimer l'intérêt de la visualisation : le plus souvent, un graphe brut est un graphe illisible et plus le nombre de relations est élevé, moins la figure produite est lisible³⁸.

L'objectif de la sémiologie graphique, développée en France par Jacques Bertin³⁹, est d'établir un certain nombre de règles visant à favoriser la lecture et l'interprétation correcte des graphiques produits, qu'il s'agisse de diagrammes, de cartes, de réseaux, etc. Selon Bertin, seule la variable visuelle taille permet de représenter des données quantitatives portant sur des stocks quand la valeur ou la couleur traduisent des relations quantitatives continues, segmentées en classes. Si la population de mon réseau est divisée en classe d'âge par exemple, il est autorisé d'utiliser des couleurs différentes (éventuellement graduées) pour colorer mes sommets mais la taille ne pourra varier qu'en fonction d'une quantité (le degré par exemple). Compléter ces règles par les trois principes de l'excellence graphique définis par Tufte⁴⁰ est un bon moyen pour produire des figures lisibles. Ces trois principes sont les suivants : « l'excellence graphique est une présentation bien conçue des données d'intérêt [...], [elle] vise à communiquer des idées complexes avec clarté, précision et efficacité ; [elle] est ce qui donne au lecteur le plus grand nombre d'idées dans le temps le plus court, tout en utilisant le moins d'encre sur un minimum d'espace »⁴¹.

Dans le cadre d'une figure destinée à transmettre un message (hors démarche exploratoire donc), il est nécessaire que cette image soit accompagnée d'une légende, tout comme il est recommandé de préciser les outils de visualisation utilisés (logiciel [24] et algorithmes). Il est fréquent de voir des images où les sommets sont de différentes formes, de différentes couleurs, les liens d'épaisseur variable et aucun indice ne permet de savoir ce que représentent ces formes, ces couleurs et ces épaisseurs...

Si la représentation de graphe peut être lisible pour quelques dizaines voire quelques centaines de sommets - et encore, si la densité de liens est faible -, il devient illusoire de penser produire une image quand le réseau étudié comporte quelques (centaines de) milliers de sommets. Si des cartes ou des

38. N. HENRY et J.D. FEKETE, « Représentations visuelles alternatives pour les réseaux sociaux », *Réseaux*, 6, p. 59-92, 2009.

39. J. BERTIN, 1967, *Sémiologie graphique. Les diagrammes - Les réseaux - Les cartes*, Paris, Éditions Gauthier-Villiar, 1967.

40. E.R. TUFTE, *The Visual Display of Quantitative Information*, Cheschire, Graphics Press, 1983.

41. *Ibid.*, p. 51, adapté par nos soins.

graphes de ce type sont régulièrement produits, l'enjeu est essentiellement technique. L'idée brillante des physiciens a été d'utiliser d'autres types de graphiques afin de révéler une propriété donnée des réseaux étudiés, par exemple montrer la distribution des degrés et tracer la courbe la résumant pour établir si le réseau est ou non sans échelle ⁴².

En guise de conclusion

Si l'analyse de réseau est un ensemble de méthodes ayant le vent en poupe actuellement, en histoire comme dans bien d'autres disciplines, il ne s'agit pourtant pas d'un outil susceptible d'apporter des réponses pertinentes pour toutes les questions de recherche. La souplesse de l'outil, l'aspect attrayant des visualisations (quand celles-ci sont travaillées et légendées), la diversité des approches possibles sont pourtant des atouts non négligeables et il est probable que son utilisation soit amenée à se développer encore dans les années à venir.

Un autre atout majeur de cette approche est sa propension à faciliter les rencontres, les discussions et les collaborations inter-disciplinaires ⁴³. De nombreuses sciences humaines et sociales, sans oublier la physique ou l'informatique, partagent aujourd'hui un vocabulaire, des méthodes d'analyse et des outils communs. Si les questions de recherche varient, et si le goût du dialogue entre disciplines n'est pas nécessairement partagé par tous les chercheurs, ce socle commun peut pourtant permettre d'enrichir nos pratiques respectives.

42. Un réseau sans échelle se caractérise par une très forte hiérarchie de degré des sommets : quelques sommets ont un degré très élevé, la très grande majorité a un degré très faible. Le degré moyen n'a donc pas de sens pour un tel réseau, d'où l'appellation de réseau sans échelle (*scale-free*).

43. La création par l'INSHS d'un groupement de recherche Analyse de réseaux en SHS début 2016 s'inscrit tout à fait dans cette dynamique - <http://arshs.hypotheses.org>.