



HAL
open science

L'analyse de réseaux bimodaux en sciences sociales. Aperçus épistémologiques et historiques

Laurent Beauguitte

► **To cite this version:**

Laurent Beauguitte. L'analyse de réseaux bimodaux en sciences sociales. Aperçus épistémologiques et historiques. 2021. hal-04002125

HAL Id: hal-04002125

<https://hal.science/hal-04002125>

Preprint submitted on 23 Feb 2023

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 bimodaux en sciences sociales

Aperçus épistémologiques et historiques

Laurent Beauguitte
CNRS, UMR Géographie-cités

Octobre 2021

Ce texte a été préparé pour la Sixième rencontre Réseaux et Histoire (Aix-en-Provence, 21-22 octobre 2021) ; des problèmes de santé m'ont empêché de m'y rendre. Ce document de travail pouvant peut-être avoir une petite utilité, je me suis finalement décidé à le rendre disponible.

L'analyse des réseaux unimodaux – étude des relations au sein d'un seul type de sommets – a de multiples origines : sociométrie (années 1930¹), analyse des flots dans les réseaux (années 1950), théorie des graphes (années 1950), analyse des réseaux sociaux en anthropologie (années 1960) et en sociologie (années 1970), analyse des réseaux d'infrastructures en géographie (années 1960), etc. Inversement, l'analyse de graphes bimodaux est plus récente et semble largement minoritaire dans les différentes sciences sociales. Il reste d'ailleurs très fréquent, quelle que soit la discipline considérée, de s'affranchir de la structure bipartite des données de départ en les transformant en un réseau unimodal valué et non orienté.

Cette communication vise trois objectifs complémentaires concernant l'analyse de réseaux bipartis : préciser le vocabulaire employé et différencier graphe biparti, bimodal et réseau d'affiliation ; retracer brièvement l'histoire de l'étude des graphes bipartis en théorie des graphes ; étudier l'apparition de l'analyse de réseaux bipartis en sociologie, en écologie² et en archéologie.

Ces quelques pages ne prétendent pas fournir un tableau exhaustif de ces différentes questions mais fournir un certain nombre de points de repères,

1. Les décennies entre parenthèses indiquent la période où le programme de recherche considéré s'autonomise et s'installe de façon durable dans la discipline concernée. Il existe bien entendu des articles relatifs à la théorie des graphes avant les années 1950 mais son autonomisation vis-à-vis de la topologie ne se fait que lors de cette décennie (Beauguitte, 2022).

2. L'écologie étudiant des écosystèmes d'êtres vivants, elle est considérée ici comme une science sociale.

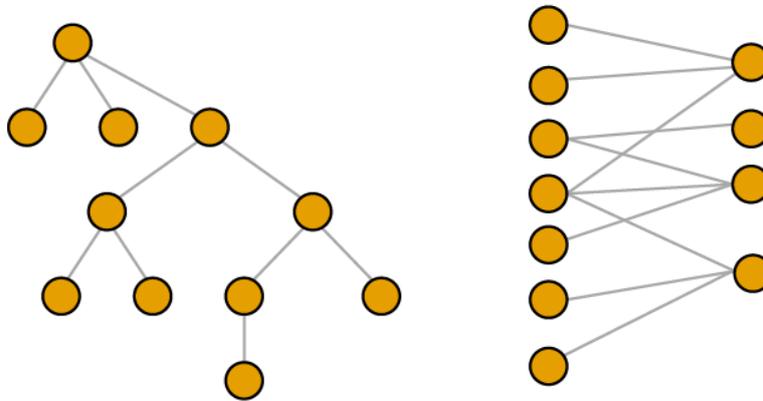
points qui tous nécessiteront des approfondissements ultérieurs. Il s'agit ici davantage de pistes de réflexion que des conclusions fermes et définitives.

1 Questions de vocabulaire

Le premier objectif concerne la clarification épistémologique relative aux termes utilisés : quelles sont les différences méthodologiques et conceptuelles entre des termes parfois employés de manière indifférenciée comme graphe biparti, graphe bimodal ou réseau d'affiliation ? Il est en effet nécessaire d'un point de vue conceptuel de distinguer structure mathématique et construction des données relationnelles.

Un graphe est dit biparti (*bipartite graph*) si son ensemble de sommets V peut être divisé en deux sous-ensembles disjoints $V1$ et $V2$ tel que chaque arête ait une extrémité dans $V1$ et l'autre dans $V2$. Cela ne suppose pas qu'il y ait une quelconque différence de nature, de statut ou d'attribut entre les sommets de ces deux ensembles. D'un point de vue mathématique, tout arbre peut par exemple être considéré comme un graphe biparti (figure 1).

FIGURE 1 – Un même graphe, deux représentations graphiques possibles



En théorie des graphes, tout arbre (graphe connexe acyclique) et toute forêt (graphe non connexe acyclique) est un graphe biparti : il est en effet toujours possible de créer une partition des sommets en deux sous-ensembles $V1$ et $V2$ telle que tout lien ait une extrémité dans $V1$ et l'autre dans $V2$.

Lorsqu'on parle de graphe bimodal (*two-mode network*), on évoque un graphe biparti comprenant un attribut permettant de différencier les deux sous-ensembles de sommets. Si j'étudie des relations sentimentales strictement hétérosexuelles, je peux les symboliser par un graphe bimodal des relations hommes – femmes. Si je m'intéresse aux relations entre enfants et

enseignantes, idem. Mes deux catégories de sommets peuvent être de nature différente (lien entre un groupe musical et un label, entre une autrice et une revue, etc.) mais ce n'est pas obligatoire.

Le terme de réseau d'affiliation (*affiliation network*) suppose également un graphe biparti entre des personnes P appartenant/affiliées à des organisations O . Il s'agit donc d'un cas particulier de graphe bimodal. Le terme de réseau d'affiliation n'est pas synonyme d'*interlock*, ce dernier terme désignant en analyse de réseaux sociaux un cas particulier de réseau d'affiliation, à savoir l'appartenance de personnes à des conseils d'administration. Wasserman et Faust signalent que d'autres termes ont été proposés pour désigner ce type de réseaux comme *membership network*, *dual network* ou *hypernetwork* (1994, p. 291 et 295).

La structure mathématique de l'objet est la même dans les trois cas (deux sous-ensembles de sommets), le choix du terme dépend du type de sommets (affiliation) ou d'attribut (bimodal) considéré. Enfin, lorsque l'on souhaite étudier les relations entre deux sous-ensembles de sommets et les relations à l'intérieur de ces deux sous-ensembles, on parlera d'analyse multiniveaux.

Le tableau 1 présente les résultats obtenus sur le moteur de recherche *Google Scholar*. Ce moteur de recherche privilégie les ressources récentes, ressources qui ne cessent pas ailleurs d'augmenter de manière exponentielle ces dernières décennies. Il est donc attendu que le nombre de références augmente fortement. Le découpage en décennies est arbitraire ; le choix de la borne 1950 est lié à la réédition du premier manuel de théorie des graphes (König, 1950³).

TABLEAU 1 – Fréquences des termes sur *Google Scholar*

	bipartite graph	two-mode network	affiliation network
1900-1949	10	0	0
1950-1959	9	0	0
1960-1969	183	0	0
1970-1979	1 070	2	7
1980-1989	3 600	7	21
1990-1999	11 200	30	57
2000-2009	22 500	415	877
2010-2019	28 900	2 680	2 480

Les cases « inclure les brevets » et « inclure les citations » ont été décochées. Requêtes effectuées le 9 août 2021.

L'outil *JSTOR Data for Research* permet d'affiner ces résultats : la prime

3. L'ouvrage, écrit en allemand, paraît en 1936 à Leipzig mais le contexte international freine sa diffusion ; le manuel de König est réédité (mais non traduit) à New York en 1950 et devient alors accessible.

aux articles récents est moins flagrante - même si l'augmentation du nombre d'articles et de revues est constante - et les requêtes permettent de savoir quelles disciplines utilisent ces termes, un même article pouvant être classé dans différentes disciplines. Le tableau 2 fournit les occurrences des différents termes par décennies et le nombre de disciplines mobilisant ces termes.

TABLEAU 2 – Fréquence des termes et disciplines associées sur *JTOR Data for Research*

	bipartite graph	two-mode network	affiliation network
1900-1949	0	0	0
1950-1959	1	0	0
1960-1969	43 (6 - 70%)	0	0
1970-1979	113 (11 - 74%)	0	1
1980-1989	100 (10 - 74%)	3 (2 - 67%)	3 (2 - 67%)
1990-1999	253 (18 - 62%)	10 (4 - 80%)	6 (7 - NA*)
2000-2009	333 (27 - 66%)	13 (6 - 46%)	38 (25 - 32%)
2010-2019	559 (38 - 63%)	102 (26 - 40%)	60 (21 - 38%)

Requêtes effectuées le 9 août 2021 : option « All content » cochée. Le nombre entre parenthèses indique le nombre de « Subject », une même ressource peut être attribuée à plusieurs thèmes. Le pourcentage qui suit indique la proportion de la discipline dominante lorsqu'il en existe une.

La lecture de la case 1960-1969/bipartite graph est donc la suivante : 43 ressources contiennent le terme « bipartite graph », elles correspondent à 6 disciplines, la plus représentée regroupe 70% des références.

**Lors de la période 1990-1999, 6 ressources utilisent le terme « Affiliation network », les 7 disciplines correspondantes n'apparaissent qu'une fois.*

Les effectifs ne sont pas comparables avec ceux de *Google Scholar* mais apportent des enseignements complémentaires. Tout d'abord, les contenus contenant le terme « bipartite graph » restent tout au long de la période majoritairement issus des mathématiques (entre 62 et 74% des contenus selon les décennies). Inversement, à partir de 1990, les termes « two-mode network » et « affiliation network » sont appropriés par différentes disciplines ; si la sociologie reste en tête pour presque toutes les décennies (elle est à égalité en 2000-2009 avec « Business » pour le terme « affiliation network »), de nombreuses autres disciplines mobilisent ces expressions. L'étude de l'analyse des graphes bipartis en théorie des graphes permet de compléter ces premières informations.

2 Graphe biparti et théorie des graphes

Plusieurs articles étudiant les propriétés des graphes bipartis peuvent être trouvés dans les revues de mathématiques à partir de la fin du XIX^e siècle.

Néanmoins, le sous-champ de la topologie étudiant les graphes n'est pas à l'époque suffisamment autonome pour que ces articles permettent de créer une école ou un programme de recherche à part entière.

Dans le premier manuel de théorie des graphes, publié en allemand à Leipzig en 1936 et réédité mais non traduit à New York en 1950, les graphes bipartis sont évoqués à plusieurs reprises par le mathématicien hongrois Dénes König (voir notamment le chapitre XI). Par contre, dans l'ouvrage de Claude Berge *La théorie des graphes et ses applications*, publié par Dunod en 1958 et traduit en anglais en 1962, le terme est très peu présent. Le terme fait peu à peu son apparition dans les manuels parus dans les années 1960 (Ore, 1962) mais de façon non systématique (il reste absent de Harary *et al.*, 1965 par exemple). Le manuel de Berge daté de 1970 et traduit en anglais en 1973, *Graphes et hypergraphes*, consacre plusieurs sections à l'étude des graphes bipartis et le terme « biparti » compte plus de 100 occurrences dans l'édition française. Le premier manuel exclusivement dédié à l'analyse des graphes bipartis est relativement récent dans la mesure où il date de 1998 (Asratian *et al.*).

La question de la transformation des graphes bipartis en graphe unimodaux valués n'est à ma connaissance jamais abordée dans les manuels de théorie des graphes : l'opération est mathématiquement très simple (multiplication de la matrice d'adjacence par sa transposée⁴) et n'a donc pas un grand intérêt, l'étude des graphes unimodaux valués est peu abordée, et surtout les théorèmes que cherchent à produire les mathématicien.nes ne supposent que rarement la transformation de l'objet étudié. Plus généralement, il semble que les problèmes que se posent les mathématicien.nes sur les graphes bipartis sont difficilement transposables aux données relationnelles analysées en science sociale – cette hypothèse mériterait évidemment un travail confirmatoire conséquent pour être éventuellement validée⁵.

Si l'analyse des graphes bipartis devient une opération classique en théorie des graphes (classique car mise en manuel), il faut une petite dizaine d'années pour que l'ensemble des manuels produits abordent cette question. Or la production de manuels en sciences sociales et s'inspirant de façon très variable de la théorie des graphes a été très rapide. Dans les années qui suivent la traduction de l'ouvrage de Berge, des manuels consacrés à l'analyse des réseaux en économie (Avondo-Bodino, 1962), en psychologie (Flament, 1963), en géographie (Haggett et Chorley, 1969) paraissent ; tous citent le manuel

4. La transposée T d'une matrice $Mn \times m$, notée $T(M)$, est la matrice $m \times n$. L'opération $T \times T(M)$ permet d'obtenir les liens valués entre les sommets n ; l'opération $T(M) \times T$ les liens valués entre les sommets m .

5. De manière générale, le travail mathématique sur l'objet graphe consiste à produire des théorèmes et à énoncer des propriétés aussi générales que possible ; l'analyse de réseau en science sociale cherche au contraire à caractériser un phénomène relationnel précis symbolisé par un graphe particulier.

de Berge (soit la version française de 1958, soit la traduction anglaise de 1962) et tous se concentrent sur les seuls graphes unimodaux. Cette focalisation sur les graphes unimodaux reste prégnante dans nombre de manuels d'analyse de réseaux en sciences sociales (voir par exemple et en se limitant à la seule analyse des réseaux sociaux : Knoke et Kuklinski, 1982 ; Berkowitz, 1982 ; Degenne et Forsé, 2004).

3 Des programmes disciplinaires indépendants : réseaux écologiques, artefacts et *interlocks*

Des graphes montrant les interactions entre espèces sont utilisés dès la fin du XIX^e siècle pour illustrer le fonctionnement de réseaux trophiques mais ces représentations n'ont eu aucun impact avant l'ouvrage de CS Elton, *Animal Ecology* paru en 1927 (Bersier, 2007). Les modèles mathématiques utilisés, notamment les modèles proie-prédateur de Lotka–Volterra développés dans les années 1920, sont fondés sur des équations différentielles (Ings et Hawes, 2018). Les réseaux trophiques (qui mange qui?) ne sont pas des graphes bipartis : il n'y a pas des proies d'un côté et des prédateurs de l'autre, chaque espèce ayant ses proies et ses prédateurs. De plus, il peut exister des cycles, des boucles (cannibalisme), des chaînes de longueur plus ou moins grandes, et les liens sont orientés. La visualisation sous forme de graphe biparti est possible pour les interactions insectes – plantes liées à la pollinisation : dans ce cas, il y a bien deux ensembles distincts de sommets et une relation entre ces deux ensembles. Néanmoins, avant les années 2000 et l'impact des travaux des physiciens, les indicateurs mis au point se basent sur le calcul matriciel et sur les équations différentielles et non sur la théorie des graphes (Holt, 1977).

L'étude des réseaux d'affiliation, et notamment l'étude dite des *interlocks* définis par Mizruchi comme étant la présence d'une personne affiliée à une organisation dans le conseil d'administration d'une autre organisation (Mizruchi, 1996), se développe dans les années 1970 dans la sociologie nord-américaine. Trois articles⁶ peuvent illustrer ce développement. En 1970, Perruci et Pilisuk cherchent à étudier les élites d'une ville moyenne (50 000 habitants) à l'aide d'un recueil de données exhaustif (liens individus – firmes) et d'entretiens où des générateurs de noms sont utilisés. Le

6. Pourquoi ces trois articles? Parce qu'ils sont accessibles en ligne et qu'ils sont cités dans la plupart des articles postérieurs relatifs à cette thématique. Ce choix revient à donner une « prime aux vainqueurs » dans la mesure où seules les propositions ayant eu du succès sont examinées. Un travail plus approfondi permettant de mettre en évidence les programmes de recherche concurrents serait nécessaire. Sans nécessairement être complètement convaincu par la possibilité de mettre en pratique le « programme fort » en histoire des sciences (ie étudier de façon symétrique la théorie qui l'emporte et celles qui échouent dans une controverse scientifique), s'intéresser à la pluralité des pistes proposées à un moment donné, et pas uniquement à celle a été retenue, me semble utile.

vocabulaire annonce celui la *Social network analysis* (SNA) à venir mais, si une représentation des liens inter-organisationnels est proposée (figure 2), les mesures sont des pourcentages liés aux propriétés des individus. Aucune référence n'est faite ni à la théorie des graphes, ni à la sociométrie.

FIGURE 2 – Une représentation graphique sans analyse de réseau

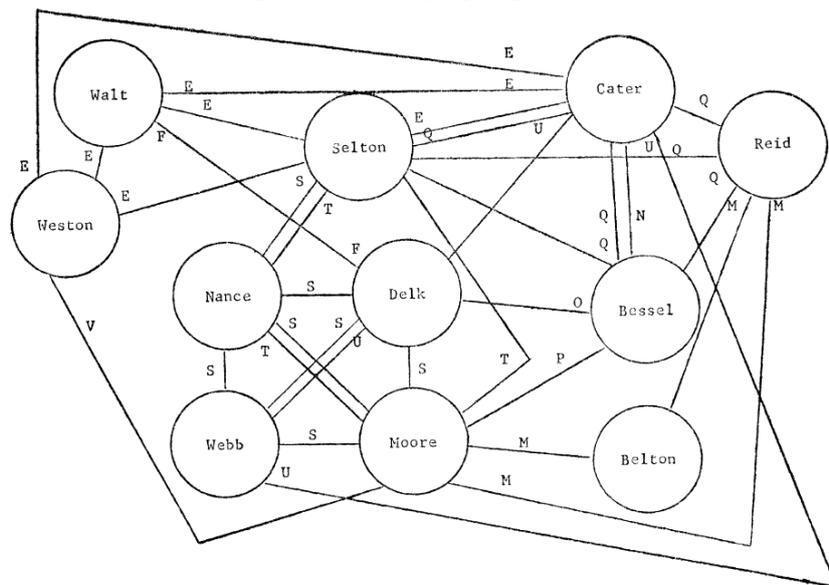


FIGURE 1. COMPLETE RESOURCE NETWORK GENERATED FROM OVERLAPS IN ORGANIZATION "S"

Figure extraite de R. Perrucci et M. Pilisuk, 1970, p. 1054 et représentant les liens entre dix personnes occupant une position de pouvoir dans la communauté étudiée. Les étiquettes des liens ont la signification suivante : « D-advisory commission to city government ; E, M, P-bank ; G, J-welfare organization ; L-voluntary association ; F, Q-hospital board ; S-educational board ; T-voluntary association ; U-educational board ; N, V-industrial firm. »

En 1972, Levine cherche à cartographier le réseau des liens entre banques et firmes. Le lien avec la SNA est explicite tant dans les remerciements de la page 14 (Philip Bonacich et Harrison White sont cités) que dans le vocabulaire (*network, node, link, connected network, shortest path*, etc.). Il cherche par ailleurs une alternative au diagramme lien-nœud (il utilise les expressions *line-drawing representation* et *point-line sociogram*) pour visualiser le réseau et propose une méthode statistique multivariée (*multidimensional scaling*). Enfin, l'article de Breiger (1974) synthétise les différentes approches antérieures (notamment les deux articles évoqués à l'instant), mobilise la théorie des graphes (Alba, Harary) et les récents travaux de la SNA (Bonacich, White) et propose un cadre théorique et méthodologique permettant l'analyse des réseaux d'affiliation tout en distinguant l'approche proposée de la sociométrie. Sa proposition, qui sera largement reprise ensuite, suppose

la transformation de la matrice rectangulaire individus-événements en une matrice individu-individu et donc la perte du caractère bimodal du réseau originel. On note enfin une forte mathématisation de la démonstration, les équations et les symboles étant présents tout au long de l'article, caractéristique fréquente dans « l'école de Harvard ».

L'archéologie à partir des années 1970 utilise la coprésence d'artefacts similaires en des lieux différents pour inférer des relations entre ces lieux (Collar *et al.*, 2015). Là encore, comme dans le cas des *interlocks*, le caractère bimodal des données de départ (matrice rectangulaire artefacts – sites) est rapidement effacé pour créer des cartes de relations supposées entre sites. L'originalité majeure de ce programme de recherche réside dans les références mobilisées qui piochent simultanément dans la théorie des graphes (Harary) l'analyse spatiale et la géographie quantitative (Christaller, Chorley et Haggett, Berry, etc.) et la *SNA* (voir par exemple Irwin-Williams, 1977 ou Rothman, 1987).

Il est important de souligner que des travaux plus anciens, notamment les travaux des français Gardin et Garelli datés du début des années 1960 et étudiés par Sébastien Plutniak (2018), mobilisaient la théorie des graphes ; leur réception a été trop tardive pour qu'ils influencent les travaux anglophones cités ci-dessus.

Conclusion provisoire

Si la théorie des graphes étudie relativement tôt les propriétés des graphes bipartis, la prise en compte des réseaux bimodaux dans les différentes sciences sociales évoquées dans cette communication est relativement tardive et incomplète. Les réseaux écologiques ne font pas appel à la théorie des graphes *stricto sensu* avant les années 2000 et s'inspirent des travaux des physiciens. En *SNA* comme en archéologie, la nature bimodale des données est le plus souvent invisibilisée et seules les relations de cooccurrence sont analysées. D'autres disciplines comme la géographie ou l'anthropologie aux mêmes périodes (années 1970) travaillent quasi exclusivement sur des réseaux unimodaux.

En ce qui concerne les généalogies disciplinaires, l'écologie se place nettement à part : si nombre de figures produites rappellent l'analyse de réseau, ses modèles se fondent sur les équations différentielles et secondairement sur le calcul matriciel. La *SNA* se différencie de la sociométrie en s'appuyant sur la théorie des graphes (vocabulaire et mesures) et développe des méthodes et concepts spécifiques (équivalence, *blockmodel*, rôle et position, etc.). Enfin, l'archéologie est plus hybride, piochant théorie et méthodes tant en géographie quantitative qu'en théorie des graphes et en *SNA*.

Ces quelques pistes mériteraient de plus larges développements, notamment en ce qui concerne l'archéologie qui a été trop rapidement abordée ici. La généalogie des influences concernant l'étude des réseaux d'affiliation pourrait être approfondie. Enfin, une étude chronologique systématique concernant l'analyse mathématique des graphes bipartis (mesure, propriété, théorème) pourrait permettre de repérer plus aisément la porosité éventuelle entre théorie des graphes et sciences sociales et entre les différentes sciences sociales.

Références

- Asratian A. S., Denley T. M. J. et Häggkvist R. (1998), *Bipartite graphs and their applications*, Cambridge tracts in mathematics, 131, Cambridge University Press.
- Avondo-Bodino G. (1962), *Economic applications of the Theory of Graphs*, Gordon and Breach.
- Beauguitte L. (2022), « Théorie des graphes et analyse de réseau en géographie : histoire d'un lien faible (1950-1963) », *Passerelles SHS*, 1. [[En ligne.](#)]
- Berge C. (1958), *Théorie des graphes et ses applications*, Dunod.
- Berge C. (1970), *Graphes et hypergraphes*, Dunod.
- Berkowitz S. D. (1982), *An introduction to structural analysis : The network approach to social research*, Butterwoth & Co.
- Bersier L.-F. (2007), « A history of the study of ecological networks », *Biological networks*, 3, 365-421.
- Breiger R. L., 1974, « The duality of persons and groups », *Social forces*, 53(2), 181-190.
- Collar A., Coward F., Brughmans T. et Mills B. J. (2015), « Networks in archaeology : phenomena, abstraction, representation », *Journal of archaeological Method and theory*, 22(1), 1-32.
- Degenne A. et Forsé M. (2004), *Les réseaux sociaux*, 2e éd. U, Armand Colin.
- Flament C. (1963), *Applications of graph theory to group structure*, Prentice-Hall.
- Haggett P. et Chorley R. J. (1969), *Network analysis in geography*, Edward Arnold.
- Harary F., Norman R. Z. et Cartwright D. (1965), *Structural models : An introduction to the theory of directed graphs*, John Wiley & Sons.

- Holt R. D. (1977), « Predation, apparent competition, and the structure of prey communities », *Theoretical population biology*, 12(2), 197-229.
- Ings T. C. et J. E. Hawes (2018), « The history of ecological networks », *Ecological networks in the tropics*, 15-28.
- Irwin-Williams C. (1977), « A network model for the analysis of prehistoric trade », in T. K. Earle et J. E. Ericson (dir.), *Exchange systems in Prehistory*, *Academic Press*, 141-152.
- Knoke D. et Kuklinski J. H. (1982), *Network analysis*, Quantitative applications in the social sciences, 28, SAGE.
- Levine J. H. (1972), « The sphere of influence », *American sociological review*, 37(1), 14-27
- Mizruchi M.S. (1996), « What do interlocks do? An analysis, critique, and assessment of research on interlocking directorates », *Annual review of sociology*, 22(1), 271-298.
- Perrucci R. et Pilisuk M. (1970), « Leaders and ruling elites : The interorganizational bases of community power », *American sociological review*, 35(6), 1040-1057.
- Plutniak S. (2018), « Aux prémices des humanités numériques? La première analyse automatisée d'un réseau économique ancien (Gardin & Garelli, 1961). Réalisation, conceptualisation, réception », *ARCS - Analyse de réseaux pour les sciences sociales*. [[En ligne.](#)]
- Rothman M. S. (1987), « Graph theory and the interpretation of regional survey data », *Paléorient*, 13(2), 73-91.
- Wasserman S. et Faust K. (1994), *Social network analysis. Methods and applications*. Structural analysis in the social sciences, Cambridge University Press.