



Analyses dynamiques et longitudinales de réseaux

Marion Maisonobe

► **To cite this version:**

Marion Maisonobe. Analyses dynamiques et longitudinales de réseaux. Groupe f.m.r. (flux, matrices, réseaux) animé par César Ducruet et Laurent Beauguitte (CNRS, U.. 2013. <halshs-00811672>

HAL Id: halshs-00811672

<https://halshs.archives-ouvertes.fr/halshs-00811672>

Submitted on 10 Apr 2013

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.

Analyses dynamiques et longitudinales de réseaux

Marion Maisonobe, CNRS, UMR LISST-CIEU
marion.maisonobe<a>univ-tlse2.fr

Version 1 - Avril 2013



Introduction

La théorie des graphes permet de rendre compte de la position d'éléments dans une structure. S'en servir suppose de faire un effort de formalisation consistant à définir rigoureusement quels sont les éléments (les nœuds ou sommets) et quelle est la nature de la relation supposée structurante qui les unit (les liens ou arcs). Les liens traduisent des relations d'appartenance à un ensemble qui est inclus dans l'ensemble plus grand que constitue la population du graphe (les nœuds). Dans un graphe de collaborations scientifiques, si les nœuds sont des chercheurs en géographie, ils appartiennent tous à l'ensemble que constitue « la communauté scientifique ». Le graphe de collaboration rend compte de l'appartenance de ces chercheurs à des sous-ensembles définis par le fait de partager une propriété supplémentaire en commun : avoir co-écrit un article scientifique. Cet exemple illustre le fait qu'en général, les attributs traduisent des relations d'appartenance [34]. En effet, dans le graphe de collaboration, l'appartenance à la communauté scientifique est une propriété des nœuds ou attribut. Parce que la relation attributaire est plus stable dans le temps, on a tendance à la prendre pour acquise¹. Dans ce qui suit, on essaye de dépasser la distinction entre réseau d'appartenance et réseau social pour lui préférer la distinction plus généralisable de Ron Atkin entre « décor » et « trafic ». Ces distinctions sont utiles pour saisir la dynamique des structures. Expérimentale, cette synthèse prétend avant tout présenter les différents corpus d'études qui inscrivent les réseaux dans le temps.

1. Dire qu'il n'y aucune différence essentielle entre relation d'appartenance et relation sociale ne fait pas l'unanimité parmi les analystes de réseaux sociaux. Cette position entre en contradiction avec celles d'Erving Goffman, James Coleman et Ronald Breiger [10].

Penser en dynamique avec Ron Atkin : *Backcloth* et *traffic*

En 1980, Peter Gould, reprenant les travaux du mathématicien Atkin, considère la théorie des graphes trop pauvre pour rendre compte de la multidimensionalité du monde et au-delà de la complexité des structures : « Mais les structures sont si complexes dans les systèmes naturels et humains que même le graphique le plus beau et sophistiqué ne suffit pas à en rendre compte... »² [17]. Pour remédier à ce problème, il propose de recourir à la *Q-analysis* ou *Polyhedral Dynamics* développée par Ron Atkin dans les années 1970. L'apport théorique majeur de cette approche consiste en la distinction entre le décor (*backcloth*) et le trafic (*traffic*). Dans une pièce de théâtre, le décor contraint le déroulement de l'action sur la scène (s'il y a un obstacle, les acteurs doivent le contourner). Pour autant, le décor n'est pas fixe : il peut être modifié pendant le spectacle et alors modifier le cours de l'action. D'un autre point de vue, c'est l'action elle-même qui a des incidences sur le décor. Appliquée à l'exemple précédent, l'appartenance à la communauté scientifique fait partie du décor et le fait de collaborer représente le trafic. L'ensemble obtenu n'est pas stable dans le temps : des membres peuvent intégrer ou quitter la communauté au cours du temps (changement de décor) ; leurs collaborations sont des événements ponctuels qui relèvent du trafic. *A priori*, tous les membres de la communauté sont susceptibles de collaborer entre eux mais il est envisageable d'ajouter une pointe de complexité en considérant que les membres de la communauté travaillant dans le même laboratoire (ou ayant été collègues auparavant), ont une probabilité plus forte de collaborer. Dans la logique d'Atkin, l'analyse du réseau de collaboration gagne en profondeur si elle est effectuée au regard de la structure sous-jacente (le « complexe simplicial » dans lequel les chercheurs se trouvent et les collaborations se déroulent). En 1980, tandis que Peter Gould publie son introduction à la *Q-analysis* pour les géographes et urbanistes, Linton Freeman et Patrick Doreian, deux sociologues intéressés par l'évolution de la structure des réseaux sociaux l'expérimentent [16][15]. Tous deux concluent qu'il n'y a rien qu'ils aient pu faire avec la *Q-analysis* qui n'aurait pu être exprimé par la théorie des graphes. En revanche, ils insistent sur le caractère heuristique de la distinction entre décor et trafic³. Aujourd'hui, les analyses de graphes bipartis se sont substituées à la

2. « *But structural matters can become so complex in natural and man-made systems that even sophisticated and beautifully drawn graphics begin to break down. The basic problem is the multidimensional nature of complexity, and the fact that the objects of our inquiries eventually force us to abandon the limited two dimensions of a page, and even the third dimension we can capture by playing a trick on our eyes with perspective.* », Gould, 1980 [17].

3. « *My final reaction is mixed. On one hand, there is nothing in this analysis that could not have been expressed—often more simply—using the terms and concepts of graph theory. On the other hand, Atkin has stressed a number of ideas—like the distinction between backcloth and traffic—that would probably not have occurred to me had I been using ordinary*

Q-analysis d'Atkin, laquelle n'est plus tellement utilisée [7].

Structures et dynamiques

Que Patrick Doreian, pionnier dans les réflexions sur la dynamique des réseaux sociaux, se soit intéressé aux travaux d'Atkin n'est pas anodin. Ces travaux invitent les analystes des réseaux sociaux à davantage de rigueur en les incitant à ne s'intéresser au trafic – et notamment à ne réaliser des analyses diachroniques – qu'une fois le décor posé. Cette distinction mérite d'être remise à l'ordre du jour compte tenu des évolutions récentes de l'analyse de réseaux. Les données mises à disposition des chercheurs au cours de la dernière décennie – à partir du suivi instantané de l'activité des internautes, des signaux d'appel générés depuis les téléphones portables et plus globalement, issues de la généralisation du *badging* – aussi qualifiées de « traces », se prêtent particulièrement bien aux analyses de trafic ; mais au risque, éventuellement, de réinventer l'eau chaude en oubliant de tenir compte du décor. Dans la logique d'Atkin, il s'agirait de se demander en quoi les contacts ou interactions observés en temps réel sont surprenants compte tenu du décor. Dans les années 2000, un vocabulaire dédié s'est développé parmi les spécialistes du traitement de données relationnelles repérées dans le temps. Un graphe temporel, dynamique ou *Time-dependent network* est un graphe dont la structure évolue dans le temps. Si le vocabulaire n'est pas stabilisé, c'est aussi le cas de la définition. La majeure partie des auteurs entend par « graphe temporel », non pas un graphe mais une série temporelle de graphes statiques (chez certains, le nombre de nœuds est fixe et seuls les liens varient d'un pas de temps à l'autre ; chez d'autres encore, la position des nœuds dans l'espace est considérée et elle évolue dans le temps). Ainsi, les travaux réalisés sur l'évolution des réseaux sont, pour l'essentiel, des études longitudinales de réseaux. Bivas Mitra *et alii* évoquent les limites de ces définitions et proposent une définition plus large, adaptée à l'étude de réseaux de citations [30]. Un seul graphe est défini au sein duquel les liens peuvent traverser le temps ; les nœuds aux extrémités d'un lien peuvent appartenir à des périodes différentes. Cette définition vaut également pour les graphes mobilisés en analyse de réseaux d'évènements, de lignages et de trajectoires biographiques (les approches qualifiées dans ce qui suit d'alternatives aux approches standards). Trois catégories de travaux se distinguent dans le domaine de l'analyse dynamique de réseaux. La synthèse aborde successivement la modélisation des dynamiques ; la recherche de « communautés » ou « groupes » cohérents dans le temps ; et les approches alternatives relevant de près ou de loin du « programme narratif » défini par Andrew Abbott [1].

network analytic tools. », Freeman, 1980 [16].

1 La modélisation

1.1 Sociométrie et analyse des réseaux sociaux

Détecter les principes d'évolution d'une structure dans le temps et prévoir son évolution est une préoccupation qui remonte aux années 1950. L'analyse des réseaux sociaux telle qu'on la connaît aujourd'hui est alors qualifiée de « sociométrie ». Mise au point par le psychiatre et psychologue Jacob L. Moreno, la sociométrie permet d'analyser les relations interpersonnelles d'un groupe à un moment donné. En 1959, Leo Katz et Charles H. Proctor sont les premiers à mobiliser des chaînes de Markov discrètes pour analyser l'évolution des affinités dans une salle de classe en comparant quatre « sociogrammes » statiques [25]. Un processus de Markov d'ordre 1 est un processus possédant la propriété suivante : il est possible de prévoir un phénomène à l'état $t+1$ (le futur) en se référant uniquement à l'état du phénomène au temps t (le présent). Cela suppose que l'état du phénomène au temps $t-1$ (le passé) n'a pas d'incidence sur son état au temps $t+1$. Dans le cas contraire, il faut mobiliser des chaînes de Markov d'ordre 2 ou plus. Une fois les probabilités de transition des relations d'un état à un autre estimées, Katz et Proctor en déduisent un modèle de Markov d'ordre 1. Ce modèle se révèle adapté pour décrire l'évolution du groupe dans son ensemble (dont la configuration a été relevée par questionnaire auprès des élèves à quatre moments de l'année). Stanley Wasserman et Dawn Iacobucci expliquent que ce type de modélisation permet de répondre à deux questions de recherche : comment évolue un phénomène dans le temps ? Dans quelle mesure l'histoire du phénomène permet de prévoir son futur ? Cela suppose d'expliquer la structure du réseau lui-même par les stratégies individuelles supposées des acteurs. Katz et Proctor se limitent à estimer l'effet du principe de réciprocité (posons deux individus A et B, si A apprécie B au temps t , quelle est la probabilité pour que B apprécie A en retour au temps $t+1$). Le modèle log-linéaire défini par Wasserman et Iacobucci en 1988, quant à lui, en plus des effets de réciprocité, permet de mesurer l'impact des effets de popularité (plus connus aujourd'hui sous le nom d'attachement préférentiel) sur l'évolution du réseau [42].

Les modèles qui ont rencontré le plus de succès en analyse de réseaux sociaux dans les années 1980 et 1990 (p1, p2) sont des modèles log-linéaires adaptés à l'étude de réseaux statiques [6]. En revanche, en analyses longitudinales de réseaux sociaux, les modèles log-linéaires ont progressivement été délaissés au profit de modèles fondés sur des chaînes de Markov continues. En effet, l'inconvénient du modèle de Katz et Proctor (chaîne de Markov discrète) et du modèle de Wasserman (modèle log-linéaire) évoqués plus haut, est de ne pas tenir compte de ce qu'il peut arriver dans le réseau entre deux enregistrements y compris si ces enregistrements sont très espacés dans le temps (tous les 4 mois dans l'exemple de Katz et Proctor). Le recours à

une modélisation par chaînes de Markov continues permet de surmonter ce problème. Il est suggéré pour la première fois par Holland et Leinhardt en 1977 [21]. Il faudra attendre les années 2000 pour que soient mis au point les modèle p^* par Wasserman et Pattison et SIENA par Tom A. B. Sniders mobilisant des chaînes de Markov continues. Ils permettent d'abandonner l'hypothèse d'indépendance des dyades (dénoncée par David Krakhardt en 1988 [26]) et ainsi d'estimer les effets triadiques comme la transitivité sur la structure du réseau (*triadic closure, clustering*). Entre autres avantages (décrits dans [14]), les modèles de Markov offrent la possibilité de tenir compte d'un nombre supérieur d'attributs si bien que SIENA est un modèle prisé par les membres de l'école de la Proximité. Ces derniers sont des géo-économistes intéressés par les effets d'homophilie (le fait de privilégier l'établissement de relations sociales ou le partage d'information avec des individus partageant des attributs semblables aux siens) par exemple, pour expliquer les effets de synergie observés dans les systèmes locaux d'innovation [9]. La considération des attributs est un moyen d'estimer les effets de décor mais, à ma connaissance, les attributs sont fixés ce qui empêche de tenir compte des changements de décor. Jusqu'au milieu des années 2000, ces modèles sont limités par le fait qu'ils ne peuvent gérer que des relations binaires (absence ou présence de la relation) et ne sont pas en mesure de traiter des liens valués. Cette limite est à présent surmontée dans les modèles SIENA (depuis 2008) et ERGM. Une autre limite, pointée dans sa thèse par Tore Opsahl en 2009, vient de la complexité d'estimation des paramètres qui s'explique par le fait que la variable dépendante (à expliquer) est le réseau entier [32]. En effet, ces modèles sont conçus pour des graphes statiques (un seul pour un modèle ERGM ; plusieurs pour un modèle SIENA). Ils estiment l'ordre dans lequel les liens s'établissent sans disposer de l'information. Tore Opsthal montre que dans le cas de graphes enregistrés en temps réels, pour lesquels chaque transition (création/disparition/augmentation/diminution) d'un lien est datée, il est préférable d'avoir recours à un modèle fondé sur une régression logistique conditionnelle [32]. Il s'agit de modéliser le choix fait par un nœud à un moment donné de se lier plutôt à un nœud qu'à un autre compte tenu des caractéristiques des nœuds (de leur position dans le réseau et de leurs attributs). Les modèles qui viennent d'être exposés sont conçus pour répondre à des interrogations sur les structures sociales. L'objectif est d'évaluer le rôle joué par différents principes d'attachement : réciprocité (pour les réseaux orientés) ; popularité ; transitivité et homophilie. Il semble néanmoins y avoir un intérêt à isoler ces mécanismes dans d'autres types de réseaux comme cela a pu se voir depuis le succès des travaux de Albert Barabási sur l'attachement préférentiel [4].

1.2 Les études de propagation : diffusion d'innovation et épidémiologie

A l'origine éloignés des préoccupations de l'analyse des réseaux sociaux, les chercheurs spécialisés dans la compréhension des processus de diffusion et contagion tendent de plus en plus à recourir à la théorie des graphes pour leurs modélisations. Les travaux fondateurs se concentraient sur des modélisations tenant compte d'un nombre limité de paramètres : la probabilité d'infection (épidémie) ou d'adoption (innovation) se déduisant de la part de la population infectée ou convertie selon le principe dit de *mass action* ou *homogeneous mixing* établi au début du XX^e siècle par des chercheurs en épidémiologie [24]. A partir de ce principe, les épidémies suivent des courbes en cloche (en raison des guérisons, le nombre d'infectés débute à zéro et retourne à zéro à la fin de l'épidémie) tandis que les diffusions d'innovation suivent des courbes en S (un plafond est atteint à partir d'un certain seuil d'adoption)⁴ [3][36]. Ainsi, les premiers modèles élaborés reposent sur l'hypothèse d'une population homogène en matière de susceptibilité et d'un espace de diffusion isotrope [35].

A partir des années 1970, la théorie des graphes apparaît pertinente pour tenir compte de l'hétérogénéité de la population. Ronald S. Burt montre que la position structurale des acteurs dans un réseau social est déterminante pour estimer leur vulnérabilité ou leur accès à l'information [11]. Dès 1962, Torsten Hägerstrand introduit dans ses modèles de Markov, la notion d'« effet de barrière »⁵ qui découle de l'inégale réceptivité des individus au contact d'une innovation et de l'inégale distribution de l'information à propos de l'innovation [20]. Il développe la notion d'« information privée » qui s'apparente à la notion de « connaissance tacite » de Michael Polanyi, qui transite par le « face à face ». Il observe que l'influence du voisinage joue un rôle crucial sur le choix d'un individu d'adopter une innovation : « En moyenne, la densité des contacts qui participent à l'information privée à la portée d'un seul individu doit décroître très rapidement à mesure qu'augmente la distance géographique »⁶ [20]. Pour Hägerstrand, la distribution de l'information privée peut être saisie par les concepts sociométriques développés par Moreno. Il anticipe sur les travaux ultérieurs mobilisant l'analyse des réseaux sociaux pour modéliser les processus de diffusion

4. Les articles fondateurs sont : pour la diffusion de l'innovation, « The diffusion of hybrid seed corn in two Iowa communities » paru en 1943 dans la revue *Rural Sociology*; pour la propagation d'épidémie, les travaux de T.J. Bailey et M.S. Barlett sur les modèles stochastiques d'épidémies.

5. L'effet de barrière n'est autre qu'un effet du décor sur le trafic.

6. p. 235. Il met ainsi le doigt sur une préoccupation devenue majeure au sein de l'école de la Proximité : estimer le poids de la proximité géographique relativement à celui d'autres types de proximité (institutionnelle, cognitive...) dans le partage d'informations et, par extension, dans la diffusion des idées ou innovations et la constitution de communautés épistémiques ou groupes d'intérêt [39].

d'innovation. Parallèlement, les différents types d'hétérogénéité (population cible, répartition des individus dans l'espace et dans le temps, différentiel de susceptibilités, mode de transmission, période d'incubation...) sont pointés par les épidémiologistes qui tâchent de les intégrer dans leurs modélisations stochastiques. Ce n'est qu'à partir du milieu des années 1980 que l'épidémiologie importe, elle aussi, les méthodes de l'analyse des réseaux sociaux pour étudier la propagation du VIH et des MST en général [31].⁷

Contrairement aux modèles épidémiologiques fondés sur des graphes, les nœuds dans les modèles de diffusion de l'innovation sont localement dépendants, autrement dit, l'effet qu'un nœud infecté aura sur son voisin dépend de l'état des autres voisins de ce dernier. En épidémiologie, les événements de contagion entre paires d'individus sont indépendants les uns des autres (indépendance des dyades) [43]. L'hétérogénéité dans un réseau social modélisé par un graphe se manifeste par la différence de position des nœuds dans le graphe et dans la répartition de leurs attributs. Dans un article sur le marketing viral, Hossam Sharara ajoute une touche de complexité en tenant compte de l'évolution dans le temps de la confiance accordée à autrui pour l'adoption d'un produit (ne pas se contenter de regarder le nombre de voisin ayant adopté le produit mais regarder plutôt le nombre de voisins de confiance l'ayant adopté) [37]. Depuis les années 2000, on parle de cascade pour qualifier le phénomène de diffusion ou de contagion qui traverse un graphe. Une cascade est difficilement prévisible car le même événement répété à plusieurs reprises peut ne pas systématiquement donner lieu à une cascade : de nombreuses conditions doivent être réunies pour que se déclenche une cascade. Duncan J. Watts montre en 2001 qu'il est plus facile d'empêcher une cascade dans un réseau peu dense que dans un réseau très dense : en effet, il suffit dans le réseau peu dense de supprimer les nœuds de très haut degré pour limiter les dégâts [43].

Pendant longtemps le problème en matière de modélisation venait des données accessibles. En effet, en l'absence de données précises sur le trafic, les modélisations s'appuient sur un état statique de la réalité. Prévoir la situation à $t+1$ étant donné la situation au temps t , sans tenir compte des multiples évolutions dans la population et ses contacts entre t et $t+1$. A partir des années 2000 on accède à des données précises sur les temps de contact et la mobilité en temps réel des individus (*Delay tolerant networks*). Des modélisations s'appuyant entre autres sur le principe de la marche aléatoire se développent [12]. Les modélisations tiennent de mieux en mieux compte du trafic [33][41].

7. Dans le même temps, le premier modèle spatial est proposé par Rvachev et Longini en 1985 sur la propagation de l'influenza : ce modèle tient compte des réseaux de transport pour expliquer la propagation [24].

2 Clusters, partitions et communautés en dynamique

Pour les études de propagation et de contagion, il peut être important de détecter les structures cohésives qui se maintiennent dans le temps (par exemple pour une mise en quarantaine). Les travaux portant sur la détection de communautés en dynamique sont extrêmement récents et se retrouvent sous l'appellation *evolutionary clustering* [13][38]. En général, résoudre ce problème revient à comparer les communautés détectées au temps t avec les communautés détectées au temps $t+1$. Les méthodes de partitionnement mobilisées le plus souvent sont donc des méthodes de partitionnement conçues pour des graphes statiques. Ces approches longitudinales ne sont pas dénuées d'intérêt et restent adaptées à la plupart des situations. Cependant, pour des instantanés rapprochés, les méthodes de détection de communautés classiques, sensibles au bruit, détectent des communautés aux structures très différentes d'un instantané à l'autre en raison du trafic alors même que la structure d'ensemble (le décor) ne se modifie pas fondamentalement. Pour éviter cela, il peut être tout simplement préférable d'agréger plusieurs instantanés et de rechercher les communautés sur cet agrégat [22]. D'autres méthodes ont été proposées pour repérer des communautés stables dans le temps en tenant compte de tous les instantanés. En 2011, Thomas Aynaud et Jean-Loup Guillaume proposent une adaptation de la méthode de Louvain aux graphes longitudinaux : plutôt que d'optimiser le critère de modularité à chaque instantané, il s'agit d'optimiser le critère de modularité globale et ainsi d'extraire les partitions de meilleure modularité globale, à savoir bonnes sur toute la période considérée [2]. La même année, A. Lancichinetti *et alii* développent une méthode dénommée *Order Statistics Local Optimization Method* (OSLOM). Au lieu d'analyser séparément les instantanés, cette méthode regroupe les informations des différentes périodes pour les traiter ensemble. L'algorithme s'appuie sur l'optimisation locale d'une fonction *fitness* exprimant la significativité des clusters en tenant compte du bruit [27]. En 2012, Romain Guigourès *et alii* mettent au point une méthode de *triclustering* pour la détection de structures temporelles dans les graphes orientés : « les nœuds sources, les nœuds cibles, ainsi que la variable temporelle sont simultanément segmentés afin de construire des intervalles de temps et des *clusters* de nœuds entre lesquels la distribution des arcs est stationnaire par intervalle de temps » [19]. Tenant compte simultanément de la structure du réseau et de son évolution, cette méthode de partitionnement est assez exigeante.

La même année, Bivas Mitra *et alii* proposent une méthode adaptée aux graphes dynamiques de type graphe de citations : les nœuds aux extrémités d'un même lien peuvent appartenir à des périodes différentes [30]. Depuis l'article de Norman P. Hummon et Patrick Doreian en 1989 qui étudie le

développement de la théorie de l'ADN à partir de l'analyse d'un réseau de citation, une littérature spécifique s'est développée sur ce type de réseaux qui s'inscrit dans un champ plus large : l'analyse évolutionniste des trajectoires [23]. L'outil « *Main Path Algorithm* » implémenté dans le logiciel Pajek a été conçu pour détecter le chemin principal dans un réseau de citation ou graphe dynamique. Utile pour détecter une trajectoire spécifique, il est mobilisé en 2007 par Andrea Mina *et alii* [29].

Les méthodes de partitionnement apparaissent utiles pour déduire la structure d'ensemble, l'armature du réseau ou décor. Une fois les partitions trouvées, il est possible de réaliser un graphe de partition ou « graphe image » et d'en étudier l'évolution dans le temps. Narciso Pizaro incite les chercheurs en sociologie à construire les graphes images qu'il appelle « réseaux de places » à partir de graphe bipartis autrement dit, en tenant compte des relations attributaires [34]. Dans ce cas, il n'est plus question de déduire le décor à partir de la structure du trafic (comme dans les méthodes exposées plus haut) mais plutôt, d'analyser le trafic compte tenu du décor (ce dernier est déduit à partir des relations d'appartenance ou attributs des entités).

Supposées plus stables dans le temps, les graphes images peuvent tout de même se modifier sous l'effet de petites fluctuations qui correspondent aux « bifurcations » au sens d'Andrew Abbott.

3 Les approches narratives

3.1 *Time matters* : une alternative à l'approche standard

En soi, l'analyse des réseaux ou « méthode structurale » est déjà une alternative aux méthodes standards dans la mesure où elle met l'accent « sur la détermination réciproque et l'interdépendance entre cas. » [1]. Les modélisations stochastiques évoquées plus haut sont aussi un moyen de relâcher les hypothèses selon lesquelles « l'ordre n'a que peu d'impact sur les dynamiques sociales ». Cependant, ces modèles sont sensiblement limités dans la mesure où ils ne cherchent à déduire le futur qu'en fonction de la seule situation présente (principe de Markov), sans compter qu'ils suggèrent une relation de causalité entre les deux situations : la situation passée expliquant linéairement la situation présente. Pour Andrew Abbott, « il n'y a aucune raison de penser qu'un indicateur donné a un quelconque statut particulier, expliqué ou explicatif » puisque « la vie sociale est une immense toile dans laquelle tout se tient ». Les méthodes standards parviennent difficilement à considérer simultanément toutes les dépendances. L'abandon de l'hypothèse d'indépendance des dyades dans les modèles stochastiques est en ce sens une avancée considérable mais ces modèles restent fondés sur l'hypothèse forte de l'individualisme méthodologique. Aussi, dans ces modèles, « la trajectoire causale serait toujours la même » or « la signification des causalités dépend du contexte » et « c'est la succession des points, envisagés dans toutes les di-

mensions à la fois, qui importe ». L'approche narrative proposée par Andrew Abbott, est descriptive et non causale : « Dans ce cadre, étudier une succession d'étapes n'implique pas de poser des questions de causalité, mais plutôt de descriptibilité ». Pour Abbott, « la dynamique sociale peut être analysée en terme de trajectoires et de bifurcations ». Les bifurcations interviennent quand « les cas passent d'une trajectoire à une autre, donc entrent dans un nouveau régime de causalité » et en ce sens, « les bifurcations permettent à la structure de jouer un rôle » [1]. L'analyse de séquences, développée notamment par Abbott, permet de comparer les trajectoires les unes par rapport aux autres en calculant leur distance et de repérer des régularités⁸.

3.2 Chaînes relationnelles et réseaux personnels

La compréhension des bifurcations se prête particulièrement aux méthodes mixtes à l'articulation du quantitatif et du qualitatif. Inspiré par le problème posé par Marc Granovetter en 1985 sur la question du rôle des relations interpersonnelles dans les activités économiques, Michel Grossetti a développé une méthode mixte dite de « narrations quantifiées » [18]. En focalisant sur des processus et non sur des personnes en particulier, cette méthode est résolument axée sur les dynamiques. Il s'agit d'analyser les séquences d'accès aux ressources dans un milieu professionnel donné. On distingue l'accès aux ressources par le biais de son réseau social de l'accès aux ressources par l'intermédiaire de dispositifs de coordination dédiés. Michel Grossetti analyse les trajectoires, leur degré de dépendance à des collectifs en mobilisant les notions d'« encastrement » et de « découplage » développées par Harrison Whyte. La méthode consiste à réaliser un traitement quantifié à partir d'entretiens semi-directifs, autrement dit, à repérer des régularités. Cette méthode n'est pas une méthode d'analyse de réseaux personnels. Le réseau social est vu, uniquement, comme le moyen d'accéder à une ressource à un moment donné. Il n'est pas question de l'analyser puisqu'il (le réseau) est envisagé comme un instrument dans l'explication du processus, lequel est réellement questionné.

S'il on retourne le problème, il est possible d'étudier les effets engendrés par une bifurcation sur la structure d'un réseau personnel. Claire Bidart a adopté cette approche quand elle a réalisé une analyse longitudinale de réseaux personnels [8]. En réalisant tous les trois ans, des entretiens auprès d'un panel de 87 jeunes en transition vers l'âge adulte, Claire Bidart étudie l'impact des changements intervenant dans leur vie entre chaque entretien sur leur réseau personnel. Un réseau personnel est un réseau obtenu auprès d'un « ego » amené à qualifier les relations qu'il entretient avec ses « alters » et les relations que ses « alters » entretiennent entre eux. Ce type d'approche permet une plus grande finesse dans l'analyse des dynamiques relationnelles

8. Le package `TramineR` pour le logiciel R a été développé à cette intention.

qu'une approche purement quantitative qui présume que l'évolution des relations se limite à des mécanismes structurels comme la transitivité. Les effets de contexte jouent un rôle crucial sur l'évolution de ces réseaux, en particulier, le passage d'un milieu homogène comme le lycée vers un milieu plus hiérarchisé, celui du travail. Pour saisir les logiques qui sont en jeu dans l'établissement ou la dilution de relations, l'espacement des entretiens ne constitue pas un frein particulier dans la mesure où l'évolution d'un réseau personnel est un processus relativement lent. Les relations étudiées ne sont pas des relations qui changent de nature d'une minute à l'autre puisqu'elles tiennent à la perception de l'individu enquêté et donc davantage au décor dans la vie d'un individu qu'au trafic qui anime la vie de cet individu.

3.3 Analyse de réseau historique

La sociologie historique utilise les méthodes d'analyse de réseaux pour les arbres généalogiques, les graphes de lignages et d'évènements. Les graphes mobilisés sont alors des graphes dynamiques au sens de Bivas Mitra *et alii* : les liens traversent le temps s'ils relient des évènements ou des cas qui se succèdent (les arcs sont orientés dans le sens du temps) [30]. La visualisation de tels graphes requiert un ordre de lecture (le plus souvent, de gauche à droite pour un réseau d'évènements, et de haut en bas pour un arbre généalogique). Ce type d'approche exige une rigueur particulière puisqu'il s'agit de reconstituer un récit, proposer une narration à l'aide d'un graphe. Des questions spécifiques y sont associées. Notamment, la question des frontières du graphe se pose en des termes différents : par où commencer ? Où s'arrêter ? Doit-on intégrer ce cas dans l'analyse ? Par exemple, l'historien qui veut fabriquer un graphe d'évènement à partir de récits personnels se retrouve rapidement encombré d'évènements plus ou moins contingents (propres aux personnes interrogées) compte tenu du phénomène plus global qu'il étudie. Les analyses qui touchent à cette catégorie de graphes sont encore assez rares mais l'intérêt croissant des historiens pour la théorie des graphes suggère qu'elles le seront bientôt de moins en moins [28].

Peter Bearman *et alii* proposent une méthode de détection des « casseurs de cas » (*case breakers*) autrement dit des évènements (nœuds) indispensable au déroulement du récit [5]. La célèbre citation de Blaise Pascal « Le nez de Cléopâtre s'il eut été plus court, la face de la terre en aurait été changée » illustre bien les enjeux d'une telle approche. Elle s'éloigne du programme narratif, tel que défini par Andrew Abbott, puisque la recherche de « casseurs de cas » repose sur l'idée d'une causalité linéaire entre les cas. En revanche, l'ordre des évènements (ou cas) est ici essentiel. Pour Andrew Abbott, chaque cas dispose d'un horizon temporel qui lui est propre. Dès lors, il peut être réducteur de réduire chaque cas à un nœud dans un graphe puisque cela revient à traiter les cas sur un pied d'égalité. On pourrait néanmoins imaginer de faire varier l'importance des cas en leur affectant, par

exemple, un poids relatif à leur horizon temporel. . .

Balazs Vedres et David Stark analysent l'évolution des liens entre 1696 entreprises hongroises de 1987 à 2001 en cumulant une analyse de réseaux sociaux statique à une analyse longitudinale de réseau historique [40]. Les liens personnels entre entreprises sont analysés pour détecter des groupes cohésifs ou partitions. La cohésion et l'évolution des groupes est ensuite analysée à partir d'un graphe dynamique de lignage. Les liens traduisent la continuité : si un ou plusieurs membres restent dans le groupe d'une année à l'autre, il existe un lien dirigé entre le groupe au temps t et le même groupe au temps $t+1$. Ce travail s'inscrit dans une vieille tradition en sociologie qui consiste à interroger la permanence des structures dans le temps.

Conclusion

Plusieurs des défis posés par les analyses dynamiques de réseaux se sont révélés traduisibles dans le vocabulaire de Ron Atkin. En effet, la difficulté majeure provient de ce que la dynamique d'une structure résulte d'allers-retours continus entre le décor et le trafic, la structure et l'action, l'espace et le processus. La compréhension de ces allers-retours est fondamentale pour une grande variété de questions de recherche n'ayant pas, en apparence, de rapport. En plein dynamisme en particulier grâce à l'apport de jeux de données inédits, l'analyse de l'évolution des réseaux est déjà richement dotée en matière conceptuelle. La variété des travaux portant sur l'évolution des réseaux est notable. Deux cas de figures se sont présentés : réseaux de relations et réseaux de contacts. Soit l'analyste ne dispose que d'instantanés séparés dans le temps (l'état du réseau à un moment donné). Soit, l'analyste dispose d'une série d'instantanés rapprochés voire a connaissance de la durée des contacts (de l'instant auquel les contacts s'établissent, changent et se résorbent). Selon que l'on soit dans un cas de figure ou l'autre, les méthodes à privilégier ne sont pas les mêmes à la fois pour la modélisation et le partitionnement. Se voulant aussi exhaustive que possible, cette synthèse n'en laisse pas moins de côté un grand nombre de travaux. On signale ceux portant sur la visualisation de la dynamique des réseaux et ceux, très récents, qui portent sur le suivi d'entités spatiales dans le temps ou entités spatio-temporelles à l'aide de la théorie des graphes.

Références

- [1] Andrew ABBOTT : Time matters Traduction de l'épilogue par Claire Lemerancier et Carine Ollivier. *Terrains Travaux*, (19) : 183-203, 2011.
- [2] Thomas AYNAUD et Jean-Loup GUILLAUME : Détection de communautés à long terme dans les graphes dynamiques. *Journée thématique : Fouille de grands graphes*, 2010.

- [3] Norman T.J. BAILEY : The total size of a general stochastic epidemic. *Biometrika*, 177–185, 1953.
- [4] Hawoong JEONG, Zoltan NEDA et Albert-László BARABASI : Measuring preferential attachment in evolving networks. *EPL (Europhysics Letters)*, 61(4) : 567, 2007.
- [5] Peter BEARMAN, Robert FARIS, et James MOODY : Blocking the Future New Solutions for Old Problems in Historical Social Science. *Social Science History*, 23(4) : 501-533, 1999.
- [6] Laurent BEAUGUITTE : Analyse de graphe et modèles statistiques Du modèle p1 à l'ERGM. 2012. *Groupe fmr*, <http://halshs.archives-ouvertes.fr/FMR/>
- [7] Laurent BEAUGUITTE : L'analyse des graphes bipartis. 2013. *Groupe fmr*, <http://halshs.archives-ouvertes.fr/FMR/>
- [8] Claire BIDART : Dynamiques des réseaux personnels et processus de socialisation : évolutions et influences des entourages lors des transitions vers la vie adulte. *Revue française de sociologie*, 49(3) : 559–583, 2008.
- [9] Ron BOSCHMA, Pierre-Alexandre BALLAND et Dieter KOGLER : A Relational approach to knowledge spillovers in Biotech network structures as drivers of inter-organizational citation patterns. *Papers in Evolutionary Economic Geography (PEEG)*, 2011.
- [10] Ronald L. BREIGER : The Duality of Persons and Groups. *Social Forces*, 53(2) : 181-190, 1974.
- [11] Ronald S. BURT : Innovation as a structural interest : Rethinking the impact of network position on innovation adoption. *Social Networks*, 2(4) : 327–355, 1980.
- [12] Tracy CAMP, Jeff BOLENG et Vanessa DAVIES : A survey of mobility models for ad hoc network research. *Wireless communications and mobile computing*, 2(5) : 483–502, 2002.
- [13] Deepayan CHAKRABARTI, Ravi KUMAR et Andrew TOMKINS : Evolutionary clustering. *KDD Conference*, 2006.
- [14] Ainhoa DE FREDERICO DE LA RUA : L'analyse longitudinale de réseaux sociaux totaux avec SIENA – Méthode, discussion et application. *Bulletin de Méthodologie Sociologique*, 84 : 1-27, 2004.
- [15] Patrick DOREIAN : On the Evolution of Group and Network Structure. *Social Networks*, (2) : 235-252, 1979.
- [16] Linton C. FREEMAN : Q-analysis and the structure of friendship networks. *International Journal of Man-Machine Studies*, 12(4) : 367–378, 1980.
- [17] Peter GOULD : Q-analysis, or a language of structure : an introduction for social scientists, geographers and planners. *International Journal of Man-Machine Studies*, 13 : 169-199, 1980.

- [18] Michel GROSSETTI : Une méthode mixte pour étudier des processus sociaux. *Terrains Travaux*, 19 (2) : 161-182, 2011.
- [19] Romain GUIGOURES, Marc BOULLE et Fabrice ROSSI : Triclustering pour la détection de structures temporelles dans les graphes. *MARAMI Villetaneuse*, 2012.
- [20] Tornsten HAGERSTRAND : *Innovation Diffusion as a Spatial Process*, Postscript and Translation by Allan Pred : Transl. with the Assistance of Greta Haag : (From Innovationsförloppet Ur Korologisk Synpunkt). The University of Chicago Press, 1967.
- [21] Paul W. HOLLAND et Samuel LEINHARDT : A dynamic model for social networks. *The Journal of Mathematical Sociology*, 5(1) : 5-20, 1977.
- [22] John HOPCROFT, Omar KHAN, Brian KULIS et Bart SELMAN : Tracking evolving communities in large linked networks. *Proceedings of the national academy of sciences of the United States of America*, 101(Suppl 1) : 5249–5253, 2004.
- [23] Norman P. HUMMON et Patrick DOREIAN : Connectivity in a citation network : The development of DNA theory. *Social Networks*, 11(1) : 39–63, 1989.
- [24] Isham, Valerie ISHAM : *Stochastic models for epidemics*. 2004. <http://www.ucl.ac.uk/statistics/research/pdfs/rr263.pdf>.
- [25] Leo KATZ et Charles H. PROCTOR : The concept of configuration of interpersonal relations in a group as a time-dependent stochastic process. *Psychometrika*, 24(4) : 317–327, 1959.
- [26] David KRACKHARDT : Predicting with networks : Nonparametric multiple regression analysis of dyadic data . *Social Networks*, 10(4) : 359-381, 1988.
- [27] Andrea LANCICHINETTI, Filippo RADICCHI, José J. RAMASCO et Santo FORTUNATO : Finding statistically significant communities in networks. *PLoS ONE*, 6(4), 2011.
- [28] Claire LEMERCIER, Sandro GUZZI-HEEB et Michel BERTRAND : Introduction : où en est l'analyse de réseaux en histoire? *Redes : revista hispana para el análisis de redes sociales*, 21 : 12–23, 2011.
- [29] Andrea MINA, R. RAMLOGAN, G. TAMPUBOLON et J.S. METCALFE : Mapping evolutionary trajectories : Applications to the growth and transformation of medical knowledge. *Research Policy* 36 : 789-806, 2007.
- [30] Bivas MITRA, Lionel TABOURIER, et Camille ROTH : Intrinsically dynamic network communities. *Computer Networks*, 56(3) : 1041-1053, 2012.

- [31] Denis MOLLISON : *Epidemic Models : Their Structure and Relation to Data*. Cambridge University Press, 1995.
- [32] Tore OPSAHL : *Structure and Evolution of Weighted Networks*. Thèse de doctorat, University of London, 2009.
- [33] André PANISSON, Alain BARRAT, Ciro CATUTTO, Wouter VAN DER BROECK, Ruffo GIANCARLO et Rossano Schifanella : On the dynamics of human proximity for data diffusion in ad-hoc networks. *Ad Hoc Networks*, 10(8) : 1532-1543, 2012.
- [34] Narciso PIZARRO : Appartenances, places et réseaux de places. *Sociologie et sociétés*, 31(1) : 143-161, 1999.
- [35] Dominique RAYNAUD : Why Do Diffusion Data not Fit the Logistic Law? A Note on Network Discreteness, Heterogeneity Anisotropy. *From Sociology to Computing in Social Networks : Theory, Foundations and Applications*, 81-96, 2010.
- [36] Everett. M. ROGERS : *Diffusion of innovations*. Free Press of Glencoe, 1962.
- [37] Hossam SHARARA, William RAND et Lise GETOOR : Differential adaptive diffusion : Understanding diversity and learning whom to trust in viral marketing. *Proceedings of the AAAI International Conference on Weblogs and Social Media (ICWSM)*, 2011.
- [38] Mansoureh TAKAFFOLI, Farzad SANGI, Justin FAGNAN et Osmar R. ZAIANE : Community Evolution Mining in Dynamic Social Networks. *Procedia - Social and Behavioral Sciences*, 22 : 49-58, 2011.
- [39] Anne TER WAL et Ron BOSCHMA : Applying social network analysis in economic geography : framing some key analytic issues. *Ann Reg Sci*, (43) : 739-756, 2009.
- [40] Balazs VEDRES et David STARK : Structural Folds : Generative Disruption in Overlapping Groups. *American Journal of Sociology* 115(4) : 1150-1190, 2010.
- [41] Matthew C. VERNON et Matt J. Keeling : Representing the UK's cattle herd as static and dynamic networks. *Proceedings of the Royal Society B : Biological Sciences*, 276(1656) : 469-476, 2009.
- [42] Stanley WASSERMAN et Dawn IACOBUCCI : Sequential social network data. *Psychometrika*, 53(2) : 261-282, 1988.
- [43] Duncan J. WATTS : A simple model of global cascades on random networks . *Proceedings of the National Academy of Sciences*, 99(9) : 5766-5771, 2002.