

Intégration de la Dimension Temporelle dans l'Analyse des Réseaux de Santé en Français *

Arnaud Casteigts^{†‡} et Louise Bouchard[‡]

[†]École d'Ingénierie et de Technologie de l'Information, Université d'Ottawa

[‡]Institut de Recherche en Santé des Populations, Université d'Ottawa

{acasteig, louise.bouchard}@uottawa.ca

31 août 2010

Résumé

Le présent document propose quelques éléments sur l'analyse des réseaux humains par le biais de graphes dynamiques. Il se destine plus particulièrement à l'évaluation de la capacité de communication des réseaux de santé francophones qui ont été implantés ces dernières années dans les provinces canadiennes, suivant le modèle *vers l'unité pour la santé* [1] recommandé par l'OMS. Notre objectif dans ce premier rapport est de discuter les éléments *quantitatifs* de la future analyse, afin d'anticiper les questions que devrait comporter notre questionnaire d'enquête visant à cartographier ces réseaux.

1 Introduction

En 2001, le Comité Consultatif des Communautés Francophones en Situation Minoritaire (CCCFSM) a présenté au ministre fédéral canadien de la santé un rapport [2] faisant état de l'accès aux soins en langue maternelle par le million de francophones vivant en situation minoritaire au Canada. Il a été constaté qu'entre 50 % et 55 % de ces francophones n'avaient pas ou peu accès à des services de santé en français. Par ailleurs, d'importantes disparités régionales et locales étaient présentes et ne dépendaient pas toujours de la densité des francophones, soulignant ainsi l'insuffisance d'une approche uniquement nationale pour résoudre le problème.

Parmi les mesures proposées pour améliorer l'état de santé individuel et collectif des francophones se trouvait la recommandation de « mettre en œuvre une initiative de réseautage communautaire pour permettre l'organisation de réseaux provinciaux et territoriaux et d'un secrétariat national, lesquelles entités constitueraient la base du développement de l'accès aux services de santé en français [et viseraient] à créer des liens entre les communautés francophones, les professionnels de la santé, les établissements d'enseignement, les établissements de santé et les associations professionnelles de langue française. » D'autres recommandations portaient sur les thèmes de l'infrastructure, la main-d'oeuvre, les nouvelles technologies, et la recherche ; le thème du réseautage a cependant été présenté comme prioritaire car pouvant servir lui-même de catalyseur à tous les autres thèmes.

En réaction à ces recommandations, le ministère fédéral de la Santé – Santé Canada – a créé un fond national dédié au développement de ces réseaux, et en a confié la gestion à l'organisme *Société Santé en français* (avec une enveloppe initiale de près de 10 M\$). Le modèle de réseau qui a été choisi est celui mis

* Rapport préliminaire pour le compte du Réseau de Recherche Appliquée sur la Santé des Francophones en Ontario (RRASFO).

en avant par l'Organisation Mondiale de la Santé, dans le cadre de sa stratégie *vers l'unité pour la santé* [1]. Il s'agit d'un modèle de réseau en forme de graphe pentagonal complet visant à faciliter la coordination et l'intégration du vaste éventail d'acteurs (décideurs politiques, gestionnaires de services, professionnels de la santé, organismes de formation, et communautés) au niveau d'une population donnée. Cette approche, où tous les acteurs sont censés interagir (voir Figure 1), est censée réduire la fragmentation due à la nature différente des acteurs, et contribuer à se rapprocher à long terme de l'idéal d'une santé de qualité pour tous à moindre prix.

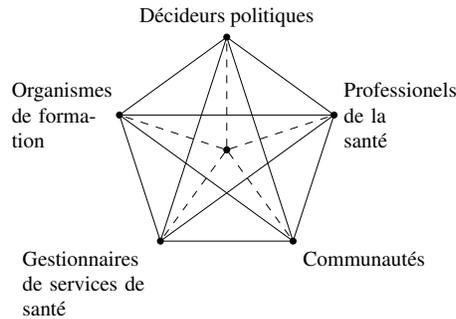


FIGURE 1 – Organisation d'un réseau à l'échelle provinciale/territoriale.

Il a été décidé d'instaurer au minimum un réseau de ce type par province ou territoire canadiens, et de nommer dans chaque réseau une personne à plein temps chargée à la fois de promouvoir les interactions locales et d'assurer la fonction de relai sur le plan national (*représentée ici par un nœud central dans le pentagone*). De nombreuses structures ayant déjà précédé le déploiement de ces réseaux à travers le Canada, une grande diversité existe aujourd'hui quant à leur nature, taille, et organisation. Nous considérons néanmoins dans ce document que chaque réseau peut effectivement être assimilé à un pentagone, et que l'allure des réseaux sur le plan national pourrait ressembler en essence à ce qui est montré Figure 2.

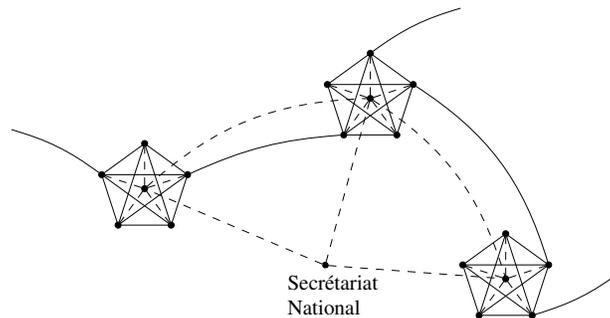


FIGURE 2 – Vue d'artiste du réseau à l'échelle nationale.

Les réseaux en général peuvent être évalués selon des critères extrêmement variés. On peut citer par exemple les questions portant sur le réseau lui-même : quelle est sa nature ? (*e.g.* réseau de connaissances, communauté de pratique, réseau virtuel) Quel est son mode de gouvernance ? A-t-il un fonctionnement transparent ? Une orientation stratégique ? Un mandat ? Si oui, ce mandat est-il réalisé ? Quel est le rendement du réseau sur le plan interne (bénéfices pour les membres) et externe (bénéfices pour la société) ? Le réseau favorise-t-il l'innovation ? *etc.* D'autres critères peuvent également être considérés à l'échelle des membres, *p.ex.* se sentent-ils impliqués ? Ont-ils un sentiment d'appartenance à une même communauté ? Quel est leur degré d'engagement dans le réseau ? Ont-ils des objectifs précis et partagés ? Des responsabilités clairement définies ? Leurs possibilités d'expression sont-elles équitables ? Leurs échanges réciproques ? *etc.* Chacun de

ces aspects permet à sa manière de comprendre le fonctionnement d'un réseau, et certains sont généralement considérés comme nécessaires à sa réussite (voir [3] et [4] pour une série de discussions sur les différents types de réseaux et leur évaluation). On peut finalement noter que l'analyse peut ne pas se limiter au réseau seul, en prenant en compte la spécificité du domaine ou de l'objet étudié, à l'instar de la théorie de l'acteur-réseau [5].

Le présent document n'a pas vocation à répondre à toutes ces questions, et se limite à un aspect transversal essentiel du réseau : la *communication*, ou plus précisément, la mesure quantitative des capacités de communication et d'interactions qu'offre le réseau sur une base régulière. Il est évident que ces mesures seules ne peuvent rendre compte de la complexité d'un réseau d'acteurs humains, et les éléments présentés ici doivent donc être vus comme un moyen de compléter une analyse qualitative plus vaste de ces réseaux.

Les études existantes sur la quantification des interactions d'un réseau se bornent généralement à étudier la *structure* du graphe le représentant (forme, diamètre, degré des nœuds, longueur des chemins, redondance, *etc.*). Ces informations seules reflètent cependant assez mal la réalité car elles n'intègrent pas la dimension *temporelle* des interactions (*e.g.* fréquence et durée). Il est peu probable, par exemple, que l'on échange plus d'information avec une personne que l'on croise une fois par mois dans un ascenseur qu'avec un collègue dont on partage quotidiennement la pause déjeuner. Ces deux relations ont pourtant le même statut lorsqu'on ne regarde que la structure d'un graphe d'interaction.

Nous proposons ici d'utiliser de nouveaux concepts issus de la recherche fondamentale en informatique sur les réseaux dynamiques, qui s'articulent en particulier autour de la notion de *distance temporelle* entre entités du réseau [6, 7]. L'application de ce type de concepts à l'analyse de réseaux sociaux a fait l'objet de travaux récents parmi lesquels [8] ou [9]. Le deuxième est presque déjà considéré comme incontournable en la matière, et c'est d'ailleurs de lui que sont dérivées la plupart des idées proposées dans ce rapport. En essence, l'apport principal du glissement vers la dimension temporelle est de révéler la forme « fonctionnelle » d'un réseau, qui peut s'avérer très différente de sa forme structurelle. De là peut être développée une nouvelle panoplie de mesures pour caractériser l'efficacité du réseau, son équilibre, la présence de points névralgiques, de sous-ensembles temporellement excentrés, *etc.*, que nous proposons ici d'appliquer à l'analyse des réseaux de santé en français au Canada.

2 Communication systémique

Il est raisonnable de se représenter un réseau comme ayant une structure assimilable à un graphe G , dont les sommets (ou nœuds) représentent les entités et les arêtes (ou liens) représentent les *possibilités* d'interactions entre ces entités (voir un exemple à la Figure 3(a)). Au dessus de cette structure prennent place des interactions, qui peuvent être vues comme des utilisations ponctuelles des arêtes de la structure. La dynamique de ces interactions joue un rôle crucial dans la manière dont circule l'information dans le réseau. Bien que certaines de ces interactions soient essentiellement irrégulières et imprévisibles (*p.ex.* se faisant en réaction à des événements ponctuels), une bonne partie d'entre elles est déterminée par les relations professionnelles ou amicales qu'entretiennent les membres du réseau, et offrent par là même une certaine stabilité. La caractérisation de cette communication *systémique* (pour reprendre le terme utilisé dans [9]) permet de comprendre comment l'information circule réellement dans le réseau, et donc d'étudier sa morphologie fonctionnelle.

Supposons que l'on arrive à identifier les fréquences moyennes d'interaction entre chaque paire de nœuds interagissant (on parlera de nœuds *voisins* dans le graphe), comme illustré à la Figure 3(b). Ce nouveau type de graphe, rajoutant des informations temporelles au dessus d'une structure G , sera généralement représenté par la lettre \mathcal{G} . A partir de ce nouveau graphe on peut calculer les délais moyens de propagation associés à un cheminement donné de l'information, comme suit : on détermine tout d'abord le délai moyen associé à chaque arête du graphe en prenant la *période* d'interaction correspondante divisée par deux. Par exemple,

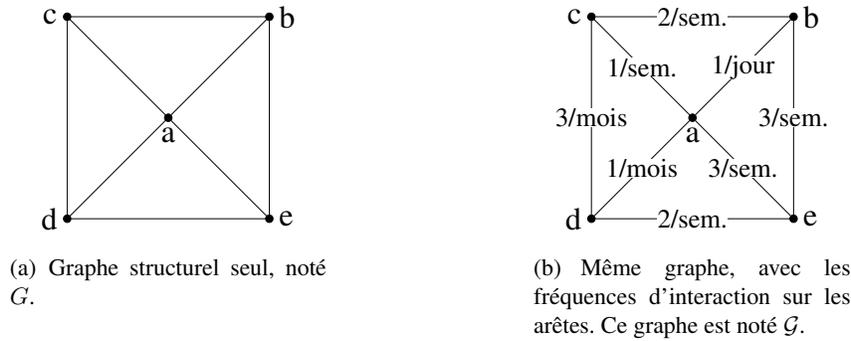


FIGURE 3 – Exemple d'un réseau quelconque, avec ou sans informations temporelles sur les arêtes.

l'arête (cd) étant activée en moyenne trois fois par mois, son délai d'attente moyen est de $(30/3)/2 = 5$ jours. Les délais correspondant aux fréquences du graphe de la Figure 3(b) sont reportés à la Figure 4(a).

Une fois les délais d'arêtes calculés, le délai d'un chemin constitué de plusieurs arêtes est simplement la somme de tous les délais de ces arêtes. En termes simples, cette durée correspond au temps moyen qu'il faut à une information pour se propager le long du chemin considéré. Il est intéressant de constater à ce stade que les chemins qui sont les plus courts sur le plan structurel ne sont pas forcément les plus courts sur le plan temporel. Par exemple, une information émise par a est susceptible de parvenir en d cinq fois plus rapidement de manière indirecte (via e) que directe.

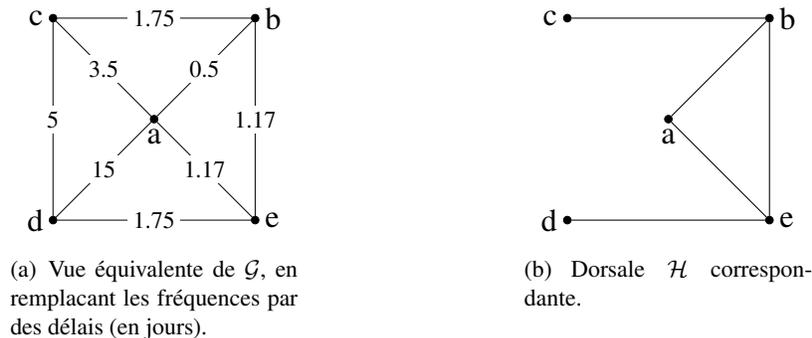


FIGURE 4 – Même exemple de réseau qu'à la Figure 3, mais présenté sous différents angles.

Sur la base de ce constat, les auteurs de [9] se sont intéressés à isoler l'ensemble des arêtes dites *essentielles*, c'est-à-dire qui ne peuvent pas être battues en vitesse par un autre chemin indirect. Ils ont baptisé le nouveau graphe déterminé par cet ensemble la *dorsale* du graphe d'origine (backbone, en anglais), que l'on notera généralement \mathcal{H} . Ce concept de dorsale s'avère très utile pour analyser les réseaux car il montre les chemins privilégiés qu'emprunte l'information dans le réseau, révélant ainsi des aspects que la structure seule ne pouvait pas montrer. Si l'on regarde le graphe structurel de la Figure 3(a), on pourrait conclure que le nœud a est le plus central, et que les nœuds b et c jouent des rôles d'importance égale dans le réseau (par exemple). Mais un coup d'oeil rapide à la dorsale du graphe Figure 4(b) nous informe que la réalité est différente, et que les nœuds b et e , bien que légèrement excentrés sur le plan structurel, sont finalement plus centraux que a sur le plan fonctionnel.

Les sections suivantes explorent ce que l'on pourrait analyser en utilisant les concepts ci-dessus dans le cadre des réseaux de santé en français. À noter que nous avons adopté ici une vision purement statistique des interactions, en considérant la structure du réseau comme stable sur le long terme, ce qui représente

une simplification importante. L'étude de l'évolution de cette structure au cours du temps est un sujet de recherche intéressant, mais il ne sera pas abordé dans ce document.

3 Analyse des réseaux

La majorité des études récentes sur les réseaux sociaux s'appuient sur un corpus de données d'où peuvent être extraits des traces d'interactions datées dans le réseau. Ces traces peuvent être par exemple des entêtes de courriels anonymisés, des historiques de connexions sur des sites collaboratifs, ou des messages laissés dans des forums communautaires en ligne. Ces données étant précisément datées, il est possible de reconstituer le déroulement des communications et d'obtenir des mesures très précises sur le réseau. La situation est plus compliquée dans le cas des réseaux de santé que nous souhaitons analyser, car nous ne disposons d'aucun historique d'interaction. Il est cependant envisageable de réaliser une enquête auprès des membres nous permettant de cartographier le réseau et obtenir des informations *statistiques* sur les interactions (leur fréquence, entre autres).

De nombreuses questions se posent alors quant aux informations que l'on souhaite récolter et utiliser. En premier lieu, quel type d'interaction veut-on analyser ? S'intéresse-t-on aux interactions par échange de courriel ? Par communication téléphonique ? Ou par rencontre physique des personnes ? Il est probable que les trois nous intéressent, mais la manière dont ils doivent être combinés n'est pas une question triviale, car ils diffèrent aussi bien en nature (*p.ex.* le courriel est unidirectionnel) qu'en intensité (les gens échangent bien plus lorsqu'ils se rencontrent physiquement). Même question pour le thème des échanges (*p.ex.* échanges de connaissances techniques, de moyens logistiques, ou interactions amicales). Souhaite-t-on étudier les différents thèmes séparément et regarder s'ils engendrent des motifs d'interaction différents ? Une autre question majeure est celle de la délimitation du réseau (aussi connue sous le nom « boundary detection ») : qui doit être considéré à l'intérieur et à l'extérieur du réseau ; jusqu'à quelle profondeur doit-on sonder ? Est-ce que des chemins passant à l'extérieur du réseau sont susceptibles d'affecter les distances temporelles entre des nœuds internes ? Nos mesures seraient-elles affectées s'il manquait des nœuds (*c.-à-d.* des personnes non-interrogées) ? Se pose finalement le problème de la granularité : souhaite-t-on étudier les interactions au niveau des personnes ou au niveau des pôles d'activités (sommets du pentagone) ? Les nœuds du graphe doivent-ils être des personnes ou des pôles (voire dans certains cas des provinces) ? Cette dernière question engendre également de nouveaux problèmes d'ordre mathématique ou technique, liés à la manière dont on peut définir les distances temporelles entre *groupes* de nœuds.

Les questions posées ci-dessus peuvent (et doivent) avoir un impact sur la conception du questionnaire que l'on veut soumettre à ces communautés. La Section 5 répertorie les informations minimales dont il faudrait disposer pour le type d'analyse discuté. Ceci étant dit, si l'on suppose avoir déjà un graphe d'interaction donné à disposition, voici une liste de questions auxquelles l'on pourrait tenter de répondre sur l'efficacité ou l'équilibre des réseaux de santé. Ces questions sont accompagnées d'indications techniques sur la démarche à suivre pour y répondre (se référer au glossaire de la Section 4 pour tout terme inconnu).

3.1 Efficacité

La mesure centrale d'efficacité que l'on considère ici est la *latence* de l'information, c'est-à-dire le temps qu'il faut à une information pour se propager au sein du réseau. Les questions ci-dessous devraient permettre de la caractériser.

Question 1 *Quelle est la vitesse moyenne de propagation dans le réseau ? (à l'échelle provinciale comme nationale)*

Réponse. Pourcentage moyen de nœuds atteignables en fonction du temps écoulé. Cela englobe deux résultats plus précis : le temps de propagation médian (temps écoulé pour 50 % de nœuds informés) et le diamètre

temporel du réseau (temps écoulé pour 100 % de nœuds informés). La moyenne est calculée sur l'ensemble des nœuds.

Question 2 *Comment ces vitesses se comportent-elles par rapport à une propagation uniforme aléatoire ? Est-ce plus performant ou moins performant ? (la réponse la plus probable est que c'est moins performant)*

Réponse. Comparer les vitesses obtenus avec \mathcal{G} aux vitesses obtenus avec un autre graphe \mathcal{G}' ayant la même structure que \mathcal{G} , la même quantité cumulée d'interactions pour chaque nœud, mais en répartissant ces interactions de manière uniforme entre tous ses voisins. Des ajustements devront être faits pour équilibrer la valeur de l'arête en fonction de ses deux extrémités (au moins dans le cas non-orienté), ces valeurs n'étant plus nécessairement égales.

Question 3 *Y'a-t-il des écarts notables entre les différents réseaux provinciaux/territoriaux ?*

Réponse. Écart type du temps de propagation médian et du diamètre temporel parmi les différents réseaux.

Question 4 *Quel est l'impact des arêtes de longue portée ? (cette question n'a probablement de sens qu'à l'échelle nationale). L'hypothèse que l'on souhaite vérifier (ou infirmer, mais c'est peu probable) est que les arêtes de longue portée raccourcissent de manière importante la latence de l'information dans le réseau. C'est en quelque sorte une mesure de l'efficacité de ce que Granovetter appelait les liens faibles.*

Réponse. Proportion d'arêtes de longue portée de \mathcal{G} qui sont conservées dans la dorsale \mathcal{H} , à comparer à la proportion d'arêtes de moyenne ou courte portée qui sont également conservées.

Question 5 *Quels types de liens sont conservés dans la dorsale ? La réponse probablement attendue est : une combinaison faite essentiellement de liens de longue portée et de liens fortement encastrés (résultat obtenu lors de précédentes études portant sur des échanges de courriels).*

Réponse. La question précédente répond en partie pour la portée. Quant à l'encastrement, on peut examiner le ratio d'arêtes conservées dans \mathcal{H} en fonction de leur valeur d'encastrement dans \mathcal{G} .

Question 6 *Existe-t-il de la redondance dans le réseau ? Ou autrement formulé, le réseau est-il robuste à la perte arbitraire de nœuds ? (e.g. un départ à la retraite ou la démission de l'un des membres)*

Réponse. Il est évident que de nombreux chemins alternatifs existent la plupart du temps au niveau structurel, mais sont-ils de performance comparable ? On aurait besoin ici de restreindre l'ensemble des arêtes considérées un peu à la façon de la dorsale, mais tout en gardant plus d'arêtes (la dorsale élimine par définition beaucoup de redondances). La solution pourrait être d'utiliser une version « graduelle » de la dorsale, dans laquelle on enlèverait seulement les arêtes pour lesquelles il existe un chemin indirect k fois plus rapide (p.ex. $k = 2$). La redondance peut ensuite être testée dans ce nouveau graphe de manière structurelle.

3.2 Équilibre

On s'intéresse ici à savoir si le réseau est équilibré, c'est-à-dire s'il n'y a pas d'écart notables entre les différentes entités à l'échelle locale (écarts entre personnes, pôles d'activités, ou couples de pôles d'activités), ou nationale (écarts entre provinces ou couples de provinces).

Question 7 *Les arêtes du pentagone sont-elles toutes utilisées ?(1) De manière régulière ?(2)*

Réponse. (1) Les arêtes du pentagone sont toutes présentes dans \mathcal{G} . (2) Leur fréquence est supérieure à un seuil donné.

Question 8 *Quel est le niveau de concentration des interactions ?*

Réponse. Proportion d'arêtes de \mathcal{G} qui sont présentes dans la dorsale \mathcal{H} (arêtes *essentielles*), c'est-à-dire $|E_{\mathcal{G}}|/|E_{\mathcal{H}}|$.

Question 9 *Y'a-t-il des pôles qui coopèrent plus fréquemment entre eux ?(1) Si oui, quels sont les pôles les plus centraux, ou au contraire, les plus excentrés ?(2)*

Réponse. (1) Distance temporelle moyenne entre chaque couple de pôles (au sein de leur pentagone). (2) Excentricité temporelle moyenne de chaque pôle dans son pentagone. On peut aussi regarder le degré des nœuds dans \mathcal{H} .

Question 10 *Y'a-t-il des provinces qui coopèrent plus fréquemment entre elles ?(1) Si oui, quelles sont les plus centrales, ou au contraire, les plus excentrées ?(2)*

Réponse. (1) Distance temporelle moyenne entre chaque couple de provinces. (2) Excentricité temporelle moyenne de chaque province dans le réseau national. On peut aussi regarder le degré des nœuds dans \mathcal{H} .

Question 11 *En considérant une granularité plus fine (au niveau des personnes), y'a-t-il dans le réseau des points névralgiques (e.g. personnes dynamique, franchiseurs de frontières) ? L'apport du directeur est-il flagrant de ce point de vue ? Existe-t-il des personnes particulièrement isolées ?*

Réponse. Même type de réponse que pour les deux questions précédentes, mais en considérant les personnes plutôt que les pôles. (Probablement à faire à l'échelle du pentagone.) Pour le franchiseur de frontière, on peut compter le nombre d'arêtes de longue portée dans \mathcal{G} qu'un nœud conserve dans \mathcal{H} .

Question 12 *Quel est le niveau de fragmentation du réseau ? (à l'échelle provinciale)*

Réponse. Regarder la proportion de nœuds critiques dans \mathcal{G} et \mathcal{H} . Si la structure pentagonale est respectée, il ne devrait pas y avoir de nœud critique dans \mathcal{G} . Il est pourtant probable que sur l'ensemble des provinces, cela arrive au moins une fois (*p.ex.* si le directeur est le seul en contact avec l'un des pôles). Il est en revanche assez probable qu'il y ait souvent des nœuds critiques dans \mathcal{H} , et cela pourrait être le reflet d'un rôle un peu trop central de certaines entités.

Question 13 *Est-ce que les nœuds à fort degré sont réellement les plus centraux ? L'idée ici est de montrer que l'importance des nœuds n'est pas strictement proportionnelle à leur degré dans le graphe structurel.*

Réponse. Degré moyen des nœuds dans \mathcal{H} en fonction de leur degré dans \mathcal{G} . On devrait obtenir une courbe sous-linéaire, qui permet de conclure que les réseaux sont en général plus équilibrés que ce que leur structure seule donne à penser.

Question 14 *Est-ce que le réseau bénéficierait d'une concentration de charge ? Ou au contraire, d'un nivellement ? Autrement dit, l'efficacité serait-elle améliorée si l'on accentue les différences de fréquences entre arêtes, ou si au contraire, on les réduit ?*

Réponse. En conservant un volume total d'interactions identique pour chaque nœud, répartir les interactions de manière à renforcer les liens les plus fréquents, ou les moins fréquents, puis mesurer l'efficacité du nouveau réseau (*p.ex.* en considérant vitesse de propagation médiane). Il est probable que l'on obtienne des résultats différents selon le type d'échange (*p.ex.* courriels ou rencontres, le premier étant a priori plus nivellé que le deuxième).

4 Glossaire

Par ordre de dépendance :

- **Degré** d'un nœud : nombre de voisins de ce nœud dans le graphe.
- **Chemin** : succession d'arêtes adjacentes, reliant un nœud à un autre nœud de manière indirecte. La longueur du plus court chemin entre deux nœuds est appelée leur **distance topologique**.
- **Portée** d'une arête (u, v) : distance topologique à laquelle seraient u et v si cette arête si elle était enlevée. On parle d'arêtes de **courte** ou de **longue** portée.
- **Encastrement** d'une arête (u, v) : proportion de voisins qu' u et v ont en commun : $|N_u \cap N_v| / |N_u \cup N_v|$. Remarquons qu'un encastrement non nul implique une portée de 2.
- **Distance temporelle** d'un nœud u à un un nœud v : temps minimal qu'il faut à une information pour se propager de u à v . On considère que cette distance ne change pas au cours du temps (car on ne manipule ici que des fréquences moyennes).
- **Excentricité temporelle** d'un nœud n : distance entre n et le nœud le plus lointain (temporellement parlant).
- Graphe **connexe** : graphe dans lequel il existe un chemin entre n'importe quel couple de nœuds.
- Un nœud est dit **critique** si sa suppression déconnecte le graphe.
- Arête **essentielle** : qui contribue à réduire la distance temporelle entre ses deux extrémités (c.-à-d. qu'il n'existe pas de chemin indirect plus rapide).
- **Dorsale** \mathcal{H} d'un graphe \mathcal{G} : sous-graphe constitué des seules arêtes essentielles.
- **Concentration** vs. **Nivellement** de charge : augmentation du trafic (à volume égal pour chaque nœud) sur les arêtes les plus fréquemment utilisées vs. les moins fréquemment utilisées.

À noter que si l'on considère que les interactions par échange de courriels sont unidirectionnelles, alors les définitions ci-dessus doivent être transposées dans un contexte de **graphe orienté** (graphe dont les arêtes sont dotées d'une direction).

5 Le questionnaire

Cette section discute la manière dont le sondage pourrait être mené à un niveau très général. Essentiellement, il pourrait être réalisé de manière physique (via l'organisation de réunions avec les réseaux) ou électronique (par l'intermédiaire de formulaires de type HTML). L'avantage de la première option est qu'elle fournit une occasion de rencontrer ces réseaux, et permet à leurs membres d'échanger avec leurs homologues d'autres réseaux. (Des rencontres de ce type ont été organisées par le passé par l'équipe de Louise Bouchard, en regroupant à chaque fois un petit nombre de réseaux sur critères géographiques.) Cela permet également d'intégrer des aspects plus *qualitatifs* à l'analyse. Le principal inconvénient de cette option est qu'elle requiert une cartographie préalable des réseaux afin de décider qui inviter (et qui interroger). La voie électronique est plus souple à ce niveau, car elle permet de cartographier les réseaux de manière complète et récursive. On pourrait par exemple commencer par les directeurs de chaque réseau, puis interroger les personnes qui sont apparues parmi les contacts de ces directeurs, puis interroger les personnes qui sont apparues parmi les contacts de ces contacts, jusqu'à une profondeur ou un seuil d'interaction limites (liés au problème de détection des frontières évoqué plus haut). Cette méthode permet aussi accessoirement d'automatiser de bout en bout la chaîne de traitement des données. Mais je pense que la meilleure solution réside probablement dans une combinaison des deux approches.

Concernant les données à récolter, si l'on ne se concentre que sur le type d'analyse proposé ici (qui rappelons-le n'a pour objectif que de compléter une étude plus vaste et qualitative, à laquelle elle ne saurait se substituer), on a essentiellement besoin de collecter les informations suivantes :

- Pour chaque membre du réseau, connaître la liste des principaux interlocuteurs.
- Pour chacun de ces interlocuteurs, connaître la fréquence d'interaction selon chaque type de médium (typiquement courriel, téléphone/visioconférence ou rencontre physique). Il est également envisageable de décliner ces choix selon plusieurs niveaux d'intensité (*e.g.* rencontre brève, discussion poussée, etc.) ou bien de demander à l'interrogé une « note » personnelle globale d'*intensité* pour chacun de ses interlocuteurs. Soulignons que la notion d'intensité n'est pas intégrée à l'analyse proposée dans ce document, qui se concentre essentiellement sur la fréquence des interactions, mais elle est certainement importante et pourrait être intégrée ultérieurement.
- De manière optionnelle, récolter des informations sur les thèmes d'interactions (*p.ex.* échanges de connaissances techniques, de moyens logistiques, interactions amicales). Ces informations pourraient être utilisées pour étudier différents niveaux de fonctionnement du réseau, mais complexifie également beaucoup le questionnaire. Ces questions seraient peut-être mieux traitées à un niveau moins formel et plus qualitatif – c'est à discuter.

Enfin, un problème central de l'élaboration du questionnaire et de l'analyse est celui de l'anonymisation des personnes et de leurs réponses. Il est *nécessaire* pour nous de demander le nom réel des interlocuteurs afin de pouvoir croiser les réponses et reconstituer la topologie du réseau. Rien ne nous oblige, en revanche, à traiter les données de manière nominative une fois le graphe construit. Il faudra insister sur ce point lors du sondage et bien expliquer aux participants que les résultats que l'on va générer seront purement statistiques (il ne faut surtout pas qu'ils se sentent évalués personnellement). Normalement le nombre de réseaux – à ce jour 17 – devrait être suffisant pour garantir que les personnes ne soient pas identifiables à partir des résultats portant sur le fonctionnement interne des réseaux. Pour la partie de l'analyse étudiant les disparités entre provinces, en revanche, il faudra être très prudent.

6 Questions ouvertes

Comme déjà évoqué plus haut, le type d'analyse proposé dans ce document s'attache à caractériser la dynamique d'interaction d'un réseau en fonction des fréquences de rencontres et de communications entre ses membres. Il pourrait être intéressant de considérer d'autres aspects tels que l'intensité ou la durée des interactions. Au-delà de cette dynamique, se pose la question de la formation et du développement des réseaux, et il serait bon à terme d'étudier la manière dont la structure elle-même évolue au cours du temps. Par exemple, les rencontres physiques ont-elles un rôle structurant ? À quel rythme les interactions se renforcent ou s'atténuent-elles généralement (si tant est que ce type de paramètres suive une quelconque forme de régularité) ? Pouvoir répondre à ces questions nécessite cependant une plus grande quantité d'informations, comme par exemple un historique des interactions et des événements dans le réseau.

D'un point de vue plus pratique, une inconnue de taille dans ce projet concerne les outils informatiques nécessaires au traitement des données. En effet, les concepts sur lesquels s'appuie l'analyse étant nouveaux, et il n'existe pas encore d'outils logiciels capables de les manipuler. Les besoins correspondants en termes de conception d'algorithmes et de développements logiciels restent à définir. Certains outils existants pourront probablement être réutilisés pour les mesures classiques issues de la théorie des graphes (diamètre, degré, redondance, plus court chemin...) et de nouveaux outils devront être élaborés, en particulier pour le traitement de données temporelles statistiques sur les arêtes. On peut aussi demander aux auteurs de [9] de nous envoyer les programmes qu'ils ont écrit dans le cadre de leurs travaux, et tenter d'en réutiliser une partie.

Pour terminer, il s'agit de ne pas perdre de vue notre objectif à long terme dans cette analyse, qui est de comprendre la façon dont les réseaux de santé fonctionnent en situation minoritaire. Il serait juste aussi

de réfléchir à ce que notre analyse peut apporter en retour à ces réseaux. Pourrait-on, par exemple, identifier des tendances particulières que l'on retrouve dans la majorité des réseaux et qui en pénalisent le fonctionnement ? Les questions sur la fragmentation, ou encore sur la concentration de charge (Question 14) pourraient déboucher sur des recommandations assez simples. Dans un sens comme dans l'autre, il convient cependant de ne pas trop « mécaniser » ces réseaux, dont le fonctionnement est probablement aussi subtile que les relations humaines sont complexes.

Références

- [1] C. Boelen. Towards unity for health. *Case studies. Geneva : World Health Organization*, 2001.
- [2] Rapport au ministre fédéral de la santé : Comité consultatif des communautés francophones en situation minoritaire. Santé Canada, Septembre 2001.
- [3] Notions sur les réseaux (numéros I à V). Fondation canadienne de la recherche sur les services de santé, www.fcrss.ca/knowledge_transfer/networks_notes_f.php.
- [4] Chronique des réseaux (numéros 32 à 35). Fondation canadienne de la recherche sur les services de santé, www.fcrss.ca/other_documents/networks_digest_f.php.
- [5] M. Callon and M. Ferrary. Les réseaux sociaux à l'aune de la théorie de l'acteur-réseau. *Sociologies pratiques*, (2) :37–44, 2006.
- [6] B. Bui-Xuan, A. Ferreira, and A. Jarry. Computing shortest, fastest, and foremost journeys in dynamic networks. *Intl. J. of Foundations of Comp. Science*, 14(2) :267–285, April 2003.
- [7] A. Ferreira. Building a reference combinatorial model for MANETs. *IEEE Network*, 18(5) :24–29, 2004.
- [8] J. Tang, M. Musolesi, C. Mascolo, and V. Latora. Temporal distance metrics for social network analysis. In *Proceedings of the 2nd ACM workshop on Online social networks*, pages 31–36. ACM, 2009.
- [9] G. Kossinets, J. Kleinberg, and D. Watts. The structure of information pathways in a social communication network. In *Proc. of the 14th ACM SIGKDD intl. conf. on Knowledge discovery and data mining*, pages 435–443. ACM, 2008.