



HAL
open science

Représentations visuelles alternatives pour les réseaux sociaux

Nathalie Henry, Jean-Daniel Fekete

► **To cite this version:**

Nathalie Henry, Jean-Daniel Fekete. Représentations visuelles alternatives pour les réseaux sociaux. Réseaux : communication, technologie, société, 2008, Communication - Technologie - Société, 26 (152), pp.59-92. 10.3166/Réseaux.152.59-92 . hal-00772498

HAL Id: hal-00772498

<https://inria.hal.science/hal-00772498>

Submitted on 11 Jan 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.

Représentations visuelles alternatives pour les réseaux sociaux

Nathalie HENRY

Jean-Daniel FEKETE

Equipe-Projet AVIZ, INRIA

Grâce à la démocratisation des technologies Internet, il devient plus facile de collecter de grandes quantités de données et d'extraire de grands réseaux sociaux afin de les étudier. Par exemple, au travers de sites tels que Facebook ou Friendster, les réseaux sociaux disponibles sont très vastes (des millions d'utilisateurs), riches (contenant de multiples relations ou informations sur les acteurs du réseau) et évoluent dans le temps.

Si la collecte des données est grandement simplifiée, le problème de l'analyse se pose de manière aigüe. Plusieurs méthodes et outils existent pour analyser les réseaux sociaux. La majorité de ces systèmes sont basés sur des méthodes statistiques et fournissent un grand nombre de fonctionnalités d'analyse et de modélisation. UCINet¹ est un bon exemple de ce type de logiciel, permettant de calculer un grand nombre de mesures, comme les différents types de centralité, ou bien de vérifier si le jeu de données correspond à un modèle statistique donné. De nombreux autres outils existent et sont présentés plus en détail dans l'article de Huisman et Duijn².

Dès les années 1930 sont apparues des représentations visuelles de réseaux sociaux. Par exemple, Jacob Moreno a été un des pionniers à utiliser la représentation nœud-lien pour communiquer sur ses travaux (Figure 1). Dans cette représentation, un nœud représente un acteur et un lien représente une relation du réseau. Si les visualisations ont très souvent été utilisées dans ce cadre, *i.e.*, communiquer des résultats ou illustrer une théorie, elles peuvent aussi servir à analyser les données. Le statisticien Tukey, en introduisant l'Analyse Exploratoire des Données³, démontre que la représentation graphique peut être un outil puissant d'analyse. Son intérêt réside dans l'observation de multiples représentations de données brutes, qui permet de soulever des questions ou établir des hypothèses qui n'auraient pas été imaginées *a priori*. Cette utilisation de la représentation visuelle pour l'exploration a donné lieu à un domaine de recherche à part entière: la visualisation d'information.

¹ Voir <http://www.analytictech.com/>

² HUISMAN et DUIJN, section du livre sur l'analyse des réseaux sociaux avec Pajek.

³ TUKEY, 1977.

**EMOTIONS MAPPED
BY NEW GEOGRAPHY**

Charts Seek to Portray the
Psychological Currents of
Human Relationships.

New York Times
April 3, 1933

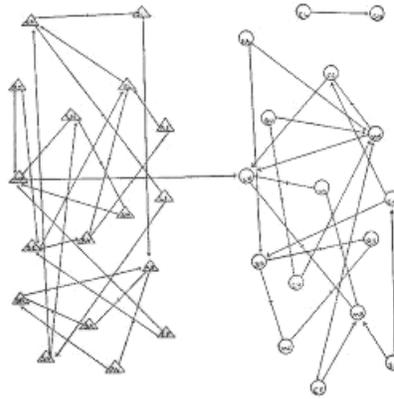


Figure 1 : Réseau d'amitiés entre garçons (triangles) et filles (cercles) par J. Moreno. Un acteur central lie les deux groupes (triangle du milieu gauche).

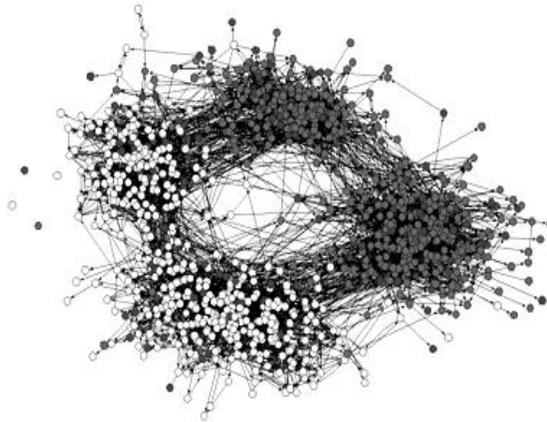


Figure 2 : Réseau d'amitiés entre lycéens par J. Moody, la couleur marque l'origine ethnique. Quatre groupes sociaux émergent après application d'un algorithme de clustering (clusters par origine ethnique et âge)

La représentation la plus courante des graphes ou des réseaux⁴ est la représentation nœud-lien⁵. Cette représentation a donc l'avantage d'être familière à la majorité des chercheurs. Cependant, elle souffre de problèmes de lisibilité lorsque les réseaux représentés sont soit grands (beaucoup de nœuds), soit denses (beaucoup de liens). Ces problèmes de passage à l'échelle sont d'autant plus importants que les données à analyser sont de plus en plus nombreuses ; ils sont devenus rédhibitoires avec l'apparition des réseaux sociaux en ligne qui sont à la fois grands et denses.

Dans les cinq dernières années, le domaine de la visualisation d'information a connu plusieurs avancées et a permis de trouver des représentations alternatives (et complémentaires) aux représentations nœud-lien. Dans cet article, nous faisons un état de l'art de ces nouvelles représentations visuelles des réseaux.

REPRESENTATIONS NŒUD-LIEN

L'analyse des réseaux sociaux a débuté il y a plus de 70 ans, avec les travaux empiriques de Jacob Moreno⁶. Wasserman et Faust⁷ présentent les diverses catégories de méthodes : analyses statistiques, structurelles et exploratoires. Freeman⁸ retrace l'historique des visualisations de réseaux sociaux et montre que les représentations visuelles peuvent être un outil efficace pour illustrer des concepts tels que les acteurs centraux ou les groupes sociaux. Les Figure 1 et Figure 2 présentent deux exemples utilisant les diagrammes nœud-lien.

Le très grand avantage des diagrammes nœud-lien est leur intuitivité : la grande majorité des lecteurs peut les comprendre. En revanche, qu'ils soient dessinés manuellement (Figure 1) ou générés automatiquement (Figure 2), leur lisibilité dépend totalement du placement des nœuds dans le plan. Ce problème épineux a

⁴ Dans cet article, nous utiliserons indifféremment les termes graphe et réseau.

⁵ En parcourant le répertoire de l'International Network for Social Network Analysis (<http://www.insna.org>), il est possible de compter plus de 50 systèmes basés sur les représentations nœud-lien.

⁶ MORENO, 1934.

⁷ WASSERMAN et FAUST, 1994.

⁸ FREEMAN, 2000.

d'ailleurs donné naissance à un domaine de recherche à part entière nommé le dessin de graphes (*graph drawing*).

Dessin de graphes et visualisation d'information

L'ouvrage de Di Battista *et al.*⁹ constitue une bonne introduction au domaine du dessin de graphes, présentant plus de trois cents algorithmes. Un état de l'art de la visualisation de graphes¹⁰, plus récent, recense également un grand nombre de techniques, parfois interactives, pour représenter les graphes et les réseaux. En général, ces techniques tentent d'optimiser un certain nombre de critères esthétiques, comme, par exemple, limiter le nombre de croisements des liens ou bien placer les nœuds pour avoir une longueur uniforme de tous les liens. Plusieurs études ont été réalisées pour déterminer quels critères sont les plus importants et ainsi tenter de découvrir le meilleur algorithme de dessin de graphe. Cependant, il existe un tel nombre de critères qu'il est difficile de former des conclusions générales.

Le domaine de la visualisation d'information a une perspective différente sur le sujet. Ce domaine de recherche se focalise sur l'exploration visuelle et la découverte ou la communication d'informations. Ainsi, les réseaux présentés en Figure 1 et Figure 2 n'optimisent certainement pas des critères esthétiques, comme la minimisation du nombre de croisements de liens, mais ils permettent de mettre en lumière certaines informations importantes sur la structure du réseau. Dans cet article, nous présentons une série de représentations alternatives aux diagrammes nœud-lien, non pas pour les remplacer mais plutôt pour les compléter. Dans le contexte de l'analyse exploratoire, nous pensons qu'il est nécessaire d'avoir plusieurs perspectives sur les données pour en découvrir toutes les arcanes.

Exploration de réseau et passage à l'échelle

Un problème récurrent se pose quand le réseau à représenter contient beaucoup de nœuds ou de liens : le diagramme nœud-lien se transforme en un amas de traits et de points difficile voire impossible à transformer en représentation

⁹ Di BATTISTA *et al.*, 1998

¹⁰ HERMAN *et al.*, 2000

lisible, ni manuellement ni automatiquement. Dans la littérature, il existe trois techniques distinctes pour tenter de résoudre ce problème :

1. **Réduire la quantité d'information représentée par filtrage ou agrégation.** Il existe plusieurs méthodes pour échantillonner les réseaux¹¹ ou bien calculer des clusters¹² (groupe d'éléments similaires), qui peuvent être ensuite agrégés en un élément représentatif du groupe.

2. **Proposer une représentation interactive pour explorer la totalité du réseau.** Plusieurs outils de visualisation ne représentent qu'une partie du réseau et proposent de naviguer interactivement pour explorer le reste du réseau. Les stratégies varient : certaines se concentrent sur un nœud particulier et permettent la navigation en suivant les connexions¹³, d'autres sont basées principalement sur du filtrage interactif¹⁴ ou bien permettent la navigation dans un réseau agrégé¹⁵.

3. **Utiliser des représentations alternatives aux diagrammes nœud-lien.** Cette stratégie consiste à utiliser différentes métaphores visuelles pour la représentation de réseaux. L'idée est de découvrir des représentations permettant d'« augmenter l'espace visuel », permettant d'afficher un plus grand nombre d'information de façon plus lisible.

La psalmodie de la visualisation d'information¹⁶ est : « vue d'ensemble d'abord, zoomer et filtrer, puis détails à la demande » (*overview first, zoom and filter, then details on demand*). Cette phrase résume le processus d'exploration visuel. Ainsi, il est primordial de présenter une vue d'ensemble du réseau afin d'initier le processus d'exploration. Les deux premières stratégies, visant à supprimer l'information ou à n'en présenter qu'une partie, ne sont donc pas satisfaisantes. Nous nous sommes donc intéressés à la troisième solution : utiliser des représentations alternatives pour explorer les réseaux.

¹¹ FRANK, 1978

¹² JAIN *et al.*, 1999

¹³ HEER *et al.*, 2005 et LEE *et al.*, 2006

¹⁴ PERER *et al.* 2006 et SHNEIDERMAN *et al.*, 2006

¹⁵ ABELLO *et al.*, 2006

¹⁶ SHNEIDERMAN, 1996

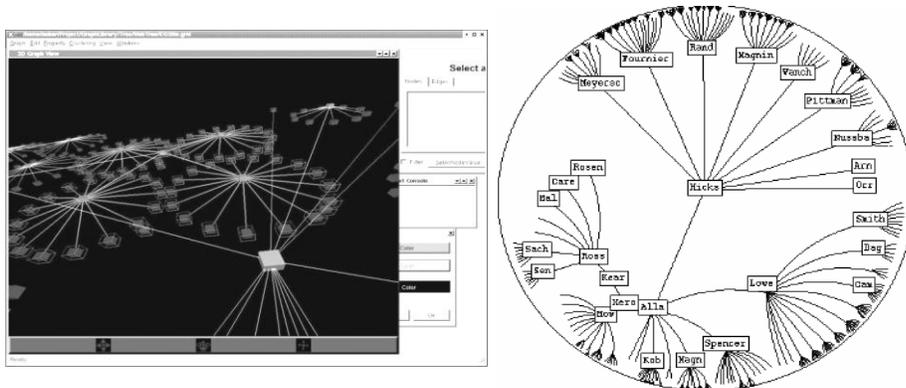


Figure 4 : Représentation d'un réseau en 3D (à gauche) et dans l'espace hyperbolique (à droite).

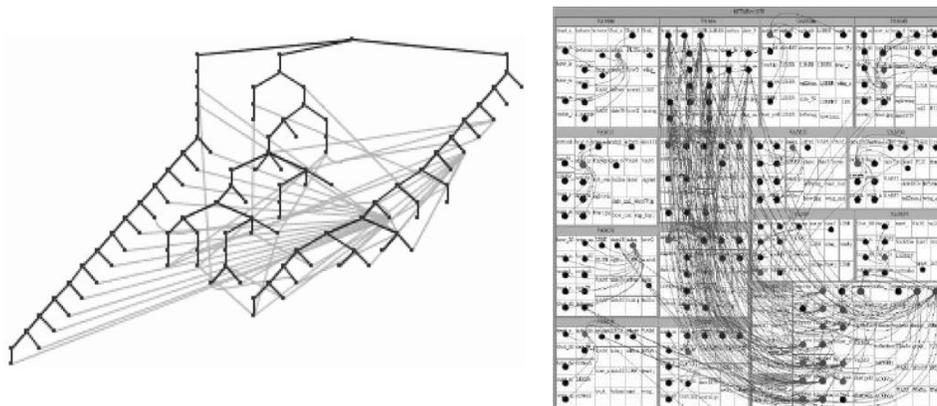


Figure 3 : Représentation d'un réseau sous forme d'arbre avec des liens supplémentaires (à gauche) et sous forme de Treemap avec des liens supplémentaires (à droite).

REPRESENTATIONS ALTERNATIVES

3D et espaces géométriques non-Euclidiens

Afin d'augmenter l'espace visuel, un certain nombre de chercheurs ont tenté d'utiliser la 3^{ème} dimension. L'idée est qu'une troisième dimension permet d'avoir plus d'espace pour dessiner de plus grands réseaux mais aussi permet d'optimiser certains critères esthétiques tels que la minimisation du nombre de croisements. Plusieurs programmes utilisent la 3D, la Figure 4 (gauche) montre un exemple créé avec la boîte à outils de dessin de graphes Tulip¹⁷.

Le principal problème lié à l'utilisation de la 3D est l'occlusion : les utilisateurs ont l'impression que les représentations sont « encombrées » et trouvent difficile d'avoir une représentation mentale de la totalité du graphe¹⁸. Pour palier à ces problèmes, plusieurs stratégies ont été adoptées¹⁹, comme l'utilisation de multiples vues ou des techniques de navigation pour voir la représentation sous de multiples angles. Malheureusement, dans la majorité des cas, ces techniques désorientent l'utilisateur et rendent les visualisations peu efficaces²⁰. Plusieurs études ont montré que, si les visualisations 3D sont attractives, elles n'améliorent pas les performances et les détériorent même pour certaines tâches²¹.

Une autre façon d'augmenter l'espace visuel est l'utilisation de la géométrie hyperbolique au lieu de la géométrie Euclidienne. Dans l'espace hyperbolique, le postulat du parallélisme est rejeté : deux lignes parallèles dans l'espace euclidien divergent l'une de l'autre. Donc, si l'on considère un disque dans l'espace hyperbolique, l'espace augmente exponentiellement lorsque l'on s'éloigne de son centre. Ainsi, un réseau dessiné sur ce disque dispose d'un espace infini sur ses bords, ce qui permet d'afficher plus d'information, comme le montre Figure 4 (droite). La métaphore s'applique aussi à la 3D²². Malheureusement, similairement à la 3D, naviguer dans un espace géométrique non euclidien

¹⁷ AUBER, 2003

¹⁸ COCKBURN *et al.*, 2002

¹⁹ ELMQVIST *et al.*, 2007

²⁰ SUTCLIFFE *et al.*, 1996

²¹ COCKBURN *et al.* 2000

²² MUNZNER, 1997

désoriente l'utilisateur. Il est de plus difficile de se constituer une carte mentale du réseau sans recours à la géométrie 2D euclidienne²³.

Représenter un graphe par un arbre

Une autre stratégie pour représenter des réseaux peu denses et augmenter leur lisibilité est de les représenter sous forme d'arbres et de le « réparer » avec des connexions supplémentaires. Les algorithmes de dessin d'arbres permettent effectivement d'ôter tous croisements voire d'exploiter au maximum l'espace utilisé si l'arbre est dessiné en utilisant des inclusions de formes.

Pour dessiner un réseau sous forme d'arbre, il est nécessaire de calculer un arbre recouvrant (généralement on prend le plus grand), de dessiner cet arbre puis d'ajouter les liens supplémentaires²⁴. Un exemple est présenté dans la Figure 3 (gauche). Cette technique fonctionne bien pour des réseaux ayant une structure différant peu de celle d'un arbre tels que les arbres généalogiques par exemple, mais la représentation devient illisible pour des réseaux plus denses.

Une alternative est la représentation sous forme de Treemap²⁵. Le principe est identique au précédent, mais la représentation utilisée pour l'arbre recouvrant diffère : il s'agit de découper l'espace successivement pour représenter la hiérarchie (imbrications de boîtes). Par exemple, un arbre très simple composé d'une racine, de deux nœuds et de quatre feuilles sur chacun d'eux, est alors représenté par un rectangle occupant tout l'espace visuel, contenant deux rectangles de taille identique (représentant les nœuds), eux-mêmes sub-divisés en quatre rectangles de taille identique. Les liens supplémentaires du réseau sont alors dessinés par dessus le Treemap²⁶ comme montré sur la Figure 3 (droite). Cette représentation devient également difficile à lire si le nombre de liens supplémentaires est important.

²³ RISDEN *et al.*, 2000

²⁴ HERMAN *et al.*, 1999

²⁵ SHNEIDERMAN, 1992

²⁶ FEKETE *et la.*, 2003

L'ajout d'interaction est une piste intéressante pour améliorer la lisibilité de ces représentations. Elle est particulièrement bien exploitée dans TreePlus²⁷, un système visualisant une portion du réseau sous forme d'arbre et permettant de naviguer interactivement dans le reste. Afin de résoudre le problème des cycles, dans cette représentation sous forme d'arbre, des nœuds sont dupliqués. Si ce système permet de mieux réaliser un certain nombre de tâches, il rend difficile la constitution d'une carte mentale de la totalité du réseau.

Représentation Matricielle

Les graphes ont deux représentations canoniques : les diagrammes nœud-lien et les matrices d'adjacence. Une matrice d'adjacence représente chaque sommet d'un réseau à la fois comme une ligne et comme une colonne. Si deux sommets sont connectés, la case correspondant à l'intersection de la ligne et de la colonne est marquée. Traditionnellement, on utilise une valeur numérique (0 marquant l'absence de connexion, 1 marquant la présence), soit par une marque graphique comme dans la Figure 5.

La table ci-dessous liste les principaux avantages et inconvénients des représentations matricielles par rapport aux diagrammes nœud-lien.

<i>Pros</i>	<i>Cons</i>
Absence de superpositions des nœuds (ce qui permet de pouvoir toujours lire les étiquettes portées par les nœuds)	Taille de l'espace visuel requis à un niveau de détail équivalent plus important que le diagramme nœud-lien
Absence de croisements des liens (ce qui permet de toujours identifier la source et destination des connexions)	Difficulté à effectuer des tâches de suivi de chemin (suivre les liens de Pierre à Paul passant par Marie)
Facilité avec laquelle il est possible d'identifier les absences de connexions	Manque de familiarité, les matrices paraissent moins intuitives que les diagrammes nœud-lien

²⁷ LEE *et al.*, 2006

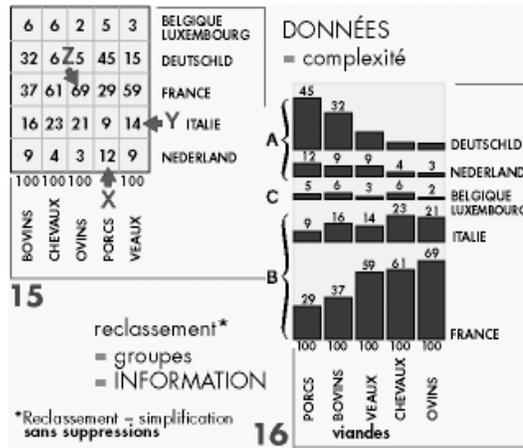


Figure 5 : Matrice réordonnable de J. Bertin, extraite de la Sémiologie graphique.

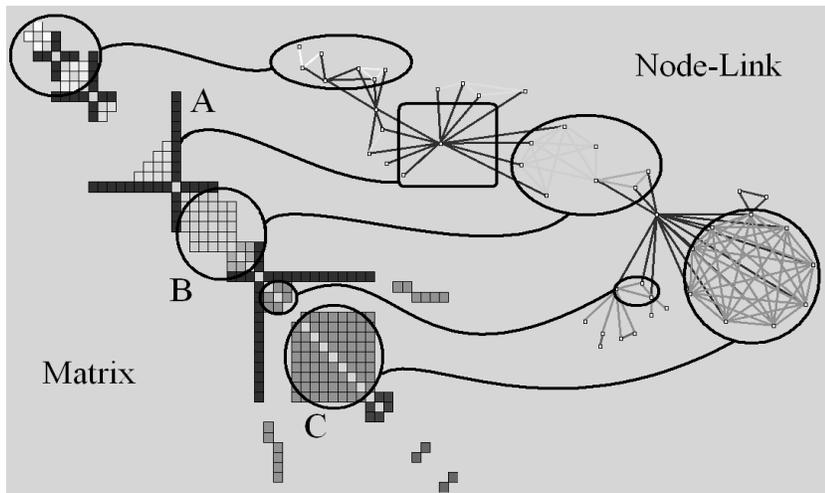


Figure 6 : Correspondance entre les motifs visuels des matrices et ceux des diagrammes nœud-lien : A montre un acteur connectant plusieurs communautés, B deux communautés et C une clique.

Plus de détails sur les performances des deux représentations de réseaux sont présentés dans les travaux de Ghoniem *et al*²⁸. Il s'agit d'une comparaison plus quantitative de la lisibilité des représentations matricielles et des représentations nœud-lien pour des graphes de différentes tailles et densités. Dans cette étude, une liste de tâches de base a été utilisée pour évaluer la lisibilité des représentations : estimer le nombre de nœuds, estimer le nombre de connexions, trouver le nœud le plus connecté, trouver le nœud portant une étiquette donnée, trouver la connexion entre deux nœuds donnés, trouver un nœud commun à deux nœuds donnés et trouver un chemin entre deux nœuds donnés.

De manière générale, les résultats de leur expérimentation montrent que les nœud-lien sont plus efficaces pour les graphes de petite taille et peu denses alors que les matrices ont de meilleures performances dès que les graphes sont plus grands (>20 nœuds !) et plus denses pour toutes les tâches excepté le suivi de chemin. Ces résultats ainsi que les pros et cons listés précédemment révèlent le vaste potentiel de la représentation matricielle pour explorer les réseaux sociaux.

UTILISER LA REPRESENTATION MATRICIELLE

Deux facteurs principaux entrent en jeu lors qu'il s'agit d'utiliser les représentations matricielles pour explorer de grands réseaux :

1. Être capable de réordonner leurs lignes et colonnes
2. Pouvoir naviguer dans des matrices de grande taille.

Réordonner les matrices

Dans son livre faisant référence en visualisation d'information²⁹, Bertin présente la matrice réordonnable. Au travers de plusieurs exemples, il démontre que réordonner les lignes et colonnes d'une matrice permet d'améliorer significativement sa compréhension.

Prenons pour exemple la Figure 5 extraite de ce livre ; la matrice présentée contient seulement 5 lignes et 5 colonnes représentant la production de 5 types de viande dans 5 pays. En transformant la matrice originale, contenant des valeurs numériques, en matrice visuelle, codant les valeurs par des formes

²⁸ GHONIEM *et al.*, 2005

²⁹ la Sémiologie Graphique, BERTIN, 1967

géométriques de différentes tailles et en réordonnant « intelligemment » ses lignes et ses colonnes, il devient beaucoup plus simple d'interpréter ce jeu de données. En effet, il est pratiquement immédiat d'observer que la France est le pays ayant la plus grande production, tout type de viande confondu ; que la Belgique a la plus petite et qu'il existe trois différents profils de production (marqué par A, B et C dans la figure).

Pour aller une étape plus loin dans l'interprétation de ce simple exemple, imaginons qu'une loi doive être votée pour limiter la production d'un type de viande produit en grande quantité par les pays. Disposant de ces données et de la représentation visuelle associée, un analyste pourra rapidement observer que la Belgique est le pays à convaincre de ne pas accepter cette loi, car son profil de production est plutôt neutre alors que les autres pays voteront dans leur propre intérêt, contraire à celui de l'autre groupe.

Cet exemple démontre l'importance du réordonnement des matrices: si les représentations matricielles dont les lignes et les colonnes sont ordonnées aléatoirement sont lisibles, une fois réordonnées elles deviennent réellement exploitables.

Algorithmes de réordonnement

Effectuer un état de l'art des techniques de réordonnement automatique est un véritable défi car un grand nombre de problèmes sont liés à la permutation des lignes et colonnes. Par exemple, un algorithme de linéarisation de graphe (alignement de tous les sommets d'un graphe de façon à minimiser un critère esthétique tel que le nombre de croisements) ou bien un algorithme d'optimisation de tables de calculs peut être utilisé pour réordonner une matrice d'adjacence. Les algorithmes sont donc disséminés dans divers domaines tels que la biologie ou le dessin de graphe; et leurs critères de qualité varient en fonction du contexte d'utilisation.

Plusieurs outils tels que PermutMatrix30 ou VisuLab31 offrent une représentation visuelle des matrices (ou de tables) et permettent d'expérimenter

³⁰ CARAUX *et al.*, 1998

³¹ voir <http://www.inf.ethz.ch/personal/hinterbe/Visulab>

avec diverses techniques de réordonnement et leurs paramètres. Nous ne détaillerons pas les diverses catégories de méthodes dans cet article, mais il est important de comprendre que la méthode de réordonnement sélectionnée a un fort impact sur l'interprétation de la représentation, similairement aux méthodes de placement de nœuds des diagrammes nœud-lien. La Figure 6 montre un exemple de motifs visuels émergeant lorsque l'ordonnement de matrice et le placement du nœud-lien sont optimisés.

Naviguer dans de grandes matrices

À niveau de détail équivalent, la représentation matricielle utilise plus d'espace que les diagrammes nœud-lien. Par exemple, sur un écran de 17 pouces, il est possible d'afficher seulement une centaine de nœuds si l'on souhaite pouvoir lire confortablement leurs noms. Ainsi, afin de tirer parti du potentiel des matrices, il est nécessaire d'avoir des outils efficaces de navigation.

De nombreuses techniques existent dans le domaine de l'Interaction Homme-Machine pour permettre de se déplacer dans de grands espaces visuels et/ou à divers niveaux de détails. Par exemple, les techniques de Focus+Context³² telles que la vue à vol d'oiseau (*bird's eye view*) permettent d'avoir une vue d'ensemble miniature et une vue zoomée d'une partie de la représentation. Ces techniques ne sont pas directement utilisables dans notre contexte, elles requièrent quelques aménagements. Il est notamment indispensable de toujours voir les étiquettes des lignes et colonnes des matrices et de faciliter les transitions entre les divers niveaux de détails pour éviter à l'analyste de se perdre lors de l'exploration. Plusieurs systèmes intègrent ces outils de navigation tels que MatrixZoom³³ et la matrice d'appels multi-niveaux présentée par van Ham³⁴.

³² CARPENDALE, 1999

³³ ABELLO *et al.*, 2004

³⁴ van HAM *et al.*, 2003

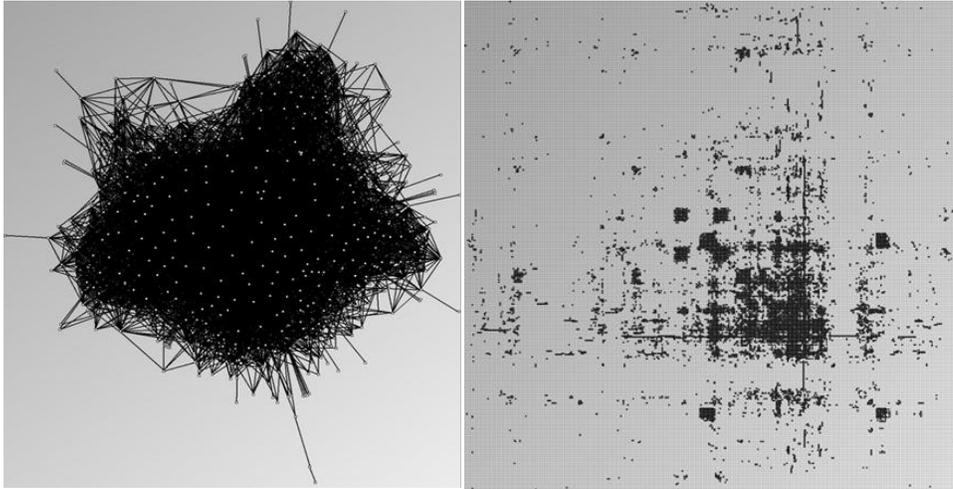


Figure 7 : Réseau d'échange de courriers électroniques entre plus de 450 chercheurs, pendant un an.

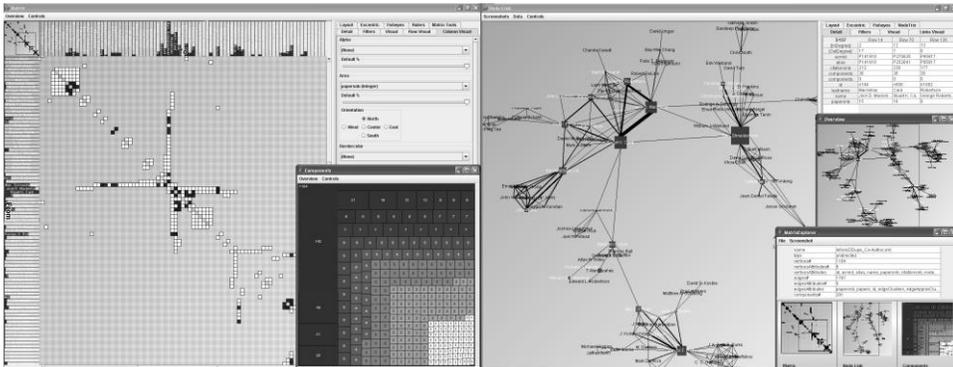


Figure 8 : Interface de MatrixExplorer, combinant matrices (grande fenêtre de gauche) et nœud lien (grande fenêtre de droite). Une représentation Treemap (petite fenêtre de gauche) montre la macrostructure du réseau, i.e. ses composantes connexes. Des vues miniatures des représentations permettent la navigation (petites fenêtres de droite).

COMBINER MATRICES ET NŒUD-LIEN

Les représentations matricielle et nœud-lien ont chacune des avantages et inconvénients à la fois pour l'exploration et la communication de résultats. Dans cette section, nous présentons notre système MatrixExplorer³⁵, qui utilise une représentation duale des réseaux et tente de combiner le meilleur de deux mondes.

Initier l'exploration

L'avantage majeur des matrices par rapport aux diagrammes nœud-lien est de permettre d'avoir une vue d'ensemble du réseau toujours exploitable et ce, même pour de très grands graphes, alors que les diagrammes nœud-lien ne montrent bien souvent qu'un amas de liens et de nœuds superposés.

Prenons l'exemple d'un réseau social contenant les données d'échanges de courrier électronique entre plus de 450 personnes pendant un an (les sommets sont les personnes et les liens sont les courriers). La Figure 7 montre deux représentations de ce réseau. Alors que la représentation nœud-lien (utilisant un algorithme de placement traditionnel par forces) permet éventuellement de former l'hypothèse que le réseau est très dense; la représentation matricielle, elle, donne beaucoup plus d'information. Pour le démontrer, analysons brièvement la matrice présentée à la Figure 7.

Dans cette représentation, chacun des points noirs représente une connexion entre une ligne et une colonne (un lien dans le réseau), le gris marquant donc l'absence de connexion. Tout d'abord, on peut clairement rejeter l'hypothèse précédente : le réseau social n'est pas très dense car on peut observer une majorité de gris dans cette matrice. Deuxièmement, si l'on regarde de plus près la représentation, on peut voir plusieurs blocs de points noirs. Il s'agit de clusters : des groupes de recherche dont les membres échangent des courriers électroniques les uns avec les autres. Une autre observation permet de voir une sorte de croix : une ligne constituée de points noirs verticale et une horizontale

³⁵ HENRY *et al.*, 2005 et 2006

(à peu près de la longueur d'une moitié de la matrice). Il s'agit d'un membre du service des missions de cet institut, celui-ci échange effectivement des courriers électroniques avec beaucoup des membres d'équipes, il a donc un grand nombre de connexions.

Ce simple exemple démontre que les matrices procurent une vue d'ensemble permettant d'initier l'exploration d'un jeu de données en mettant en évidence plusieurs éléments saillants. Le point important est d'apprendre à les décoder, car ces représentations sont beaucoup moins familières que les diagrammes nœud-lien. La Figure 6 donne quelques exemples de correspondance entre les motifs visuels courant dans les réseaux sociaux à la fois pour les représentations nœud-lien et matricielles.

Explorer interactivement

À l'issue d'entretiens avec des analystes et de séances d'observation, nous nous sommes rendu compte que le processus d'exploration visuel est constitué d'opérations successives pour configurer les représentations (ajuster la position des éléments graphiques et les attributs visuels), filtrer certains éléments, les grouper et éventuellement les agréger. Il est important de créer de vues multiples d'un jeu de données, l'analysant sous divers aspects et plusieurs niveaux de détail afin d'identifier sa structure. Par exemple, une représentation matricielle peut permettre d'identifier plusieurs communautés alors qu'un diagramme nœud-lien égo-centré peut permettre l'analyse détaillée des connexions d'un acteur important. Dans MatrixExplorer, nous avons donc décidé de proposer plusieurs vues sur les données (Figure 8) et d'offrir un grand nombre d'outils pour manipuler directement³⁶ les représentations.

Tout d'abord, nous avons synchronisé les deux visualisations (matricielle et nœud-lien) afin de combiner leurs avantages et de faciliter l'identification des motifs visuels. Il est ainsi possible d'utiliser les matrices pour certaines tâches et les diagrammes nœud-lien pour d'autres. Si un analyste souhaite avoir une vue d'ensemble sur un réseau et, par exemple, identifier ses principales communautés, il est préférable qu'il utilise une représentation matricielle. Par

³⁶ Manipulation directe, SHNEIDERMAN, 1983 et requêtes dynamiques, AHLBERG *et al.*, 1992

contre, lors d'une analyse plus détaillée, s'il est nécessaire d'identifier un plus court chemin entre deux acteurs du réseau, le diagramme nœud-lien constitue une meilleure alternative.

De plus, synchroniser les deux représentations par la sélection permet de comparer la représentation d'un motif visuel dans l'une et l'autre représentation, facilitant ainsi leur compréhension (et l'apprentissage des matrices).

Pour manipuler interactivement les deux représentations nous avons mis en place un certain nombre de fonctionnalités :

1. Spécification interactive des attributs visuels comme la couleur des nœuds, des en-têtes des lignes et de colonnes dans les matrices, ou bien encore l'épaisseur des liens ou la couleur des cellules de la matrice.
2. Méthodes automatiques et interactives pour le placement des nœuds des diagrammes nœud-lien et le réordonnancement des lignes et colonnes des matrices. En particulier, nous avons développé plusieurs outils d'aide au réordonnancement en proposant d'appliquer des algorithmes automatiques sur des portions de matrices ou bien de bloquer l'ordre de plusieurs lignes et colonnes lorsque l'analyste souhaite les conserver.
3. Filtrage interactif de nœuds ou de connexions en fonction d'une sélection ou d'un attribut particulier présent dans les données (l'âge ou le sexe des personnes contenues dans le réseau social par exemple).
4. Clustering interactif permettant d'identifier les différents groupes ou communautés détectées pour les marquer par un attribut visuel tel que la couleur.
5. Finalement, plusieurs vues d'ensemble sont proposées : miniatures des visualisations dans lesquelles l'analyste peut se déplacer ou représentations sous forme de Treemap de la macrostructure du réseau (voir Figure 9).

En combinant les deux représentations et permettant de les manipuler interactivement, les utilisateurs de MatrixExplorer peuvent créer de multiples vues sur leurs données et les explorer à divers niveaux de détail.

Trouver un consensus

Comme chaque visualisation présente une perspective différente, nous avons tenté de favoriser l'établissement de consensus à l'aide d'interactions simples. La

Figure 9 donne un exemple. Pour faciliter cette recherche de consensus, nous avons également introduit la sélection approximative : possibilité de marquer le degré d'appartenance d'un élément à un cluster (représenté par une couleur plus claire par exemple) ; et le clustering multiple : possibilité de marquer l'appartenance d'un élément à plusieurs clusters.

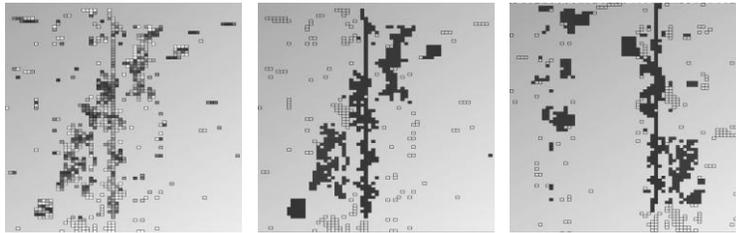


Figure 9 : L'image du milieu montre l'identification de plusieurs clusters de la matrice initiale (à gauche). Après réordonnancement, nous constatons que l'un des clusters les plus foncés est conservé alors que l'autre se retrouve divisé en trois morceaux. Il y a donc seulement consensus sur une partie des clusters identifiés (celui de couleur claire et l'un des foncés), il faut continuer à explorer pour en identifier éventuellement d'autres.

Présenter l'information

Alors que la matrice permet principalement d'explorer les données et d'en manipuler une grande quantité, le diagramme nœud-lien est indispensable pour présenter des résultats et communiquer des découvertes. Ainsi, une fois le réseau filtré et éventuellement agrégé, MatrixExplorer permet d'extraire directement des images nœud-lien du résultat.

REPRÉSENTATIONS HYBRIDES

Combiner deux représentations présente beaucoup d'avantages, mais aussi plusieurs inconvénients. Tout d'abord, la taille de l'espace visuel : une utilisation confortable nécessite au moins deux écrans (un par représentation), un

troisième offrant un confort supplémentaire en permettant de consulter les données détaillées éventuellement textuelles ainsi que plusieurs vues d'ensemble. De plus, nous avons observé que le passage d'une représentation à l'autre entraînait un coût cognitif non négligeable : un nœud unique se transforme en deux éléments dans la matrice (ligne et colonne) et un lien (visuellement représenté par un trait) devient une cellule (un rectangle) dans la matrice. Pour limiter ce coût cognitif, nous avons donc tenté d'améliorer les représentations matricielles pour limiter le recours au diagramme nœud-lien lors de l'exploration.

Augmenter les matrices

Comme l'ont montré les travaux sur la lisibilité de la matrice³⁷, la tâche de base pour laquelle cette représentation est moins performante que les diagrammes nœud-lien est le suivi de chemin. Pour remédier à ce problème, nous avons créé MatLink³⁸, une représentation matricielle augmentée (Figure 10).

Le principe de MatLink est de dessiner des liens sur les bords de la matrice afin de mieux représenter la connectivité et faciliter le suivi de chemin. Deux catégories de liens sont ajoutées sur la matrice : des liens statiques (en blanc sur l'image) et des liens interactifs (en plus foncé sur l'image) qui apparaissent uniquement lors d'un survol ou d'un clic de la souris sur les en-têtes des lignes et colonnes. Lorsque l'utilisateur sélectionne deux éléments, les liens apparaissent montrant un plus court chemin.

Nous avons réalisé une expérimentation contrôlée similaire à celle de Ghoniem *et al.* afin d'étudier si cette représentation améliorerait significativement la représentation matricielle standard. Nous avons testé un lot de tâche de lisibilité de base ainsi que plusieurs tâches spécifiques à l'analyse de réseaux sociaux : trouver un point d'articulation, trouver une clique et identifier plusieurs communautés (groupe fortement connecté).

³⁷ GHONIEM *et al.*, 2005

³⁸ HENRY *et al.*, 2007

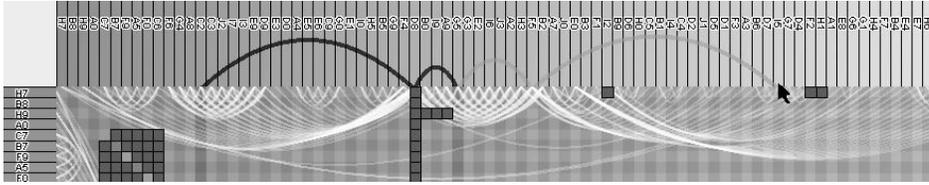


Figure 10 : MatLink augmente les matrices en ajoutant des liens statiques (en blanc) et des liens interactifs (en plus foncés sur les en-têtes)

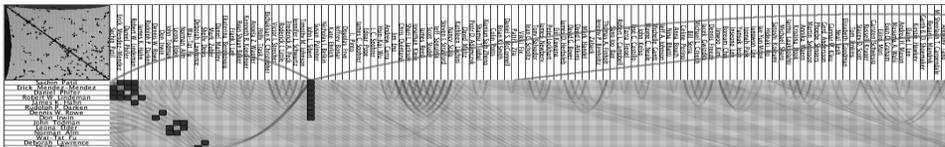


Figure 11 : Les liens permettent à l'analyste d'être conscient que des éléments connectés aux nœuds qu'il étudie existent en dehors de la vue courante.

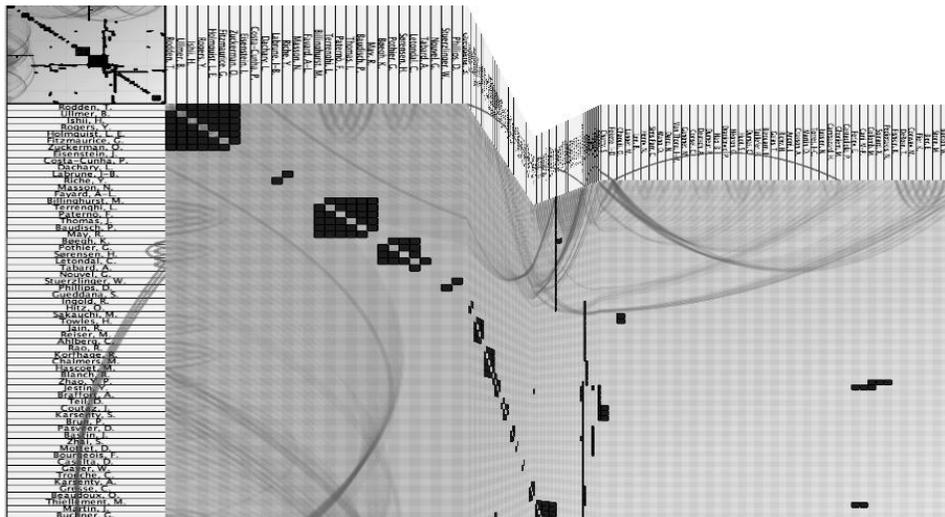


Figure 12 : Mélange permet de « plier » l'espace pour permettre de voir deux régions éloignées de la matrice.

D'une manière générale, nous avons démontré que MatLink améliore significativement la représentation matricielle standard, en particulier pour la tâche de suivi de chemin, auparavant très laborieuse à réaliser. Pour ce cas précis, MatLink montre même des performances supérieures au diagramme nœud-lien pour les graphes denses (et ce, même augmenté de surlignage interactif) à cause du grand nombre de superpositions et de croisements. La seule tâche pour laquelle les diagrammes nœud-lien restent plus performants est l'identification de points d'articulation. Cette tâche requiert assez grande familiarité avec la représentation matricielle, ce qui n'était pas le cas de nos participants (qui, en moyenne, disposaient de 30 minutes d'entraînement avec les matrices).

Naviguer dans de grandes matrices

Le second avantage de MatLink est l'amélioration que la technique offre pour la navigation dans de grandes matrices. Un des inconvénients des matrices est la grande taille requise pour les représenter. Ainsi, il peut (souvent) arriver qu'un acteur soit connecté à plusieurs autres acteurs placés en dehors de la vue. Il est nécessaire de naviguer sur toute la largeur (ou hauteur) de la matrice pour identifier tous les voisins d'un acteur particulier, ce qui est laborieux. Grâce aux liens affichés dans MatLink, et par un simple clic sur l'acteur étudié, l'analyste a un retour visuel immédiat sur tous les éléments connectés à cet acteur et possède une indication pour ceux placés hors de l'écran. La Figure 11 présente un exemple.

En outre, afin de faciliter l'exploration de ces grandes matrices (et le déplacement jusqu'aux éléments placés hors de la vue), nous avons mis au point une technique de navigation « pliant l'espace » : Mélange³⁹. Cette technique permet de visualiser deux parties de la matrice dans une même vue, en offrant un aperçu de l'espace intermédiaire, comme illustré dans La Figure 12. Mélange offre même la possibilité de zoomer sur l'une des régions d'intérêt indépendamment de l'autre.

³⁹ ELMQVIST *et al.*, 2008

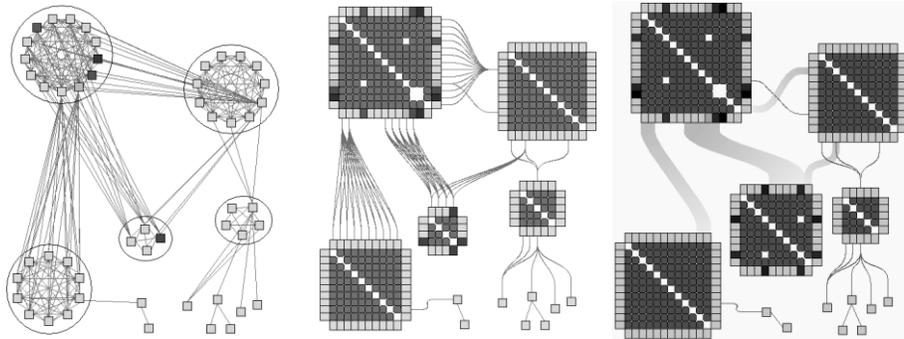


Figure 13 : Représentation nœud-lien (à gauche) dans laquelle il est difficile d'identifier que les nœuds foncés ne sont pas connectés. La représentation NodeTrix (au milieu) visualise les sous-parties denses en matrices, il devient alors possible de voir les absences de connexions (matrice du haut gauche avec des cellules manquantes). Enfin, dupliquer des nœuds (à droite) permet d'améliorer la lisibilité en supprimant des croisements de liens inter-communautés.

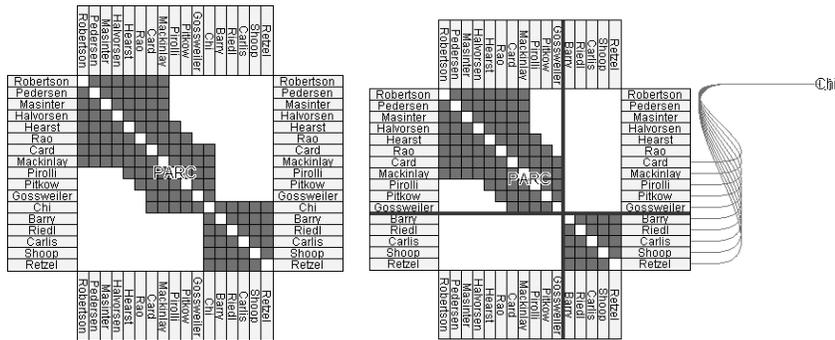


Figure 14 : Interagir avec la représentation matricielle pour mieux comprendre la structure de la communauté. Ici, en ôtant une personne de la communauté, il est possible de voir qu'il s'agit d'un point d'articulation.

Fusionner Matrices et Nœud-lien

La structure des réseaux sociaux varie beaucoup : certains exhibent une structure d'arbre (ou presque) très peu dense tels les arbres généalogiques, d'autres ont des structures quasi complètes (réseau bipartite très dense) tels que les réseaux d'import/export entre pays du monde et enfin, certains ont une structure intermédiaire, localement dense mais globalement creuse, il s'agit des réseaux petit-monde⁴⁰ tels que les réseaux d'amitiés ou encore de communication.

Les représentations visuelles que nous avons présentées précédemment sont plus ou moins performantes selon le type de réseau à visualiser. Par exemple, les diagrammes nœud-lien ainsi que les représentations basées sur des arbres avec liens supplémentaires (Treemap+liens par exemple) sont particulièrement efficaces pour les réseaux peu denses présentant une structure d'arbres. Au contraire, si le réseau à représenter est plutôt dense ou bipartite et présentant une structure quasi complète, la représentation matricielle est alors tout indiquée. Par contre, le choix n'est pas si simple pour les réseaux petit-monde.

Le problème que posent les réseaux petit-monde réside dans la variation de la densité locale. Si l'on choisit un diagramme nœud-lien, alors les parties denses (communautés) ne sont pas lisibles. Inversement, si l'on choisit une matrice, alors la majorité de l'espace est vide et l'exploration requiert beaucoup de navigation. Une possibilité est alors d'utiliser MatrixExplorer, combinant les deux représentations, mais le coût cognitif est plus élevé.

Pour ne plus avoir à choisir, nous avons mis au point une seconde représentation hybride, fusionnant matrice et diagramme nœud-lien : NodeTrix⁴¹.

⁴⁰ WATTS *et al.*, 1998

⁴¹ HENRY *et al.*, 2007

Le principe de NodeTrix est de représenter le réseau global comme un nœud-lien mais d'utiliser une représentation matricielle pour ses sous parties les plus denses. Deux exemples sont présentés Figure 13 et Figure 15.

Explorer

Afin de faciliter la création, l'exploration et l'édition des matrices représentant les parties denses du réseau, nous avons développé un ensemble d'outils interactifs basés sur la technique du glisser-déposer. Il suffit de sélectionner au lasso un ensemble de nœuds fortement connecté pour transformer cette partie du réseau en matrice. Des informations sur la connectivité intra-communauté apparaissent immédiatement comme, par exemple, l'absence de connexion entre deux membres de la communauté, auparavant difficile à percevoir à cause de la multitude de croisements de liens à cet endroit.

En outre, la représentation matricielle offre l'avantage de placer linéairement en ligne et colonnes les membres de la communauté ; il devient plus aisé d'identifier les membres connectés à des acteurs extérieurs. Pour permettre l'exploration de cette représentation, il est possible d'ajouter ou d'ôter interactivement un acteur à la communauté. L'interaction est simple : il suffit de cliquer sur l'élément (nœud ou bien en-tête d'une ligne ou colonne dans la matrice) et ensuite le glisser-déposer à l'intérieur ou à l'extérieur de la matrice (Figure 14). D'autres interactions permettent de fusionner deux matrices ou bien de les casser pour revenir à une représentation normale. Enfin, pour aider l'utilisateur à comprendre le changement de représentation, une courte animation permet de suivre la transformation⁴².

Les matrices des parties denses peuvent être créées manuellement ou automatiquement en utilisant un algorithme de clustering. NodeTrix étant intégré à MatrixExplorer, il est également possible de glisser-déposer des parties d'une matrice dans une fenêtre NodeTrix. Cela permet de mieux comprendre la connectivité inter-communauté, ce qui est surtout important si les deux sous-matrices ne sont pas côte à côte dans la représentation matricielle standard.

⁴² Une vidéo est disponible sur <http://www.aviz.fr/~nhenry>

Un inconvénient de cette représentation est la représentation concrète des communautés. Si un élément est membre de deux communautés, il faut faire un choix et le placer soit dans l'une des matrices, soit le laisser entre les deux (c'est un problème classique posé par l'agrégation hiérarchique de réseaux). Placer un acteur dans une des communautés peut biaiser l'interprétation de la représentation. Par contre, le laisser entre les deux communautés dégrade la lisibilité en introduisant beaucoup plus de liens et de croisements. Pour résoudre ce problème, nous avons proposé de dupliquer les acteurs pour les placer dans deux ou plusieurs communautés à la fois !

Notre étude utilisateur⁴³ a montré que l'impact potentiellement négatif des duplications sur l'interprétation du réseau peut être minimisé en représentant explicitement le lien de duplication. Après avoir exploré l'espace de conception possible, nous avons décidé de représenter ce lien par une bande (un lien plus épais) de couleur claire permettant de minimiser les interférences avec la représentation des connexions « réelles » du réseau. Un exemple est présenté dans la Figure 13 (droite). Notre évaluation portait sur six tâches de lisibilité du réseau (dont certaines spécifiques à l'analyse de réseaux sociaux) pour deux réseaux de densités différentes. Les résultats ont démontré que les duplications permettaient de significativement améliorer la lisibilité de représentation et d'offrir une vue non biaisée de chaque communauté.

Présenter l'information

Le grand avantage de NodeTrix est sa versatilité : il est possible de l'utiliser pour explorer un réseau et découvrir sa structure mais aussi pour en faire une représentation très compacte et l'utiliser pour présenter des résultats. Un exemple est présenté Figure 16.

⁴³ HENRY *et al.*, 2008

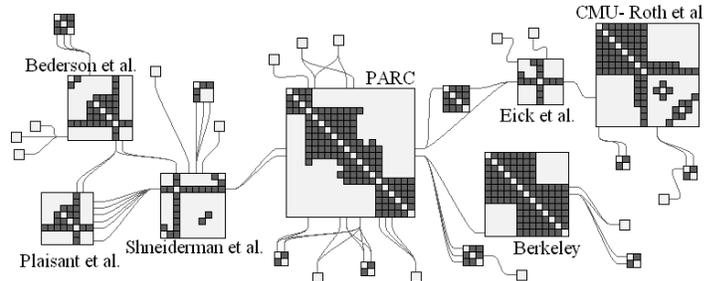


Figure 16 : Représentation NodeTrix compacte pour la communication (réseau de co-publication d'InfoVis contenant plus d'une centaine d'acteurs)

Cette souplesse d'utilisation est liée à l'intégration des représentations matricielles et à leur capacité à produire à la fois une vue d'ensemble et une vue de détail. En mode présentation des résultats, il est alors possible de réduire au maximum la taille des communautés représentée par des matrices, faisant disparaître les étiquettes individuelles de chacun des membres (mais en conservant une étiquette pour la communauté par exemple) mais permettant de conserver un aperçu de la connectivité à l'intérieur de la communauté. La Figure 15 et la Figure 16 sont un bon exemple, toutes deux représentent le même réseau, à des niveaux de détail différents.

Conclusion

Les représentations visuelles sont importantes à la fois pour l'analyse des réseaux sociaux et la communication des résultats d'analyse sur ces réseaux. Actuellement, la très grande majorité des représentations de réseaux sont des diagrammes nœud-lien. Si ces représentations sont familières et paraissent très intuitives, elles souffrent d'importants problèmes de lisibilité lorsque le réseau comporte un grand nombre de nœuds ou de liens. Dans cet article, nous avons présenté plusieurs représentations alternatives permettant de résoudre ces problèmes de lisibilités.

En particulier, nous nous sommes concentrés sur les représentations basées sur la matrice d'adjacence de graphe. Cette représentation présente plusieurs

avantages par rapport aux nœud-lien, le plus important étant de supprimer toute superposition de nœuds et croisements de liens ce qui est en particulier utile pour les réseaux de forte densité. Ces représentations ont aussi plusieurs inconvénients comparées aux diagrammes nœud-lien : il est notamment très laborieux d'utiliser les matrices pour identifier des chemins entre plusieurs nœuds.

Combinant les avantages des deux représentations, nous avons présenté MatrixExplorer, un système interactif permettant d'explorer les réseaux sociaux en manipulant matrices et diagrammes nœud-lien pour avoir de multiples perspectives sur les données. Nous avons également présenté deux représentations hybrides, fusionnant les deux représentations : MatLink, permettant d'améliorer significativement les tâches de suivi de chemin dans les matrices ; et NodeTrix, particulièrement efficace pour visualiser les réseaux petit-monde.

Ces représentations offrent de nouvelles perspectives sur la représentation visuelle des réseaux sociaux, permettant à la fois de les explorer et constituant également une solution pour communiquer de l'information et présenter des résultats d'analyse. Ils constituent des solutions à l'analyse visuelle des nouveaux réseaux sociaux en ligne.

BIBLIOGRAPHIE

ABELLO, J. AND F. VAN HAM (2004). Matrix zoom : A visual interface to semi-external graphs. In *Proceedings of the 2004 IEEE Symposium on Information Visualization* (INFOVIS'04), Austin, Texas, pp. 183-190. IEEE Computer Society.

ABELLO, J., F. VAN HAM, AND N. KRISHNAN (2006). Ask-graphview : A large scale graph visualization system. *IEEE TVCG journal* 12 (5), 669-676.

AHLBERG, C., C. WILLIAMSON, AND B. SHNEIDERMAN (1992). Dynamic queries for information exploration : An implementation and evaluation. In *Proceedings of the ACM CHI'92 Conference on Human Factors in Computing Systems*, New York, NY, USA, pp. 619-626. ACM Press.

AUBER, D. (2003). Tulip : A huge graph visualisation framework. In *Graph Drawing Software*, pp. 105-126. Springer-Verlag.

BERTIN, J. (1967). Sémiologie graphique : Les diagrammes - Les réseaux - Les cartes (Les réimpressions ed.). Paris, France : Editions de l'Ecole des Hautes Etudes en Sciences.

CARAU, G. AND S. PINLOCHE (2005). Permutmatrix : A graphical environment to arrange gene expression profiles in optimal linear order. *Bioinformatics* 21, 1280-1281.

CARPENDALE, M. S. T. (1999). RA Framework for Elastic Presentation Space. Ph. D. thesis, Simon Fraser University.

COCKBURN, A. AND B. MCKENZIE (2000). An evaluation of cone trees. *People and Computers XV (Proceedings of the 2000 British Computer Society Conference on HumanComputer Interaction)*.

COCKBURN, A. AND B. MCKENZIE (2002). Evaluating the effectiveness of spatial memory in 2d and 3d physical and virtual environments.

COCKBURN, A. AND B. J. MCKENZIE (2001). 3D or not 3D ? evaluating the effect of the third dimension in a document management system. In *Proceedings of SIG CHI 2001*, pp. 434-441.

DI BATTISTA, G., P. EADES, R. TAMASSIA, AND I. G. TOLLIS (1998). *Graph Drawing : Algorithms for the Visualization of Graphs*. Prentice Hall PTR. 14

ELMQVIST, N., N. HENRY, Y. RICHE, AND J.-D. FEKETE (2008). Mélange : space-folding for multi-focus interaction. In *CHI '08 : Proceedings of the SIGCHI conference*

on *Human Factors in computing systems*, New York, NY, USA. ACM.

ELMQVIST, N. AND P. TSIGAS (2007). A taxonomy of 3d occlusion management techniques. In *Proceedings of the IEEE Conference on Virtual Reality 2007*, pp. 51-58.

FEKETE, J.-D., D. WANG, N. DANG, AND C. PLAISANT (2003, October). Overlaying graph links on treemaps. *IEEE Symposium on Information Visualization Conference Compendium* (demonstration).

FRANK, O. (1978). Sampling and estimation in large social networks, *social networks*, 1, 91-101.

FREEMAN, L. C. (2000). Visualizing social networks. *Journal of social structure* 1 (1).

GHONIEM, M., J.-D. FEKETE, AND P. CASTAGLIOLA (2005). On the readability of graphs using node-link and matrix-based representations : a controlled experiment and statistical analysis. *Information Visualization* 4 (2), 114-135.

HEER, J. AND D. BOYD (2005). Vizster : Visualizing Online Social Networks. In *Proceedings of the IEEE Symposium on Information Visualization*, pp. 5.

HENRY, N., A. BEZERIANOS, AND J.-D. FEKETE (2008). Improving the readability of clustered social networks by node duplication. In *IEEE Transactions on Visualization and Computer Graphics (Proceedings of Visualization / Information Visualization 2008)*, Volume 14.

HENRY, N. AND J.-D. FEKETE (2006a). MatrixExplorer : a Dual-Representation System to Explore Social Networks. *IEEE TVCG (Infovis'06 proceedings)* 12 (5), 677-684.

HENRY, N. AND J.-D. FEKETE (2007). Matlink : Enhanced matrix visualization for analyzing social networks. In *Lecture Notes in Computer Science (Proceedings of the 13th IFIP TC13 International Conference on Human-Computer Interaction, INTERACT'07)*, Volume 4663, pp. 288-302.

HENRY, N. AND J.-D. FEKETE (September 2006b). Matrixexplorer : Un système pour l'analyse exploratoire de réseaux sociaux. *Proceedings of IHM2006*, International Conference Proceedings Series, 67-74.

HENRY, N., J.-D. FEKETE, AND M. MCGUFFIN (2007). Nodetrix : Hybrid representation for analyzing social networks. *TVCG (Proceedings of IEEE Information Visualization conference)* 13 (6), 1302-1309.

HERMAN, I., G. MELANÇON, M. M. DE RUITER, AND M. DELEST (1999). Latour | A tree visualisation system. *Proceedings of the Graph Drawing symposium (GD'99)*,

LNCS 1731, 392-399.

HERMAN, I., G. MELANÇON, AND S. MARSHALL (2000). Graph visualization and navigation in information visualization : A survey. *IEEE TVCG* 6 (1), 24-43.

HUISMAN, M. AND M. A. J. DUIJN (2005). Software for social network analysis (chapter 13). In V. B. W. de Nooy, A. Mrvar (Ed.), *Exploratory Social Network Analysis with Pajek*.

JAIN, A. K., M. N. MURTY, AND P. J. FLYNN (1999). Data clustering : a review. *ACM Computing Surveys* 31 (3), 264-323.

LAMPING, J. AND R. RAO (1996). The Hyperbolic Browser : A focus + context technique for visualizing large hierarchies. *Journal of Visual Languages and Computing* 7 (1), 33-35.

LEE, B., C. S. PARR, C. PLAISANT, B. B. BEDERSON, V. D. VEKSLER, W. D. GRAY, AND C. KOTFILA (2006). Treeplus : Interactive exploration of networks with enhanced tree layouts. *IEEE TVCG (Infovis'06 proceedings)* 12 (6), 1414-1426.

MORENO, J. L. (1934). Who shall survive ?

MUNZNER, T. (1997). H3 : Laying out large directed graphs in 3d hyperbolic space. *Symposium on Information Visualization (InfoVis '97)*, 2-10.

PERER, A. AND B. SHNEIDERMAN (2006). Balancing Systematic and Flexible Exploration of Social Networks. *IEEE TVCG (Infovis'06 proceedings)* 12 (5), 693-700.

RISDEN, K., M. CZERWINSKI, T. MUNZNER, AND D. COOK (2000). An initial examination of ease of use for 2D and 3D information visualizations of web content. *International Journal of Human Computer Studies*, 53 (5), 695-714.

SHNEIDERMAN, B. (1983). Direct manipulation : A step beyond programming languages. *IEEE Computer*, 16 (8), 57-69.

SHNEIDERMAN, B. (1992). Tree visualization with tree-maps : A 2-D space-filling approach. *ACM Transactions on Graphics* 11 (1), 92-99.

SHNEIDERMAN, B. (1996). The Eyes Have It : A Task by Data Taxonomy for Information Visualization. *Visual Languages*, 336-343.

SHNEIDERMAN, B. AND A. ARIS (2006). Network visualization by semantic substrates. In *IEEE TVCG (Infovis'06 proceedings)*, 12 (5).

SUTCLIFFE, A. G. AND U. PATER (1996). 3D or not 3D : is it nobler to the mind ? In

British Human Computer Interaction Conference, pp. 79-94. Cambridge University Press.

TUKEY, J. (1977). *Exploratory Data Analysis*. Addison-Wesley.

VAN HAM, F. (2003). Using multilevel call matrices in large software projects. In *Proceedings of the 2003 IEEE Symposium on Information Visualization*, Seattle, WA, USA, pp. 227-232. IEEE Press.

WASSERMAN, S. AND K. FAUST (1994). *Social Network Analysis*. Cambridge Univ. Press.

WATTS, D. J. AND S. H. STROGATZ (1998). Collective dynamics of “small-world” networks. *Nature* 393, 440 - 442.

ZIEGLER, J., C. KUNZ, AND V. BOTSCH (2002). Matrix browser visualizing and exploring large networked information spaces. In *Extended Abstracts of Human Factors in Computing Systems (CHI'02)*, pp. 602-603.