

Apports de l'analyse des réseaux d'alliances stratégiques en intelligence économie

Bernard Dousset*, Brigitte Gay**

* Institut de Recherche en Informatique de Toulouse, IRIT-SIG
Université Paul Sabatier, 118, Route de Narbonne, F-31062 Toulouse Cedex 9
dousset@irit.fr

** ESC Toulouse, 20, boulevard Lascrosses, BP 7010, F-31068 Toulouse Cedex 7
b.gay@esc-toulouse.fr

Mots clefs :

Recherche d'informations, fouille de données, analyse de données, réseau, innovation, alliances, centralité, cartographie, position, structure, stratégie

Keywords:

Information retrieval, datamining, data analysis, network, innovation, alliances, centrality, mapping, position, structure, strategy

Palabras clave :

Recuperación de datos, datamining, análisis datos, red, innovación, alianzas, carácter central, cartografía, posición, estructura, estrategia

Résumé :

Du fait de l'ubiquité des alliances stratégiques, les entreprises seront toujours plus dépendantes de leur capacité à créer et à gérer des alliances. Pourtant, très peu d'attention a pour l'heure été consacrée à la structure des réseaux et aux configurations des interactions entre des sous-groupes de firmes. Des termes comme 'l'entreprise réseau' ou 'la firme virtuelle' sont de plus en plus utilisés pour décrire des formes organisationnelles comprenant des réseaux d'entreprises. Le concept d'encastrement ("embeddedness") de l'entreprise dans des réseaux 'macro' complexes a été introduit. La compréhension des dynamiques de formation des réseaux et des interdépendances entre l'acteur/entreprise individuel et le réseau global est vue comme un aspect crucial de la stratégie. En utilisant pour exemple un segment important de l'industrie des biotechnologies, nous montrons l'importance de développer des logiciels et outils permettant l'analyse de la dynamique de ces formes organisationnelles en émergence.

Abstract :

Because of the ubiquity of strategic alliances, companies will increasingly be dependant on their ability to create and manage alliances. Very little attention has yet been given to network structure, the pattern of interactions between subsets of firms. Terms such as 'the networked firm' or 'the virtual organization' have been increasingly used to describe an organizational form including a network of firms. Also, the concept of embeddedness, of a firm in a complex macro network structure, has been introduced. Understanding the dynamics of network formation and the interdependence between the individual actor/firm and the overall network is perceived as a crucial aspect of strategy. Using as case study an important segment of the biotechnology industry, we demonstrate here the importance of developing software and tools that allow the analysis of the dynamics of these emerging organizational forms.

1 Introduction

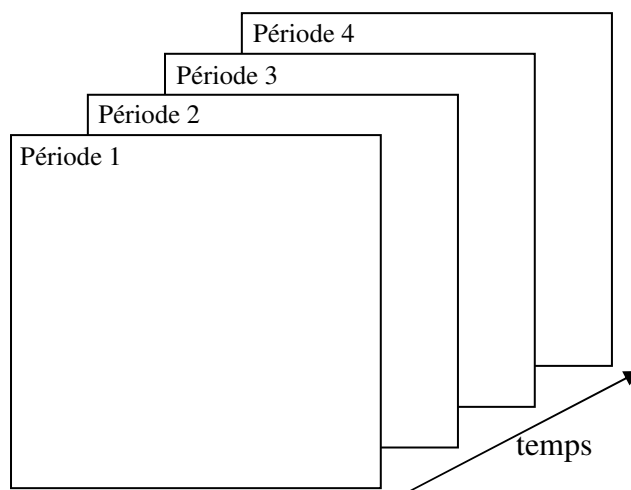
En complément de notre plate-forme bibliométrique Tétralogie, nous avons été amenés, dans le cadre d'analyses relationnelles orientées intelligence économique, à nous intéresser aux dessins de graphes d'alliances et plus particulièrement à l'étude de leur évolution sur une période donnée. Dans un premier temps nous nous sommes focalisés sur la comparaison des graphes issus d'une partition de la période en intervalles réguliers : années, groupes d'années successives ou autres [5. Dousset & Karouach 2005]. Mais la nature de certaines relations rencontrées au cours de nos analyses stratégiques n'est pas correctement modélisée par ce type de discrétisation rudimentaire. La persistance d'une relation attribuée à une époque particulière peut parfaitement déborder en aval et donc contribuer à l'établissement d'une structure pertinente dont l'édification s'étale sur plusieurs segments de temps consécutifs. La visualisation des différentes instances d'un graphe, ne permet pas de détecter à coup sûr ce type de structure souvent temporaire, celle du graphe statique non plus, car il est beaucoup plus complexe et donc difficile à visualiser (cumul de l'ensemble des instances).

Nous présentons, dans cette communication, un certain nombre de propositions afin de remédier à l'inconvénient de la partition stricte du temps et nous illustrons, par plusieurs exemples, l'utilité de prendre en compte l'effet mémoire dans l'analyse relationnelle. La source principale de nos exemples sera l'étude des alliances interentreprises dans le domaine des biotechnologies (joint venture, licences, acquisitions, participations,...).

2 Méthode

2.1 Principe général de l'étude de l'évolution

L'étude, sur plusieurs périodes, de l'évolution d'un réseau d'acteurs ou d'un réseau sémantique nécessite de décomposer la matrice correspondant à son graphe en plusieurs instances. Nous obtenons alors une matrice 3D dont chaque plan correspond à une des périodes choisies.



Une première approche de l'étude de l'évolution peut être apportée par les visualisations séparées ou enchaînées des différentes instances. Mais tout dépend de la nature des liens qui sont pris en compte. Lorsqu'il s'agit de liens temporaires comme par exemple des collaborations matérialisées par des co-signatures d'articles scientifiques ou des participations à des contrats de recherche, chaque instance prise séparément a un sens et l'on peut ainsi mettre en évidence des rapprochements, des fusions ou des cissions entre équipes de recherche. A l'opposé, pour un graphe sémantique, les liens anciens entre les termes subsistent souvent et le réseau existant est simplement complété par l'ajout de nouveaux termes et de nouveaux liens. On ne doit donc pas perdre l'essentiel de l'ancienne structure. Ici les graphes s'ajoutent plus qu'ils ne se remplacent au cours du temps : on parle d'intégration dans le temps. Une troisième situation, intermédiaire entre les deux précédentes, nous est apparue lors de

l'analyse des alliances interentreprises. En effet, une alliance a une certaine durée de vie ou plutôt une persistance et aucun des modèles précédents ne convient, puisqu'il faut tenir compte du passé récent et très peu ou pas d'un passé plus ancien. Il faut donc simuler un effet mémoire, si possible paramétrable, afin de vérifier si les hypothèses émises sur la structuration du domaine sont robustes vis-à-vis des différents réglages possibles.

Nous proposons une méthode générique d'association des différentes instances afin de prendre en compte cette problématique, méthode qui doit couvrir les trois cas envisagés : enchaînement, intégration, effet mémoire paramétrable.

2.2 Enchaînement des instances d'un graphe

Cette méthode consiste à remplacer une instance par la suivante. On saute donc d'une période à l'autre. Mais en vue d'augmenter la lisibilité de l'ensemble, les sommets du graphe restent fixes et leur position est définie par un algorithme de tracé de graphe qui prend en compte la totalité des instances, c'est-à-dire le graphe statique portant sur l'ensemble des données utilisées. Deux sommets qui seront liés à un moment ou à un autre ne seront donc pas trop éloignés l'un de l'autre, idem pour des regroupements. Cette approche marche assez bien pour des graphes de petite taille ou de taille moyenne (quelques centaines de sommets). Pour des graphes plus grands, une décomposition préalable en classes homogènes et un tri du graphe réduit permettent d'améliorer sensiblement, la situation. En voici une illustration sur deux périodes pour un réseau d'auteurs scientifiques (co-signatures).

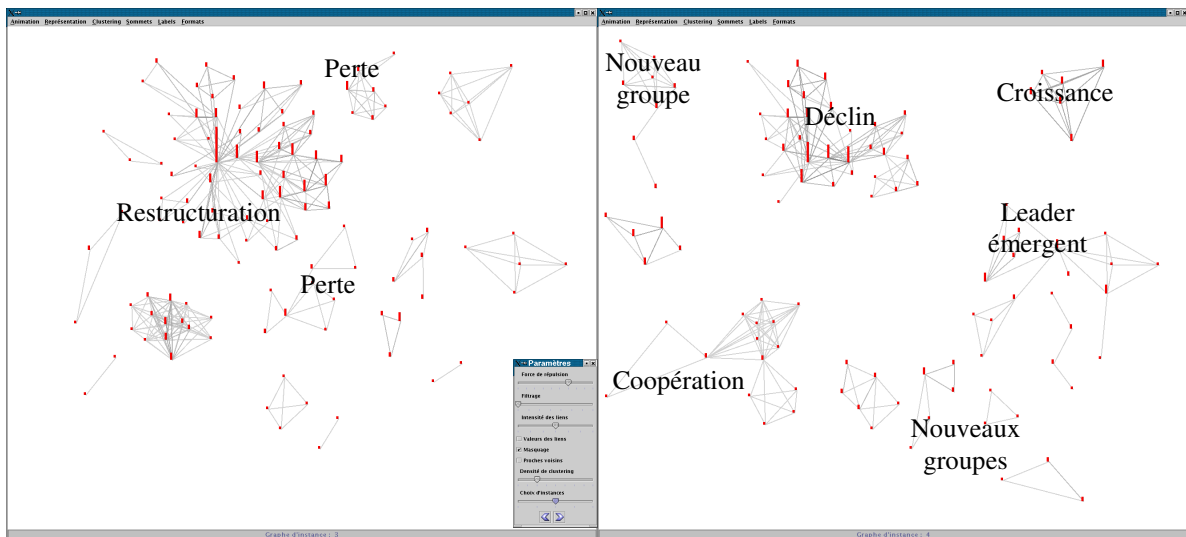


Figure 1 : Evolution d'un réseau d'auteurs entre deux périodes successives

2.3 Intégration des instances d'un graphe

Dans ce cas, les instances se cumulent afin d'aboutir au graphe statique obtenu sans faire intervenir la décomposition en plusieurs périodes. L'intérêt est de voir se constituer le graphe final à partir d'un noyau initial et d'apports successifs soit au niveau des sommets, soit au niveau des liens (apparition d'un lien ou évolution de son intensité). Cette approche convient bien aux réseaux intégratifs comme les réseaux sémantiques, les réseaux de communication, l'apprentissage. L'illustration suivante montre la naissance d'un signal sémantique (signal faible au début, véritable sujet vers la fin).

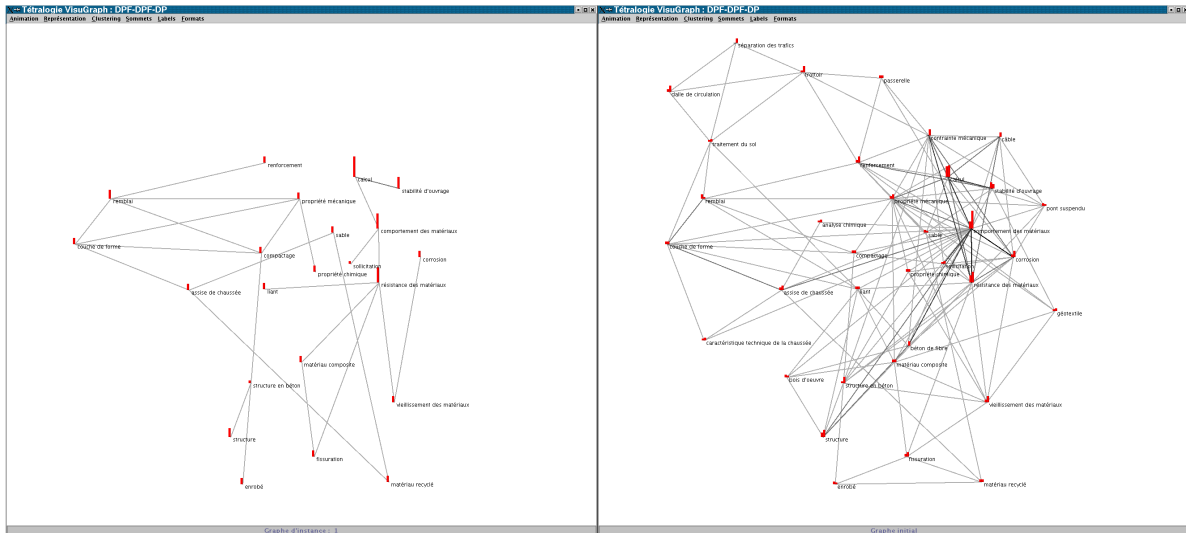


Figure 2 : Réseau sémantique émergent : signal faible -> signal fort.

2.4 Effet mémoire paramétrable

Ici, nous essayons de modéliser, par un même outil, les deux cas précédents et divers cas intermédiaires qui font intervenir la mémoire du système avec plus ou moins d'importance.

Si les instances sont équiréparties dans le temps (ce qui n'est pas toujours le cas) nous pouvons simuler nos trois cas de la façon suivante :

- **Graphe statique**

Toutes les instances sont alors cumulées, le graphe correspondant tient donc compte de l'ensemble des données disponibles.

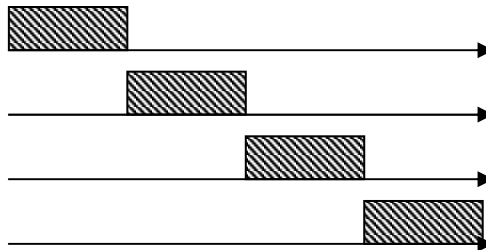
- **Une instance**

Une instance représente la limitation de l'analyse relationnelle à une période donnée (année, groupe d'années, ou toute autre période définissable à partir de données datées).



- **Enchaînement des instances d'un graphe**

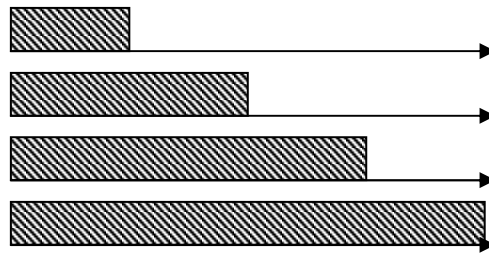
La première méthode consiste à enchaîner les instances du graphe sans tenir compte d'un éventuel effet mémoire. Nous considérons donc que les relations n'ont de valeur que de façon instantanée, elles ne subsistent pas au-delà de la période correspondante. Le lien n'est donc pas considéré comme durable : liaison passagère.



- **Intégration des instances d'un graphe**

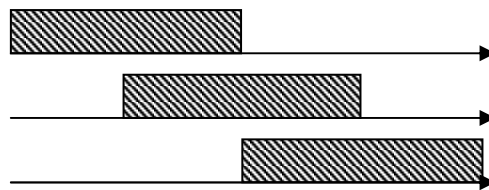
Cette seconde méthode est tout à fait l'opposée de la précédente, les liaisons sont supposées définitives, elles restent donc actives dans les périodes qui suivent leur apparition. On peut parler de phénomène d'intégration, la dernière instance correspondant au graphe statique. La visualisation de

l'évolution permet de savoir comment s'est construit le graphe final (par exemple un réseau de transport).



- **Effet mémoire paramétrable (exemple sur deux périodes glissantes)**

Ici, seule la période précédente est prise en compte, les autres sont ignorées. Nous pouvons remarquer qu'il n'y a plus, dans ce cas, que trois instances différentes, la moyenne de la dernière période est donc moins « récente » que dans le cas de l'enchaînement des instances, ceci peut avoir des répercussions sur les capacités prospectives de ce type d'étude.



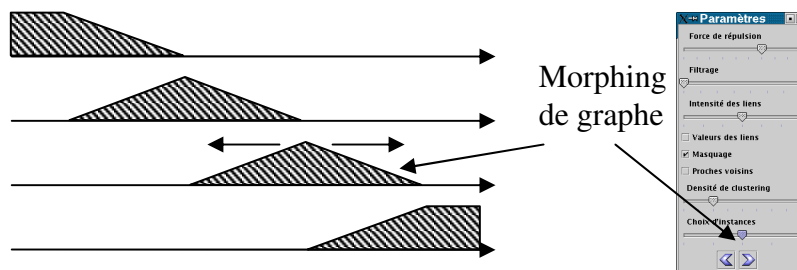
- **Courbe de persistance d'un lien**

Un dernier cas peut être envisagé lorsque nous connaissons la date exacte de chaque lien et non pas seulement son année ou sa période. Une courbe de persistance peut alors être utilisée (modélisant l'usure de la relation) et il est même alors possible de visualiser en continu l'évolution du graphe. Malheureusement peu d'informations sont aussi précisément datées, mais dans les cas de l'étude des alliances il est possible de travailler ainsi.



2.5 Simulation d'une évolution en continu

Afin de suggérer, à l'utilisateur, la dynamique du réseau étudié, nous pouvons simuler une évolution en continu du graphe par un passage progressif d'une instance à l'autre. Pour cela des fonctions « chapeau » peuvent être utilisées ce qui va permettre à la fois de simuler l'apparition ou la disparition de sommets ou d'arêtes ainsi que le renforcement ou la diminution de certains liens. Le but final de notre étude étant de proposer un couplage de ce principe avec un tracé optimisé pour faire converger ou diverger les sommets en fonction de leurs liaisons temporaires (morphing de graphe), mais de nombreuses difficultés demeurent.



Dans la figure précédente, nous pouvons remarquer que le graphe n'évolue pas pendant une demi-période au début et à la fin. Entre ces deux extrêmes, nous associons dans des proportions variables deux périodes successives pour recréer des instances intermédiaires en nombre suffisant pour donner une impression de mouvement continu. Cette impression est renforcée si les icônes représentant les sommets sont eux aussi soumis à une évolution de leur taille. La taille d'un sommet étant par exemple proportionnelle à la valeur cumulée de ses liens ou à une mesure qui lui est propre comme le nombre de publications pour les co-signatures ou la fréquence pour les mots associés. Comme dans la figure 1, l'icône peut aussi servir à visualiser, sous forme d'histogramme, l'évolution de l'importance du nœud sur l'ensemble des périodes.

Nous illustrons, ci-dessous, cette méthode en visualisant un graphe de départ (principale classe connexe de la première instance : période 2000-2001), une position intermédiaire, et le graphe d'arrivée (seconde instance : période 2002-2003). Les arêtes qui ne sont présentes que dans la première période s'estompent peu à peu, celles qui ne sont présentes que dans la seconde apparaissent progressivement, les autres arêtes s'adaptent au niveau de leur épaisseur et, de leur côté, les sommets apparaissent, disparaissent ou migrent en fonction de l'évolution des liens qui les unissent. Ce processus peut, bien entendu, être utilisé pour enchaîner un nombre quelconque d'instances d'un même graphe en utilisant, éventuellement, l'effet mémoire.

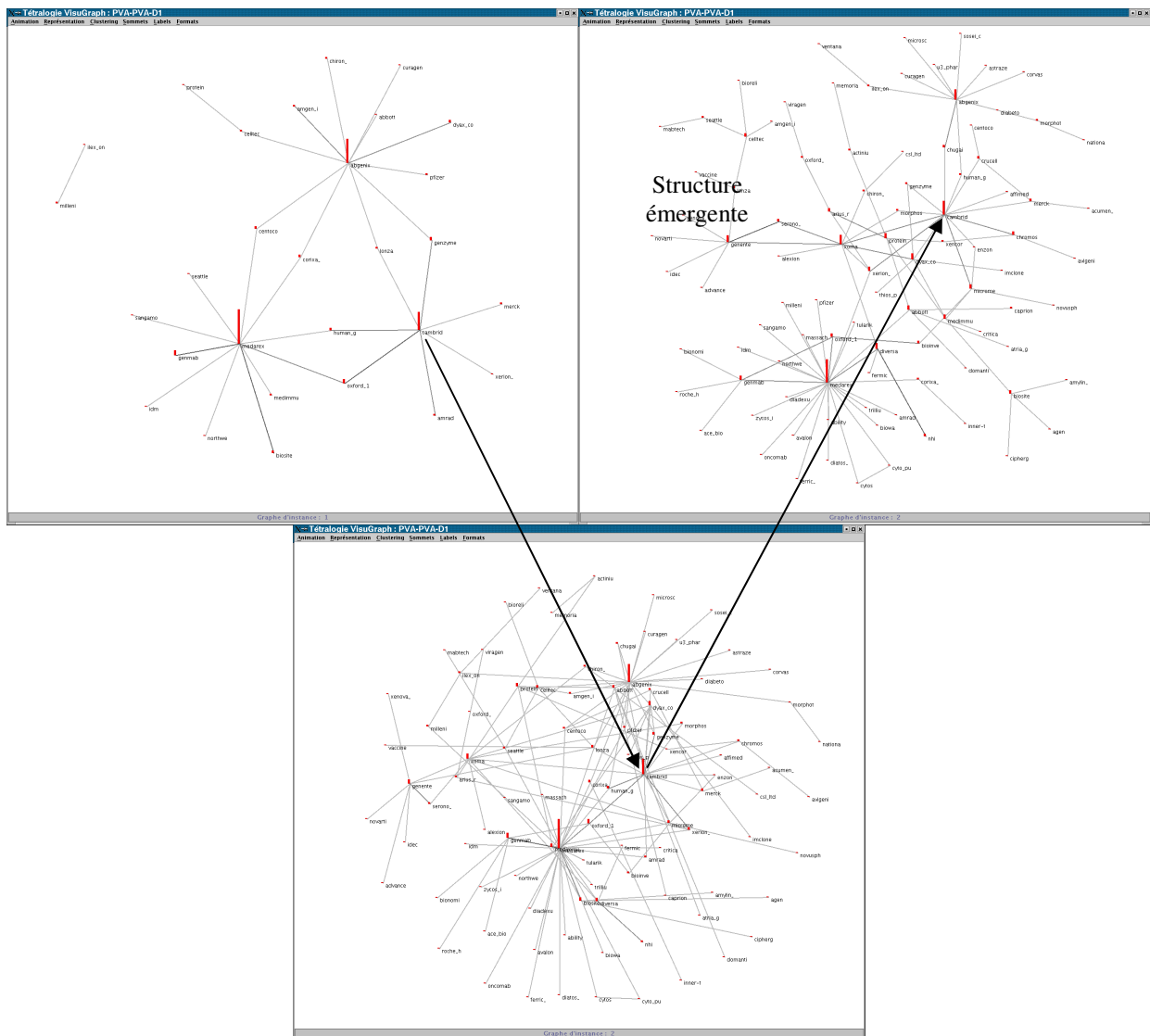


Figure 3 : Morphing de graphe entre deux périodes consécutives.

3 Expérimentation

3.1 Graphe d'alliances

Le terme de réseau est utilisé dans la recherche sur les alliances pour décrire les relations inter-entreprises. D'un point de vue méthodologique, le nœud est alors l'entreprise et le lien entre deux entreprises représente l'alliance. D'un point de vue économique, stratégique, sociologique, socio-économique, etc, le terme de réseau a été utilisé de manière assez imprécise, nommant parfois la dyade seule, parfois l'acteur focal et son réseau égocentrique (réseau de niveau 2 qui décrit les interactions de l'acteur focal avec ses partenaires et les interactions de ses partenaires entre eux) et plus généralement de son voisinage proche (mesure de la contrainte, de la redondance de cliques, etc). Les résultats des différentes recherches apparaissent donc souvent contradictoires ou sont potentiellement trompeurs.

Depuis cinq ans environ, les études plus récentes, qui s'intéressent non pas à un acteur mais à l'ensemble des acteurs et de leurs liens dans des réseaux complexes (World Wide Web), sont le fait de physiciens et de mathématiciens [14. Newmann 2003] et s'intéressent aux propriétés génériques de réseaux macros et à leur modélisation. Une perspective « réseau » en recherche sur les alliances implique en fait une analyse non seulement des liens au voisinage de l'entreprise, mais aussi celle de l'ensemble des liens formés par toutes les entreprises dans une industrie ou un secteur industriel. En effet, l'entreprise décide d'interactions spécifiques avec des partenaires industriels donnés à des fins stratégiques. Dans l'industrie des biotechnologies, l'entreprise s'appuie sur sa compétence idiosyncratique (scientifique, technologique) ou financière (dans ce cas, elle accède essentiellement à des ressources technologiques, scientifiques, ou des produits) pour former des alliances [10. Gay & Dousset 2005]. Par ailleurs, l'ensemble des liens formés par l'ensemble des entreprises dans un secteur donné donne une « image » de ce secteur à un instant t : l'ensemble des entreprises qui ont investi dans ce secteur, comment elles interagissent entre elles, et celles qui prennent une position centrale. Bien qu'elles soient essentielles, il y a eu encore peu d'études sur les réseaux d'alliances, leur dynamique et la position évolutive de l'entreprise dans ces réseaux. Nous avons déjà démontré que les positions les plus centrales étaient obtenues par des acteurs possédant des technologies clés et que ces positions n'étaient pas pérennes mais suivaient plutôt les différents cycles technologiques à l'intérieur d'un secteur.

Dans la figure 4, la prise de pouvoir par les acteurs A, C se fait de 1997 à 2001 puis décroît. Par contre l'acteur B démarre plus tardivement et maintient sa centralité sur un réseau toutefois de plus en plus cohésif à son voisinage. Il est donc important d'essayer de détecter le plus vite possible la centralité émergente de nouveaux entrants dans un secteur de manière à pouvoir à la fois détecter l'amorce de nouvelles phases technologiques et à investir éventuellement dans ces acteurs en devenir. La prise en compte de l'effet « mémoire » est dans ce cas utile parce qu'elle introduit suffisamment de redondance dans les données (chevauchement sur un an ou plus) pour percevoir la centralité émergente de nouveaux entrants et leur dynamique de croissance, sans donner une intégration des données trop longue. On arrive ainsi à la captation des nouvelles phases technologiques à partir de la cohérence des structures émergentes.

Dans le cadre de l'étude des alliances interentreprises dans le domaine des biotechnologies, nous avons pris pour exemple la principale classe connexe concernant la période 1997 à début 2004. Cette classe comprend initialement 201 entreprises. Dans un premier temps les branches de cette classe connexe sont élaguées afin de réaliser une étude macro (obtention d'un graphe simplifié de 70 sommets). Nous avons divisé le temps en 6 périodes recouvrantes 1997-98, 1998-99 à 2002-03 afin de simuler une persistance des accords sur 2 ans. Le graphe de cette classe connexe correspondant à la période de référence (1997-2003) a tout d'abord été décomposé en clusters homogènes par une méthode de Markov Clustering [13. Karouach 2003], le graphe réduit a ensuite été dessiné en utilisant un algorithme basé sur la notion d'attraction et de répulsion afin de répartir au mieux les classes dans le plan. Un sommet de chaque classe est alors fixé et le graphe de départ est redessiné en utilisant le même algorithme de forces. Une intervention manuelle est encore nécessaire pour éviter certains chevauchements de libellés. Cette démarche a aussi été appliquée au graphe total de 201 sommets afin d'affiner l'analyse et de valider nos premières intuitions.

3.2 Graphe simplifié

Afin de faire ressortir la notion de réseau, nous pouvons simplifier le graphe initial en élaguant ses branches de façon récursive. Dans un premier temps les sommets qui n'ont qu'une connexion sont éliminés, cette action est répétée jusqu'à l'obtention d'un graphe stationnaire. Il ne reste alors, dans le graphe, que des sommets servant de connecteurs complexes (inclus dans au moins un cycle). Dans notre exemple, nous n'avons plus que 70 sommets sur 201, la lisibilité est donc largement améliorée et les connecteurs qui ont une importance stratégique apparaissent clairement.

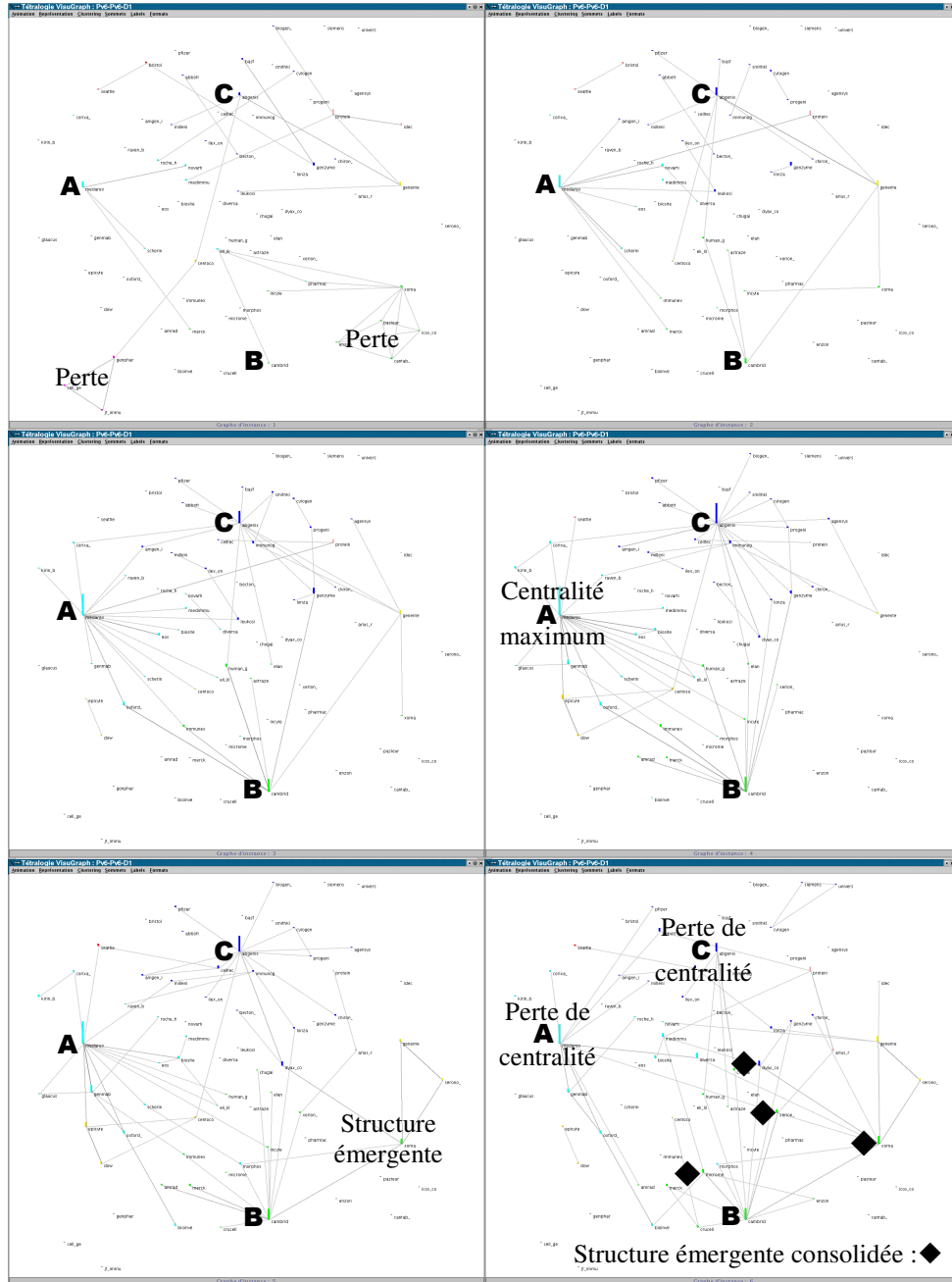


Figure 4 : Instances recouvrantes du graphe sans ses branches (70 sommets).

3.3 Graphe entier

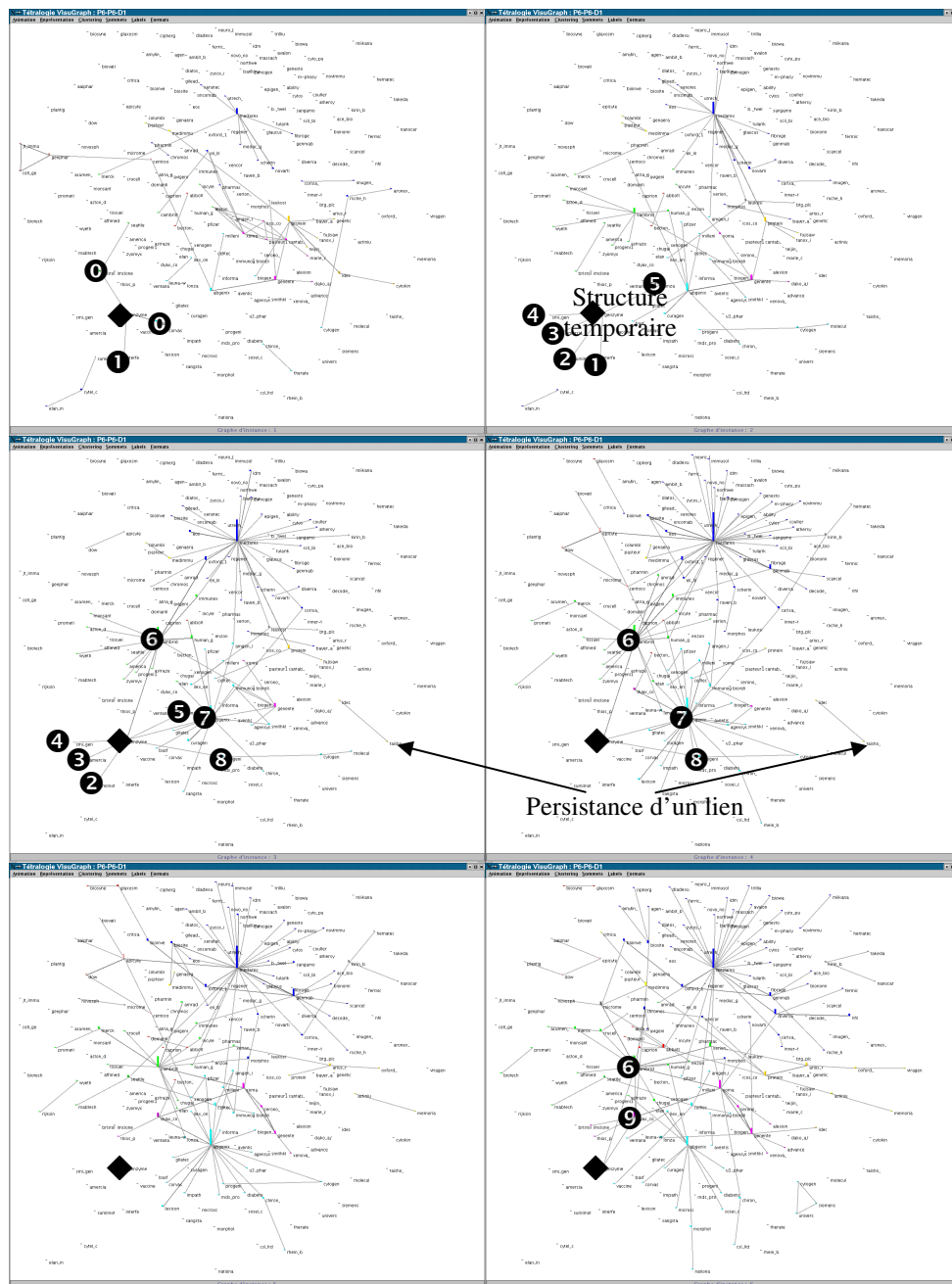


Figure 5 : Graphe des différentes instances recouvrantes

L'analyse du graphe total, quant à elle, permet d'évaluer l'encastrement de l'acteur dans le réseau macro et la pertinence stratégique de sa position et de son réseau égocentrique dans la structure globale. L'effet mémoire permet de capter la cohésion (situation de dépendance) en même temps qu'il contribue à mesurer la centralité effective de l'acteur (nombre de liens directs non redondants) qui, elle, signale l'autonomie, le pouvoir de l'acteur, et sa capacité à contrôler le réseau. Par exemple dans la figure 5, l'acteur désigné par un losange possède trois liens non redondants au début de l'analyse, mais n'appartient pas à la classe connexe principale. Sa centralité effective augmente ensuite et permet son intégration dans le réseau connexe, notamment à travers deux acteurs centraux (6 et 7 dans le graphe). Sa centralité effective disparaît progressivement. En avant dernière période l'acteur est totalement déconnecté, et se reconnecte en dernière période à un acteur émergent (9) et un acteur central (6). L'analyse indique une position forte dès la deuxième période, puis une perte rapide de pouvoir qui

n'est pas encore rattrapée en dernière période, malgré l'accès à de nouvelles technologies au travers des acteurs centraux 6, 7 et 9. Nous pouvons constater que les caractéristiques précédentes sont beaucoup plus délicates ou même impossibles à découvrir à partir du graphe statique présenté ci-dessous, qui cumule les informations de l'ensemble de la période étudiée (1997-2003). D'où l'intérêt de ce type d'approche pour la compréhension de phénomènes complexes liés à l'intelligence économique.

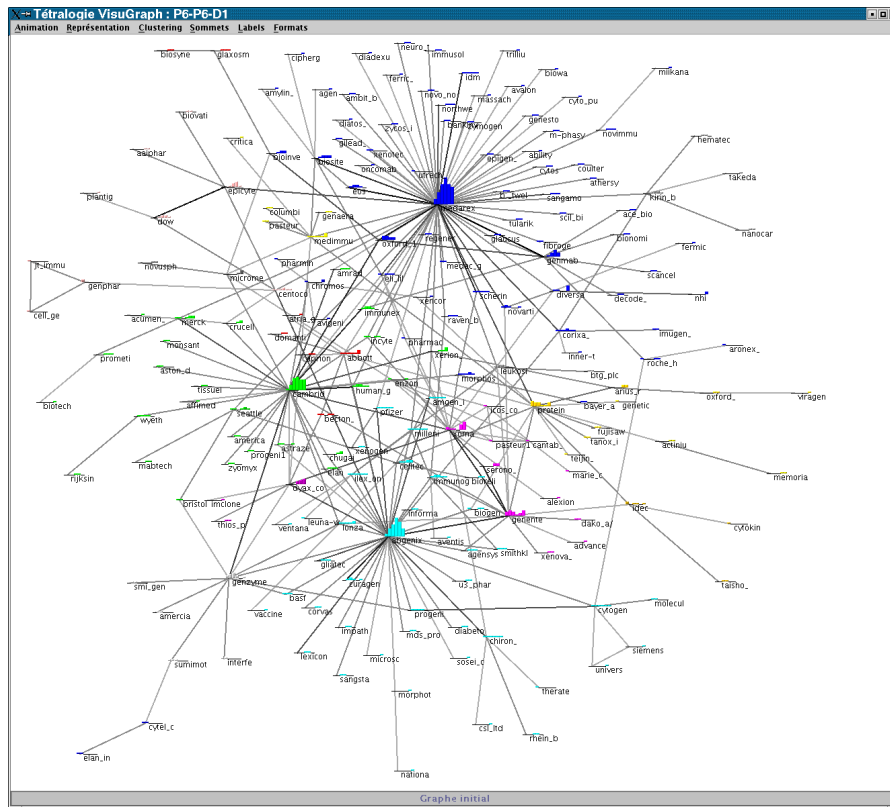


Figure 6 : Graphe de la principale classe connexe pour la période 1997-2003

4 Conclusion

Après avoir apporté des solutions pour l'étude et la visualisation des graphes de grande taille (clustering, graphe partiel, élagage, ...) ainsi que pour l'étude de l'évolution des graphes sur une partition du temps, nous venons de proposer une réflexion sur la façon de discrétiser une période afin de prendre en compte les effets mémoire en analyse relationnelle évolutive qui est un des piliers de l'intelligence économique. Les expérimentations que nous avons conduites ont été réalisées sur des outils existants qui n'ont pas été initialement prévus pour ce type d'utilisation. Mais la pertinence de certaines découvertes de structures pertinentes sur le plan stratégique nous encourage à adapter ces outils en conséquence. Bien entendu le partitionnement restera la méthode la plus utilisée, il n'y a qu'en présence de liens persistants ou intégrables que se pose le problème de l'effet mémoire. Dans certains cas la taille des périodes peut être adaptée à la nature des données (deux ans par période pour les publications scientifiques sont préférables à un an car il y a des oscillations entre année de recherche et année de publication, de plus certains grands congrès sont souvent bisannuels). Mais cette adaptation ne suffit pas si la structure à mettre en évidence est le résultat d'une stratégie de mise en place à moyen ou à long terme. Un autre problème est l'étude de la robustesse des résultats d'une analyse stratégique en fonction du réglage de l'effet mémoire : il faut arriver à apprécier ce qu'apporte la prise en compte de cet effet à la compréhension des stratégies sous-jacentes. La pertinence d'une étude peut être tributaire de la prise en compte de ces phénomènes, notamment si la structure finale est le résultat de plusieurs étapes successives liées aux règles de l'art du domaine étudié.

5 Bibliographie

- [1.] Carmel, C., Harel, D., Koren, Y., *Combining Hierarchy and Energy Drawing Directed Graphs*, (2004), IEEE Transactions on Visualisation and Computer Graphics 10 (1), p. 46-57.
- [2.] Chen C., Carr L., *Visualizing the evolution of subject domain: A case study*, (1999), InfoVis'99, IEEE Computer Society Press, 1999, p. 449-452.
- [3.] Collberg C., Kobourov S. G., Nagra J., Pitts J., and Wampler K., *A system for graph-based visualization of the evolution of software*, (2003), ACM Symposium on Software Visualization.
- [4.] Doz, Y. L., Hamel, G., *Alliance advantage*, (1998), Boston MA: Harvard Business School Press.
- [5.] Dousset, B., Karouach, S., *Manipulation de graphes de grande taille pour l'étude des réseaux d'acteurs et des réseaux sémantiques*, (2005), Colloque SFBA de l'Ile Rousse CD-Rom.
- [6.] Eades P., *A heuristic for Graph Drawing*, (1984), Congressus Numerantium, vol. 42, p. 149-160.
- [7.] Erten C., Harding P. J., Kobourov S. G., Wampler K. and Yee G., *Exploring the computing literature using temporal graph visualization*, (2004), Conference on Visualization and Data Analysis.
- [8.] Freeman, L. C., *A set of measures of centrality based on betweenness*, (1977), Sociometry, vol. 40, n° 6, p. 35-41.
- [9.] Freeman, L. S., *Centrality in social networks: Conceptual clarification*, (1979), Social Networks, vol. 1, p. 215-239.
- [10.] Gay, B., Dousset, B., *Innovation and network structural dynamics: Study of the alliance network of a major sector of the biotechnology industry*, Research Policy, In Press, (2005).
- [11.] Herman, I., Mélançon, G., Marshall, M., *Graph visualization and navigation in information visualization : a survey*. (2000), IEEE Transactions on visualization and computer graphics, vol., n° 1, p. 24-43.
- [12.] Jouve, B., Kuntz P., Velin, P., *Extraction de structures macroscopiques dans des grands graphes par une approche spectrale*, (2001), Extraction de connaissances et apprentissage (ECA), vol. 1, p. 173-194.
- [13.] Karouach, S., *Système de visualisations interactives pour la découverte de connaissances*. (2003), Thèse en informatique de l'Université Toulouse III.
- [14.] Newman, M. E. J., *The structure and fonction of complex networks*, (2003), SIAM review 45, p. 167-256.
- [15.] Van Dongen, S., *Graph clustering by flow simulation*, (2000), PhD thesis, University of Utrecht, Germany.