



**HAL**  
open science

## Mesurer la cohésion d'un groupe d'apprentissage en formation à distance

Christophe Reffay, Thierry Chanier

► **To cite this version:**

Christophe Reffay, Thierry Chanier. Mesurer la cohésion d'un groupe d'apprentissage en formation à distance. Environnements Informatiques pour l'Apprentissage Humain 2003, Apr 2003, Strasbourg, France. pp.367-378. edutice-00000152

**HAL Id: edutice-00000152**

**<https://edutice.hal.science/edutice-00000152>**

Submitted on 5 Nov 2003

**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.

---

## Mesurer la cohésion d'un groupe d'apprentissage en formation à distance.

**Comment utiliser les « réseaux sociaux » pour mesurer  
la cohésion en FAD ? Premières propositions.**

**Christophe Reffay\*, Thierry Chanier\***

*\*LIFC – UFR ST*

*16 route de Gray*

*25030 Besançon Cedex*

*Christophe.Reffay@univ-fcomte.fr*

*Thierry.Chanier@univ-fcomte.fr*

---

*RÉSUMÉ. Cette communication se propose de faire le point sur la notion de cohésion, proposée par les réseaux sociaux (Social Network Analysis), en vue de l'utiliser pour le suivi de la formation à distance, basée sur la production collaborative. Certains auteurs ayant montré l'impact de la cohésion du groupe sur son efficacité, nous nous intéressons à la définition, à la mesure et à la représentation de cette cohésion en vue de proposer des greffons sur les plates-formes de téléformation. Nous appliquons la notion de clique de niveau  $s$  à deux types de graphes de communication sur des données tirées d'une réelle formation : représentant les courriels d'une part, et les forums de discussion d'autre part. Nous montrons enfin qu'à partir de l'analyse des cliques de niveau  $s$ , il est possible de mesurer l'état de cohésion du groupe et la position relative de ses membres dans la communication du groupe ; information essentielle dans le suivi d'une formation collaborative.*

*MOTS-CLÉS : Formation à Distance, Mesure des interactions, Cohésion, Réseaux Sociaux, cliques.*

---

## 1. Introduction

Les interactions entre apprenants restent peu exploitées dans les plates-formes actuelles de formation à distance (FAD) alors qu'elles constituent un catalyseur dans le processus d'apprentissage et un facteur clé dans la cohésion d'un groupe d'apprenants [WOODRUFF 99]. Dans l'apprentissage collaboratif, la cohésion du groupe est l'élément déterminant qui motivera, entraînera et donc fera produire, discuter, négocier, reformuler les notions à apprendre. Si la cohésion n'est pas suffisante, les activités risquent fort d'être reçues par les acteurs comme une contrainte fastidieuse tandis que si la cohésion est à un niveau jugé satisfaisant, les activités se dérouleront naturellement et l'entraide entre les pairs semblera évidente. L'existence d'une forte cohésion dans un groupe d'apprenants crée en chaque acteur le besoin de retrouver les autres, pour affirmer l'objectif du groupe (dans notre cas : apprendre). Il semble donc primordial pour un tuteur ou un responsable de formation de pouvoir apprécier la cohésion ou la non-cohésion d'un groupe d'apprentissage. Dès lors se pose le problème de la mesure de la cohésion du groupe. Intuitivement, dans une classe en face à face, elle peut être appréciée par de nombreux signes visuels ou auditifs comme l'agitation, la communication des groupes, ou l'application des élèves. Dans le contexte de la FAD, ces signes n'étant pas perceptibles, et les communications étant relayées par la plate-forme, nous devons formaliser cette notion pour en donner une représentation.

Pour ce faire, nous avons interrogé les « réseaux sociaux » [DEGENNE & FORSE 94] (Social Network Analysis [WASSERMANN & FAUST 94]). Les réseaux sociaux sont issus de la sociométrie, de la dynamique des groupes, de la théorie des graphes et de leur intersection avec l'anthropologie [SCOTT 00]. La cohésion peut être, selon le contexte, mesurée suivant différents modèles. Nous en avons étudiés certains pour voir ceux qui peuvent le mieux servir le suivi des interactions en FAD. Alors que nous n'en sommes qu'au début de l'exploration des réseaux sociaux, nous avons déjà repéré plusieurs méthodes prometteuses. Nous nous proposons d'illustrer l'une d'elle dans cette communication.

Certains travaux tels que [WORTHAM 99] ont appliqué cette démarche à une micro-activité (i.e. : un bavardage d'une heure entre cinq acteurs non intégrés dans une formation), nous pensons (comme souligné dans [FJUK & LUDVIGSEN 01]), qu'il est nécessaire d'atteindre un niveau de complexité suffisant dans l'expérience pour pouvoir discuter de la pertinence des modèles fournis par les réseaux sociaux. Nous nous appuyons donc sur l'expérience Simuligne [CHANIER 02] pour tester et discuter ces différents modèles de la cohésion et apprécier la nature et la qualité des informations qu'ils mettent en évidence.

Après une nécessaire validation sur des données in situ, ces modèles pourront être intégrés aux plate-formes standards comme des greffons [MBALA & al. 02] permettant d'aider les acteurs de la FAD dans le suivi (« *monitoring* » selon [JERMANN & al. 01]) d'une téléformation collaborative. Il serait en effet prématuré

d'envisager un système « expert » (« *guiding* ») analysant ces modèles et fournissant des conseils aux acteurs. Nous limitons donc (à court terme) notre but à fournir des représentations de cette cohésion pour qu'elles soient évaluées par les acteurs humains de la FAD.

Enfin, compte tenu du nombre important de messages échangés dans une formation de taille réelle (plusieurs milliers), ce travail se propose d'étudier le flux statistique des messages entre les individus constituant les données sociométriques des groupes étudiés. L'étude des contenus sémantiques des messages, abordée par Sébastien George dans sa thèse [GEORGE 01] sur les conversations synchrones (mieux adaptée au contexte de la formation initiale) guidées par des introducteurs d'actes de langage, est une approche complémentaire, basée sur [PLETY 98], permettant de mieux définir qualitativement le rôle (profil) d'un acteur dans le groupe. La présente étude rend compte quantitativement des relations entre tous les acteurs (repérés individuellement) et permet d'en extraire une mesure pour le groupe dans son ensemble. Cependant, les représentations qui en sont extraites peuvent aussi être analysées pour chaque individu ou pour des sous-groupes.

Dans le contexte d'une formation collaborative, la communication est d'une importance capitale. Les outils de communication asynchrones textuels tels que le courriel et le forum, évitant les contraintes bloquantes de rendez-vous, restent à ce jour le meilleur compromis entre souplesse et interactivité pour les apprenants dans le contexte de la formation continue. C'est la raison pour laquelle ces deux outils ont été privilégiés dans notre étude. Leurs traces étant rémanentes pour les utilisateurs, nous pensons que leurs effets sur la cohésion sont plus durables. Nous nous proposons donc de mesurer la cohésion d'un groupe à partir des traces de communications asynchrones (courriers électroniques et messages des forums).

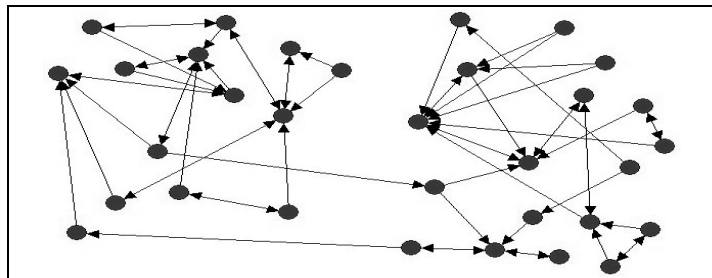
Dans l'interface d'une plate-forme de téléformation grand public servant une communauté d'apprentissage [CHANIER 01], on n'a en général pas accès à l'ensemble des informations décrivant complètement les messages déposés par un utilisateur dans les différents espaces (courriel, forum) au cours d'une période donnée. Il est donc nécessaire de rechercher, comptabiliser, et recomposer l'information utile à partir des traces elles-mêmes. Pour ce faire, nous passons par une base de données stockant pour chaque message : l'*auteur*, la *date d'envoi*, les *lecteurs*, et le *volume textuel* qu'il contient. Dès lors, on construit un graphe des échanges par courriel ou forum entre les acteurs pendant une période choisie.

La section 2 introduit le domaine des « réseaux sociaux » en précisant en particulier la définition des cliques. Nous présentons dans la section 3 l'expérimentation « Simuligne » sur laquelle nous avons testé les modèles, les mesures qu'ils renvoient et les interprétations que l'on peut en faire. La fin de la section 3 compare ces résultats au vécu des tuteurs de Simuligne, tandis que la section 4 nous permet de conclure sur les pistes prometteuses qu'ouvrent les réseaux sociaux (SNA : Social Network Analysis) dans l'étude des interactions.

## 2. Interrogeons les réseaux sociaux sur la cohésion

L'analyse structurale des réseaux sociaux [DEGENNE & FORSE 94] (Social Network Analysis [WASSERMANN & FAUST 94]) est issue de l'anthropologie structurelle et fonctionnelle d'une part, et de la sociométrie, de la dynamique des groupes et de la théorie des graphes d'autre part [SCOTT 00]. Plutôt qu'une analyse des différents attributs des individus d'une population définie a priori, l'analyse structurale focalise son étude sur les relations entre les individus pour en extraire des informations sur les groupes (densité, cohésion, multiplicité, hiérarchie, équivalence, ...) et les positions remarquables que les individus peuvent occuper au centre ou à la périphérie de ces groupes (distance, centralité, intermédialité). Enfin, selon le contexte de l'étude, ces propriétés remarquables peuvent être interprétées pour évaluer le fonctionnement du groupe ou détecter le rôle des individus repérés dans les positions remarquables (leaders, exclus, etc.)

Dans la plupart des études sociométriques proposées par les spécialistes des réseaux sociaux au siècle dernier, les relations étudiées sont non valuées. En 1943, Moreno s'intéressait aux réseaux affinitaires et il allait dans les classes d'école de différents niveaux et leur disait : « Si vous pouviez choisir votre voisin de gauche et votre voisin de droite, qui désigneriez vous ? ». Il récoltait ensuite les réponses, et construisait le réseaux des affinités entre les enfants d'une classe d'école. La relation affinitaire était représentée sur le graphe par un arc orienté non valué.



**Figure 1 : Exemple de graphe sociométrique de Moreno, 1943.**

Dans le sociogramme donné à la figure 1, les 14 garçons sont représentés dans la partie gauche et les 18 filles dans la partie droite. On note immédiatement le peu de relations entre les deux sexes.

Sur un graphe, on définit un chemin comme une séquence de sommets adjacents deux à deux. La distance du plus court chemin reliant deux sommets définit la *distance géodésique* séparant les deux sommets. On peut alors définir la *cohésion* comme l'inverse de la distance géodésique entre deux sommets. Dans ce cas, la cohésion est synonyme de « proximité ». Ainsi, selon [BORGATTI & al. 02], Ronald Burt définit la « proximité »  $p_{ij}$  (comprise entre 0 et 1) à partir de la distance

géodésique  $d_{ij}$  (comprise entre 1 et  $n-1$ ) entre les deux sommets  $i$  et  $j$  d'un graphe à  $n$  sommets comme suit :

$$p_{ij} = 1 - \frac{d_{ij} - 1}{n - 1}, \text{ pour } i \neq j$$

Mais la « proximité » (selon les chemins géodésiques) n'est pas la seule mesure possible pour la cohésion. Plus généralement, *l'indice de cohésion entre deux sommets représente la force du lien qui relie ces deux sommets*. La signification de sa valeur est directement rattachée à celle de la relation d'une part et à la manière de la calculer d'autre part.

La mesure de la cohésion n'est généralement pas une fin en soit, elle n'est qu'un outil pour caractériser les sous-groupes de différentes natures tels que les « cliques » et les « cliques de niveau  $s$  » définies comme suit :

- **Clique** : sous-graphe complet maximum : à l'intérieur de ce sous-graphe, chaque sommet est directement relié à tous les autres.
- **Clique de niveau  $s$**  : pour les graphes symétriques valués : tout couple de sommets d'une « clique de niveau  $s$  » est relié par un arc de valeur au moins «  $s$  ».

Des outils tels que UCINET 6 [BORGATTI & al. 02] permettent, en outre, d'extraire ces deux types de sous-réseaux.

### 3. Application au contexte de la FAD

Ayant présenté dans la partie précédente, les notions issues des réseaux sociaux, utilisées dans la suite, nous montrons ici, comment elles peuvent servir la FAD.

#### 3.1. Simuligne : des données réelles d'interactions en FAD

Dans le cadre du projet ICOGAD [CHANIER 02], une session de formation à distance nommée Simuligne a été conçue par la collaboration de chercheurs en sciences du langage d'une part et en informatique d'autre part. Simuligne est une formation collaborative [HENRI & CAYROL 01], 100% à distance, extensive (environ 4 heures de travail requis par apprenant et par semaine) pour 40 apprenants, 4 tuteurs/enseignants, sur 10 semaines, dans le domaine du français langue étrangère (FLE). Les apprenants étaient répartis en 4 groupes de 10, chaque groupe ayant à disposition 2 ou 3 natifs (francophones de langue maternelle) : référents/experts pour la langue et la culture française. La formation s'est déroulée sous la forme d'une simulation globale en 4 étapes selon le calendrier suivant :

- E0 (2 semaines) : Constitution du groupe et prise en main de la plate-forme ;
- E1 (3 semaines) : Construction du décor et des personnages ;
- E2 (3 semaines) : Mise en scène des personnages et jeu de rôle ;
- E3 (2 semaines) : Vote, remise des prix et bilan.

Cette simulation globale s'est déroulée pour chaque groupe (Aquitania, Gallia, Lugdunensis et Narbonensis) dans une « classe virtuelle » différente interdisant les communications entre les groupes. D'autres espaces, communs à tous ou seulement aux formateurs permettaient d'échanger dans des activités culturelles plus libres. Au total, pour une formation ayant duré 10 semaines (E0 à E3), nous avons enregistré :

- Forum (asynchrone) : 879 015 caractères dans 2686 messages, représentant 45,11% du volume total de communication en caractères ;
- Courriel (asynchrone) : 834 753 caractères dans 4062 messages, soit 42,84% du volume total de communication en caractères ;
- Bavardage (synchrone) : 234 694 caractères dans 5680 tours de parole, soit 12% du volume total de communication en caractères.

A partir de ces données réelles, nous montrons comment l'analyse structurale des réseaux sociaux peut mettre en évidence la cohésion des sous-groupes.

### 3.2. *Communications, relations et graphes en FAD*

Pour chaque outil de communication de la plate-forme, il est nécessaire de définir le graphe qui représente les flux informationnels qu'il engendre. Il est important de préciser comment (quantitativement) et quand, un message émis par un acteur a à un acteur b, est pris en compte dans le calcul du graphe sociométrique.

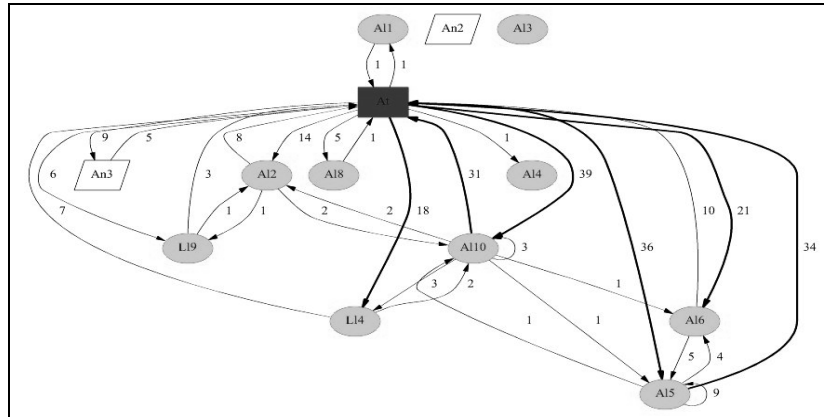
#### **Définition du Graphe $G_F(A,M,P)$ des échanges par Forum de discussion**

Nous définissons le graphe orienté et valué  $G_F(A,M,P)$  sur l'ensemble A des acteurs avec la relation  $M : A \times A \rightarrow N$ , qui à chaque couple (a,b) de  $A \times A$  associe le nombre de messages **postés par a et ouverts par b** pendant la période P dans les **forums** de discussion.

#### **Définition du Graphe $G_C(A,M,P)$ des échanges par Courriel**

$G_C(A,M,P)$  définit le graphe orienté et valué sur l'ensemble A des acteurs avec la relation  $M : A \times A \rightarrow N$ , qui à chaque couple (a,b) de  $A \times A$  associe le nombre de **courriels émis par a** et reçus et **ouverts par b** pendant la période P.

Sur le graphe de la figure 2, le tuteur (At) apparaît dans un rectangle, les natifs (An2 et An3) dans des parallélogrammes et les 10 apprenants (Al\* et Ll\*) dans des ellipses. Cette figure montre bien qu'au cours de cette formation, le courriel est utilisé essentiellement dans les relations Tuteur-Apprenant. Il fait immédiatement apparaître les individus n'ayant échangé aucun courriel : An2 et Al3. En faisant abstraction de l'orientation et de la valeur des arcs, on y trouve 4 cliques de plus de 2 individus : (Al2, Ll9, At), (Al2, Al10, At), (Ll4, Al10, At) et (At, Al10, Al5, Al6). On note que le tuteur (At) figure dans toutes les cliques, ce qui lui confère une position centrale dans le groupe. Al10 étant membre de 3 des 4 cliques citées plus haut, il a lui aussi une position relativement centrale. De façon duale, cette centralité peut être mesurée pour chaque individu, par le nombre de cliques auxquelles il appartient.



**Figure 2 : Graphe des communications par courriel dans le groupe Aquitania sur la période couvrant l'ensemble de la formation.**

Si maintenant on tient compte de l'orientation et de la valeur des arcs, on peut aussi vérifier qu'au sein des 4 cliques définies ci-avant, les échanges entre apprenants sont en nombres très variables. Par exemple, A16 entretient une relation bien plus soutenue avec A15 (9 messages) qu'avec A110 (1 message). Nous voyons ici que la force de la relation n'est pas prise en compte dans la notion de clique telle que définie dans la partie 2. La présence de A110 dans la clique à 4 sommets ne tient qu'à l'échange d'un seul message entre A110 et A16 d'une part et un dans chaque sens entre A110 et A15 d'autre part. L'analyse des cliques simples ne tient pas compte de la différence des forces de liaison entre les sommets. Elle se borne à l'existence/absence de relation entre les sommets. Ce problème a été mis en évidence dans [REFFAY & CHANIER 02] en particulier sur l'évolution d'un indice de cohésion sur l'ensemble du groupe au cours de la formation. C'est justement pour lever cette difficulté que nous utilisons la notion de *clique de niveau s* présentée dans la suite de ce papier.

### 3.3. Recherche de « clique de niveau s » dans un graphe orienté et valué

A partir d'un graphe orienté et valué tel que celui donné à la figure 2, on peut symétriser les relations du graphe en prenant la somme des valeurs des deux arcs opposés ainsi :  $s_{ij}=m_{ij}+m_{ji}$ . En appliquant un filtre de seuil  $s$ , on peut décider de supprimer les arcs de valeur inférieure à  $s$ , puis ignorer la valeur des arcs restants pour se ramener au cas du graphe non orienté et non valué, et donc à la recherche de cliques simples. Les cliques simples obtenues sur le graphe symétrisé non valué correspondent aux cliques de niveau  $s$  dans le graphe initial.

En prenant comme seuil  $s=3$ , dans notre contexte, cela signifie que les membres d'une même clique de niveau 3 auront dû échanger au moins 3 courriels avec chacun des autres membres de cette clique de niveau 3. La figure 3 compare la liste



des cliques simples obtenues précédemment avec la liste des cliques de niveau 3 sur le graphe des courriels du groupe Aquitania pour toute la période. Les différences entre ces 2 ensembles de cliques apparaissent en gras dans la figure 3.

	<b>(At, AI1)</b>	<b>(At, AI2, LI9)</b>	<b>(AI1)</b>	<b>(At, AI2)</b>	
	<b>(At, AI4)</b>	(At, AI2, AI10)	<b>(AI4)</b>	<b>(At, LI9)</b>	(At, AI2, AI10)
(AI3)	(At, AI8)	(At, LI4, AI10)	(AI3)	(At, AI8)	(At, LI4, AI10)
(An2)	(At, An3)	(At, AI5, AI6, <b>AI10</b> )	(An2)	(At, An3)	(At, AI5, AI6)
<b>Cliques simples</b>			<b>Cliques de niveau 3</b>		

Figure 3 : Comparaison « cliques simple » - « cliques de niveau 3 »

Comme le montre la figure 3, les cliques simples (At, AI1) et (At, AI4) à deux sommets (tuteur-apprenant), dont la valeur de la relation symétrisée devient respectivement 2 et 1 (inférieure à 3) ne sont pas des cliques de niveau 3. Ceci permet de mettre en évidence deux nouveaux apprenants isolés : AI1 et AI4. De même, la clique simple (At, AI2, LI9) se divise en deux cliques de niveau 3 (At, AI2) et (At, LI9), et AI10 est exclu de la clique (At, AI5, AI6, AI10).

Cette notion de seuil permet ainsi d'éliminer les relations négligeables pour constituer une clique de niveau acceptable. Le seuil à choisir est la valeur en deçà de laquelle la communication entre deux acteurs est jugée insuffisante. Bien que ce seuil puisse sembler arbitraire, le concepteur de la formation (ou le tuteur en charge de son suivi) n'aura aucune peine à en fixer la valeur. Le résultat obtenu au bout du calcul des cliques de niveau 3 est synthétisé sur la figure 4 par un diagramme du nombre de cliques (de niveau 3) de 1, 2, 3 ou 4 individus. Une clique de 1 individu est un individu isolé. Ainsi, pour le groupe Aquitania, sur l'ensemble de la période, on a 4 singletons (cliques à un individu), 4 cliques de 2 et 3 cliques de 3 individus. Le diagramme de cette figure permet de comparer ces valeurs avec celles des autres groupes sur toute la période.

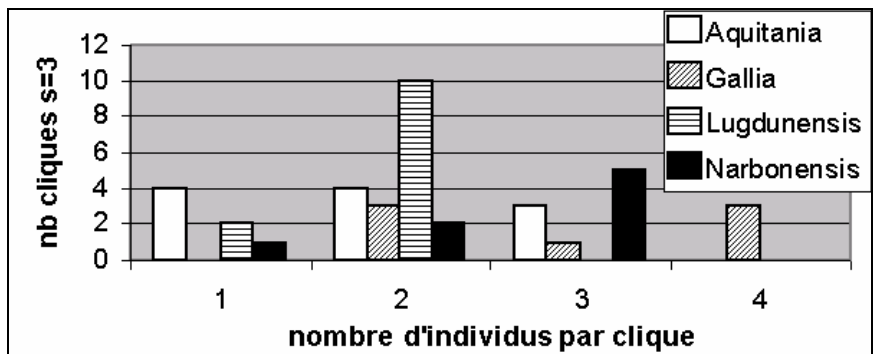
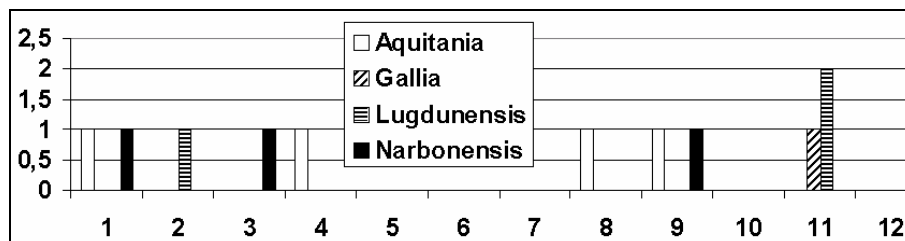


Figure 4 : Nombre de cliques de niveau 3 pour le courriel

Sur le diagramme (fig. 4), un groupe de  $n$  individus isolés (cohésion minimale) aurait  $n$  cliques de 1, tandis qu'un groupe dont les  $n$  individus auraient échangé au moins 3 courriels avec tous les autres (cohésion maximale) aurait une seule clique à  $n$  individus. On repère donc un groupe de cohésion plus forte au fait qu'il atteigne des tailles de cliques plus importantes. Gallia est le seul groupe qui comporte des cliques (de niveau 3) à 4 sommets, et il en comporte 3 ! On note également qu'il n'a aucun élément isolé. À l'inverse, le groupe Lugdunensis n'a aucune clique (de niveau 3) de plus de deux sommets. Cela signifie qu'aucun apprenant de ce groupe n'a échangé plus de 2 courriels avec un autre apprenant.

Si l'on compare ces résultats avec le vécu de la formation Simuligne (difficile à rapporter ici), nous estimons que l'analyse des cliques de niveau 3 sur les courriels rend assez bien compte du niveau de cohésion des groupes. Précisons par exemple que le groupe Lugdunensis ayant compté trop d'abandons n'a pas pu terminer la formation. Nous attirons cependant l'attention du lecteur sur le fait qu'un groupe donné pourra compter d'autant plus de cliques qu'il a de sommets au départ, et que donc, sans normalisation, cet outil ne peut être employé que pour comparer des groupes de tailles semblables.

Le courriel étant un outil de communication inter-individuel, il est normal que, en dehors de la relation apprenant-tuteur, les échanges entre les autres acteurs soient plus rares (mais donc remarquables). En effet, l'apprenant s'adresse bien souvent à l'ensemble du groupe (dans le forum), ou au tuteur (par courriel). C'est pourquoi nous avons étudié de la même façon, les forums de discussion pour les 4 groupes. Le résultat, donné sur la figure 5 pour un seuil de 3 montre que l'on atteint des cliques comptant jusqu'à 11 membres (i.e : la totalité pour Gallia).



**Figure 5 : Nombre de cliques de niveau 3 pour les forums de discussion**

Dans le cas des forums de discussions dans chaque groupe, si l'on cumule pour chaque paire (A, B) d'individus, le nombre de messages postés par A et lus par B avec le nombre de messages postés par B et lus par A, on obtient une moyenne de 38 messages chez Aquitania, 44 chez Gallia, 9 chez Lugdunensis et de 24 chez Narbonensis. Ces valeurs moyennes représentent un bon indicateur global pour le niveau d'échange entretenu dans les forums, et nous montrent que le seuil choisi (3) n'est pas en rapport avec ces niveaux moyens d'échange. En tant que formateur, on pouvait attendre un niveau minimum d'échanges de l'ordre de 10.

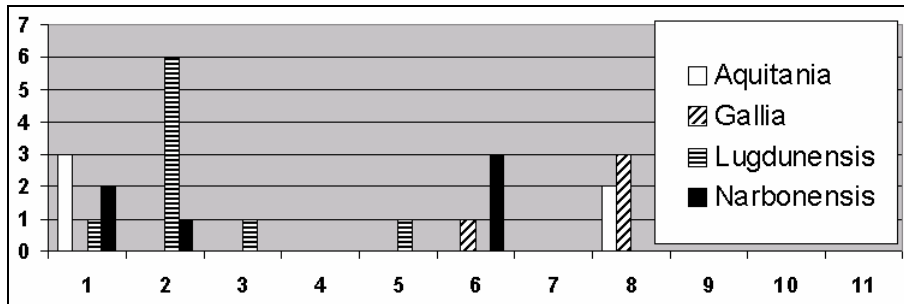


Figure 6 : Nombre de cliques de niveau 10 pour les forums de discussion

Les figures 5 et 6 montrent l'importance du choix du seuil. En effet, sur la figure 5, pour un seuil égal à 3, le groupe Lugdunensis semble être d'aussi forte cohésion que le groupe Gallia, étant les seuls à admettre au moins une clique de 11 individus. En remontant le seuil de 3 à 10 messages, on voit nettement sur la figure 6 que ce groupe Lugdunensis est le seul à n'admettre aucune clique de plus de 5 individus, ce qui le place très en dessous des autres pour la cohésion du groupe. Ceci montre aussi la difficulté de représenter simultanément la quantité de messages échangés et la manière dont ces échanges sont répartis dans le groupe. L'analyse en cliques de niveau  $s$  nécessite de fixer a priori un seuil  $s$  couperet supprimant ou conservant les arcs selon que leur valeur est respectivement inférieure ou supérieure à cette valeur. Notons que l'on ne distingue plus ceux dont les arcs étaient nuls à l'origine de ceux qui se trouvaient en deçà du seuil. Les arcs de valeur égale ou supérieure au seuil sont aussi plongés dans une seule catégorie sans distinction de leur valeur initiale. C'est pourquoi le choix de ce seuil est très important. Si, comme c'est le cas dans Simuligne, il correspond à une réalité dans la formation (définie par des consignes ou évidente pour tous les acteurs) alors cette analyse des cliques de niveau  $s$  devrait rendre compte de la structure du groupe et donc de son niveau de cohésion.

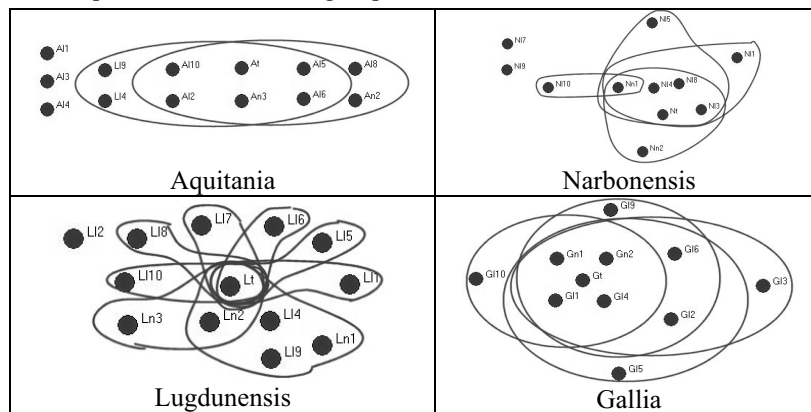


Figure 7 : Cliques de niveau 10 pour les forums de discussion de Simuligne

La figure 7 donne une illustration manuelle des cliques de niveau 10 calculées automatiquement. On y perçoit par exemple la position singulière des apprenants L14 et L19 (initialement dans Lugdunensis) qui ont été transférés dans le groupe Aquitania au moment de l'arrêt de Lugdunensis.

On voit aussi sur cette figure 7 que l'unité est plus grande dans Gallia (aucun élément isolé) avec un noyau de 5 éléments membres de toutes les cliques. Dans Narbonensis, on détecte comme singularité la clique à 2 éléments signifiant une préférence de l'élément extérieur (N110) pour Nn1. La structure des deux cliques dans Aquitania manifeste une dichotomie dans le temps : L14 et L19 sont entrés dans Aquitania quand A18 et An2 ont diminué leur activité, tandis que les six autres sont restés du début à la fin. Enfin, la structure de Lugdunensis en étoile autour du tuteur, résulte d'un silence de la part des apprenants sur le forum (sauf pour L14 et L19 qui sont membres de la clique de 5 les réunissant à 2 natifs et au tuteur). Les résultats fournis par l'analyse des cliques de niveau s sur les messages soulignent les phénomènes observés au cours de la formation Simuligne.

#### 4. Conclusion

Partant de l'hypothèse que la cohésion d'un groupe dans une formation collaborative est un facteur clé dans la réussite de l'apprentissage, nous avons consulté le champ disciplinaire des réseaux sociaux pour mieux saisir le concept de cohésion. Nous en avons rapporté les notions de cliques et de cliques de niveau s. Nous avons ensuite proposé, comme un préalable à toute mesure de la cohésion dans un groupe, une définition de graphes de communications pour représenter les courriels d'une part et les forums de discussion d'autre part. Nous avons montré que, dans Simuligne, les cliques de niveau 10 étaient plus adaptées aux forums de discussion que les cliques de niveau 3. La représentation graphique des cliques extraites et leur recouvrement permet de mettre en évidence de nombreuses singularités recoupant étroitement les phénomènes vécus dans Simuligne. Il nous reste à tester une valuation de nos arcs par le volume des messages plutôt que par leur nombre. Mais au préalable, nous souhaitons comparer la notion de cliques de niveau s à celle des clusters hiérarchiques [REFFAY & al. 03]. Nous avons d'ores et déjà acquis la certitude que les notions issues des réseaux sociaux donnent une bonne lecture sur l'ensemble complexe que constitue une formation en ligne, au niveau de l'individu ou du groupe. Cette avancée pourrait améliorer les outils de réification pour l'apprenant ou d'évaluation pour les tuteurs/enseignants en FAD.

#### Remerciements

Nous tenons à remercier le ministère de la recherche pour le soutien du projet ICOGAD<sup>1</sup> dans le programme Cognitique 2000.

---

<sup>1</sup> ICOGAD est un projet transdisciplinaire entre chercheurs en informatique de Besançon, en Psychologie de Nancy2 et en Sciences du Langage de l'Open University (UK).

## 5. Bibliographie

- [BORGATTI & al. 02] BORGATTI, S.P., EVERETT, M.G. and FREEMAN, L.C. 2002. UCINET 6 for Windows. Harvard: Analytic Technologies. <http://www.analytictech.com/>
- [CHANIER 01] CHANIER, T., « Créer des communautés d'apprentissage à distance ». *Les dossiers de l'Ingénierie Educative*, n°36, Centre National de Documentation Pédagogique(CNDP) : Montrouge. pp 56-59, 2001.
- [CHANIER 02] CHANIER, T., « Rapport de fin de recherche », *Cognitive2000*, rapport de fin de projet, projet n°23, ministère de la recherche, 2002.
- [DEGENNE & FORSE 94] DEGENNE A., FORSE M., *Les réseaux sociaux*, Collection U, Série Sociologie, Armand Colin, Paris, 1994.
- [FJUK & LUDVIGSEN 01] FJUK, A., LUDVIGSEN, S., « The Complexity of Distributed Collaborative Learning: Unit of Analysis ». (*EU-CSCL'01*), Maastricht, 2001.
- [GEORGE 01] GEORGE, S., « Apprentissage collectif à distance. SPLASH : un environnement informatique support d'une pédagogie de projet », *Thèse de Doctorat en informatique*, Université du Maine, 2001.
- [HENRI & CAYROL 01] Henri, F., Lunbgrén-Cayrol, K., *Apprentissage collaboratif à distance. Pour comprendre et concevoir les environnements d'apprentissage virtuels*. presses de l'Université du Québec, 2001.
- [JERMANN & al. 01] Jermann, P., Soller, A., Muehlenbrock, M., « From mirroring to guiding: a review of state of art technology for supporting collaborative learning ». *Proceedings of EU-CSCL'01*, Maastricht, 2001.
- [MBALA & al. 02] MBALA, A., REFFAY, C., CHANIER, T., « Integration of automatic tools for displaying interaction data in computer environments for distance learning », *Intelligent Tutoring System Conference*, Biarritz, France, 2002, p. 841-850.
- [PLETY 98] PLETY, R., *Comment apprendre et se former en groupe*, Retz, 1998.
- [REFFAY & CHANIER 02] REFFAY, C., CHANIER, T., « Social Network Analysis Used for Modelling Collaboration in Distance Learning Groups », *Intelligent Tutoring System Conference*, Biarritz, France, 2002, p. 31-40.
- [REFFAY & CHANIER 03] REFFAY, C., CHANIER, T., « How social network analysis can help to measure cohesion in collaborative distance learning? », *Proceeding of the Computer Supported Collaborative Learning Conference. (CSCL'03)*, Bergen, 2003.
- [SCOTT 00] SCOTT, J., *Social Network Analysis : a handbook*, 2<sup>e</sup>ed., SAGE, London, 2000.
- [WASSERMANN & FAUST 94] WASSERMANN S., FAUST K., *Social Network Analysis : Methods and Applications*, Cambridge University Press, New York, 1994.
- [WOODRUFF 99] WOODRUFF, E., « Concerning the Cohesive Nature of CSCL Communities », *Proceedings of CSCL '99*, 1999, p. 677-680.
- [WORTHAM 99] WORTHAM, D.W., « Nodal and matrix analyses of communication patterns in small groups ». *Proceedings of CSCL'99 Palo Alto, CA, (1999)* 681-686