

2000 Cahiers du GERAD

Z. Atoui, G. Caporossi,
S. Perron

G-2013-35

Mai 2013

2000 Cahiers du GERAD

Zeinab Atoui
Gilles Caporossi
Sylvain Perron

GERAD & HEC Montréal
Montréal (Québec) Canada, H3T 2A7

zeinab.atoui@gerad.ca
gilles.caporossi@gerad.ca
sylvain.perron@gerad.ca

Mai 2013

Les Cahiers du GERAD
G-2013-35

Copyright © 2013 GERAD

Résumé : Le nombre de Cahiers du GERAD vient de dépasser 2000. C'est un nombre qui en dit long sur l'activité du centre. Nous profiterons de cette occasion pour faire une analyse succincte de l'évolution de ces publications. Depuis 1979, lors de la création du GERAD, le nombre de publications, mais aussi de membres a évolué. Nous nous proposons ici de regarder non seulement le nombre de Cahiers chaque année, mais de voir comment la structure de la collaboration a changé avec le temps. Une analyse des communautés que l'on retrouve sera effectuée à l'aide du logiciel **Constellations** par tranches de 5 ans depuis 1980 jusqu'à aujourd'hui.

1 Introduction

Avant d'étudier la structure de collaborations et son évolution, nous nous proposons d'effectuer une brève description du nombre de publications au fil du temps. Étant donné que le nombre de membres du GERAD a fortement changé au fil de ces années, il est certainement intéressant de mettre en relation le nombre de membres du GERAD et le nombre de Cahiers publiés. La figure 1 montre le nombre de membres et de Cahiers publiés chaque année.

Le premier constat n'est pas très surprenant : le nombre de Cahiers du GERAD publiés semble globalement augmenter. Mais force est de constater que le nombre de membres a cru dans une forte proportion, de sorte que le nombre de Cahiers par membre, s'il a augmenté au début, est en diminution depuis les années 1990, comme l'illustre la figure 2.

Au delà du simple constat du nombre de Cahiers, nous nous proposons maintenant d'analyser la structure de collaboration, à l'aide de techniques d'analyse des réseaux sociaux.

2 Description des données

Ce travail repose sur la liste complète des Cahiers du GERAD. Un travail préliminaire de nettoyage des données a été nécessaire, comme chaque fois que des données sont utilisées dans un but autre que la raison pour laquelle elles ont été recueillies. Ici, la principale difficulté consistait à ajuster les noms des auteurs, car il arrivait que le même auteur apparaisse avec jusqu'à quatre différentes notations. La cause de ces changements n'était que rarement due à des erreurs dans l'écriture du nom, mais dans la manière de noter les initiales, surtout pour les prénoms composés. Par exemple, Jean-Philippe pouvait apparaître comme J.P., J.-P., J.Ph. ou J.-Ph., ce qui ne pose pas de problème en général, mais ces différentes notations sont considérées comme des individus différents lors d'un traitement informatique. Nous avons donc uniformisé ces notations. Il arrive fréquemment que des Cahiers du GERAD soient révisés, en quel cas les titres et les noms des auteurs peuvent changer. Étant donné que le numéro du Cahier ne change pas, nous avons considéré qu'il ne s'agissait que d'un seul Cahier, surtout qu'il arrive qu'un Cahier subisse plusieurs révisions et que seules la première et dernière révisions apparaissent dans le fichier de données. L'analyse a été effectuée à partir de graphes de co-auteurs. Il y a divers moyens de représenter les collaborations par un graphe. Tout d'abord, il est possible

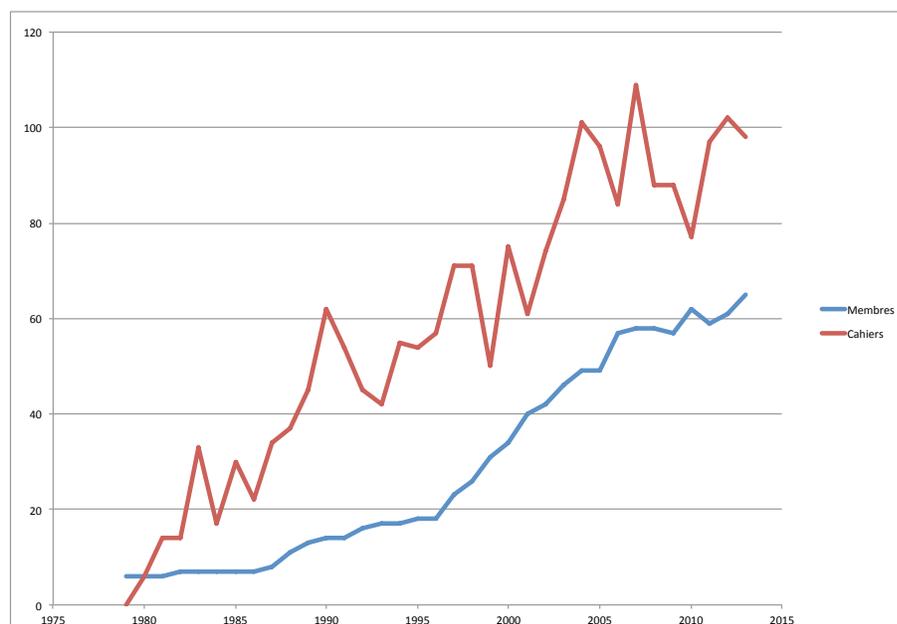


Figure 1 – Nombre de Cahiers du GERAD et nombre de membres au fil du temps

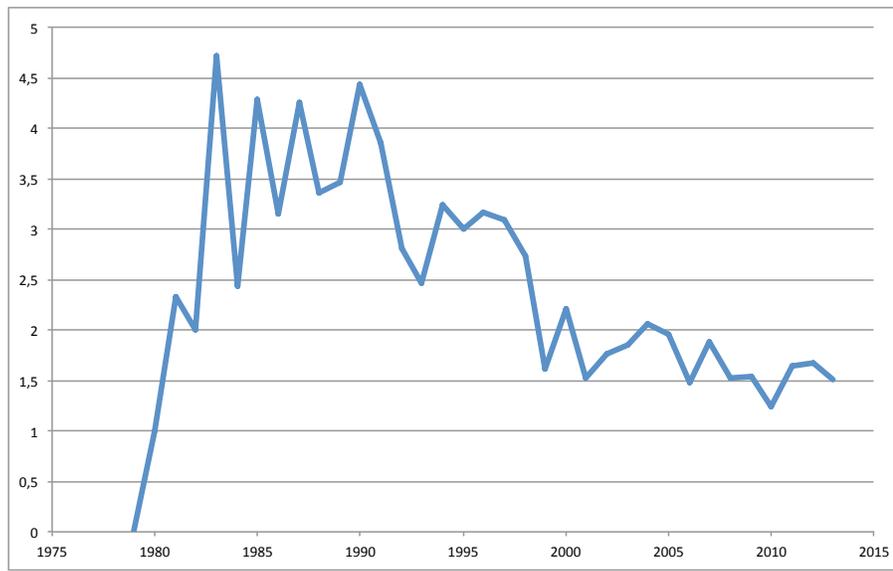


Figure 2 – Nombre de Cahiers du GERAD par membre au fil du temps

de travailler avec un graphe comportant simultanément les auteurs et les Cahiers, chaque nœud représentant un auteur ou bien un Cahier. Dans ce graphe, un arc représente la relation auteur-Cahier, de sorte qu'il n'y aurait pas d'arc entre auteurs ni entre Cahiers. Un tel graphe serait biparti. Une seconde alternative consiste à ce que chaque auteur soit représenté par un nœud et qu'un arc relie deux nœuds si les auteurs correspondants ont écrit au moins un Cahier ensemble durant la période considérée. Il est alors possible de pondérer les arcs pour tenir compte du nombre de collaborations entre auteurs. Le choix que nous avons fait ici est celui d'utiliser un graphe avec seulement les auteurs, et sans utiliser de pondération. Ce choix tient au fait que nous cherchons ici avant tout à avoir un profil des collaborations, de voir si, oui ou non, les individus collaborent.

3 Évolution de la structure de collaboration

Le GERAD existe depuis 1979, et il est évident qu'il a changé considérablement depuis cette date. Pour refléter cette évolution, nous avons décidé d'analyser la structure de collaboration non pas de manière globale, mais au fil du temps, en séparant les données par tranches successives de 5 ans.

Pour effectuer cette analyse, nous utiliserons ici le logiciel *Constellations* en développement au GERAD depuis les quelques dernières années par G. Caporossi et S. Perron. Le critère utilisé pour la détection de communautés est la modularité [3, 4]. *Constellations* est basé sur l'utilisation d'une version parallèle de la métaheuristique de recherche à voisinages variables [1, 2]. Bien sûr, cette approche de résolution ne garantit pas l'optimalité, mais donne généralement d'assez bons résultats pour pouvoir baser nos analyses dessus.

Afin d'analyser l'évolution de la structure, nous avons choisi de décomposer les données par tranche de 5 ans. Le temps nécessaire pour mener à bien une recherche étant parfois assez long, il nous a semblé qu'une fenêtre de 5 ans serait plus appropriée qu'une étude année par année, comportant une trop grande variabilité.

Le sommaire des résultats de ces analyses se trouve au tableau 1. Il est à noter que l'année 2013 est ici tronquée pour ne considérer que les 2000 premiers Cahiers. La dernière période d'analyse ne comporte donc que les Cahiers des trois dernières années et demie. Par ailleurs, certains Cahiers en ayant remplacé d'autres, selon la manière de compter, un écart de deux Cahiers peut intervenir, ce qui explique que seulement 1998 Cahiers du GERAD sont considérés ici.

La première remarque en voyant ces chiffres, est que le nombre de communautés semble globalement augmenter avec le temps, ce qui peut faire penser que des groupes nouveaux s'ajoutent au GERAD. Cette

Table 1 – Sommaire recherches de communautés

Période	Nombre de Cahiers	Nombre de communautés	Nombre de composantes connexes	Modularité
1980–1984	84	13	5	0,732772
1985–1989	168	23	9	0,785962
1990–1994	258	21	4	0,784697
1995–1999	303	21	7	0,827246
2000–2004	396	33	17	0,866651
2005–2009	465	29	12	0,87735
2010–2013	324	31	20	0,906614

hypothèse semble confortée par le fait que la valeur de la modularité croît aussi. De fait, il semble non seulement qu'il y ait de plus en plus de communautés, mais aussi que celles-ci soient plus distinctes, comme le laisse penser la modularité dépassant 0,9 pour la dernière période. Le critère choisi pour cette remarque, la maximisation de la modularité, peut toutefois être remis en question pour plusieurs raisons, et nous nous devons d'éviter les conclusions hâtives. La maximisation de la modularité est un critère difficile à optimiser et la technique utilisée, la recherche à voisinages variables, si elle fonctionne généralement bien, n'assure pas l'optimalité. Il est donc possible que nous soyons confrontés à un artefact. Par contre, le nombre de composantes connexes n'est pas obtenu par optimisation, et tend à donner la même information.

4 Conclusion et travaux à venir

Le but dans ce court portrait est purement descriptif et ne montre nullement que les chercheurs du GERAD ne collaborent pas, ou pas assez. Tout d'abord, le nombre de chercheurs croît sans cesse, et la collaboration n'est jamais immédiate, ce qui peut être une explication de ce phénomène.

Les recherches étant souvent longues, il est possible aussi que les collaborations se forment, puis changent dans le temps et que l'analyse d'une courte tranche de travail ne montre pas cette dynamique. Un travail plus approfondi serait alors nécessaire.

Références

- [1] Hansen, P. and Mladenović, N., Variable neighborhood search: Principles and applications. *European J. Oper. Res.*, 130, 449–467, 2001.
- [2] Mladenović, N. and Hansen, P., Variable neighborhood search. *Comput. Oper. Res.*, 24, 1097–1100, 1997.
- [3] Newman, M. and Girvan, M., Finding and evaluating community structure in networks. *Physical Review E*, 69, 026133, 2004.
- [4] Newman, M., Modularity and community structure in networks. *Proceedings of the National Academy of Sciences, USA*, 8577–8582, 2006.

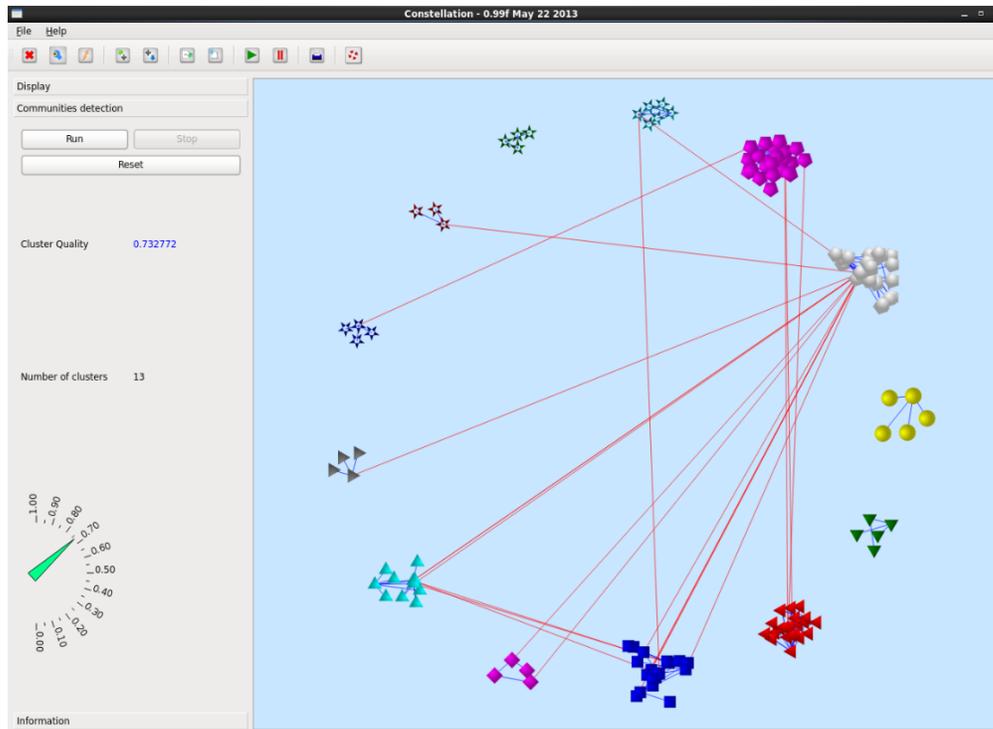


Figure 3 – Graphe des collaborateurs de 1980 à 1984

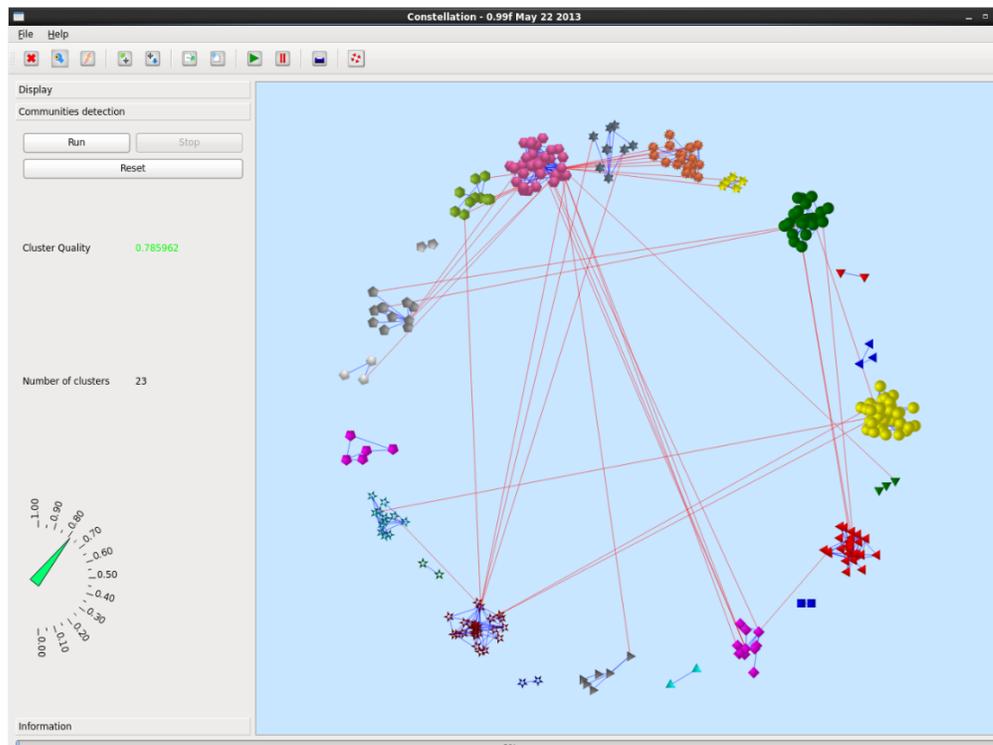


Figure 4 – Graphe des collaborateurs de 1985 à 1989

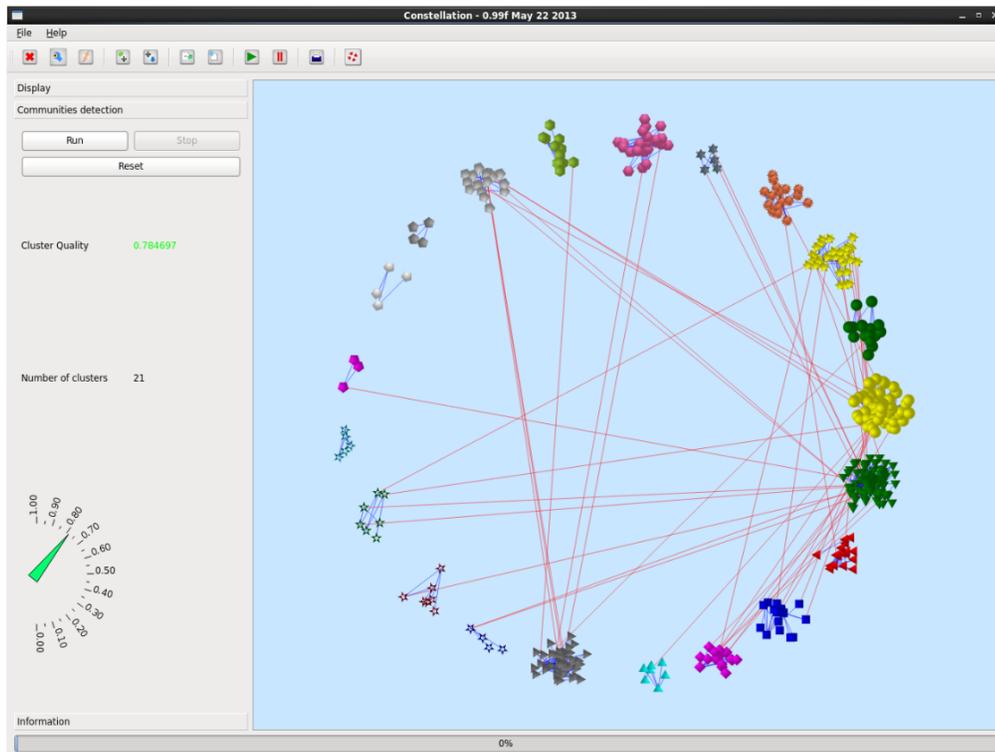


Figure 5 – Graphe des collaborateurs de 1990 à 1994

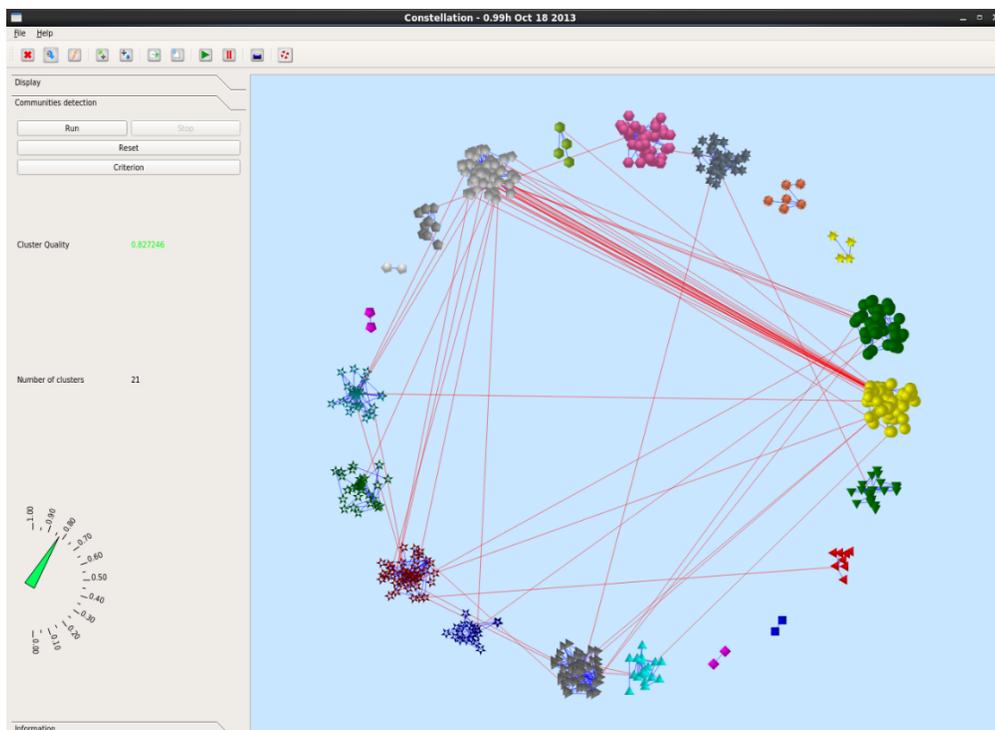


Figure 6 – Graphe des collaborateurs de 1995 à 1999

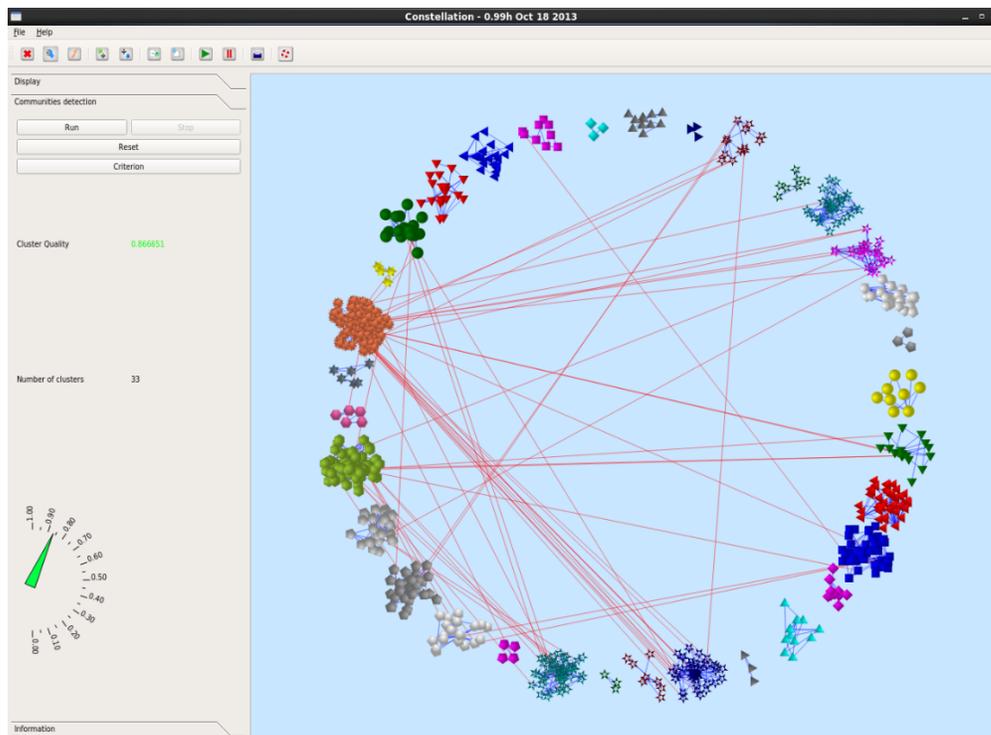


Figure 7 – Graphe des collaborateurs de 2000 à 2004

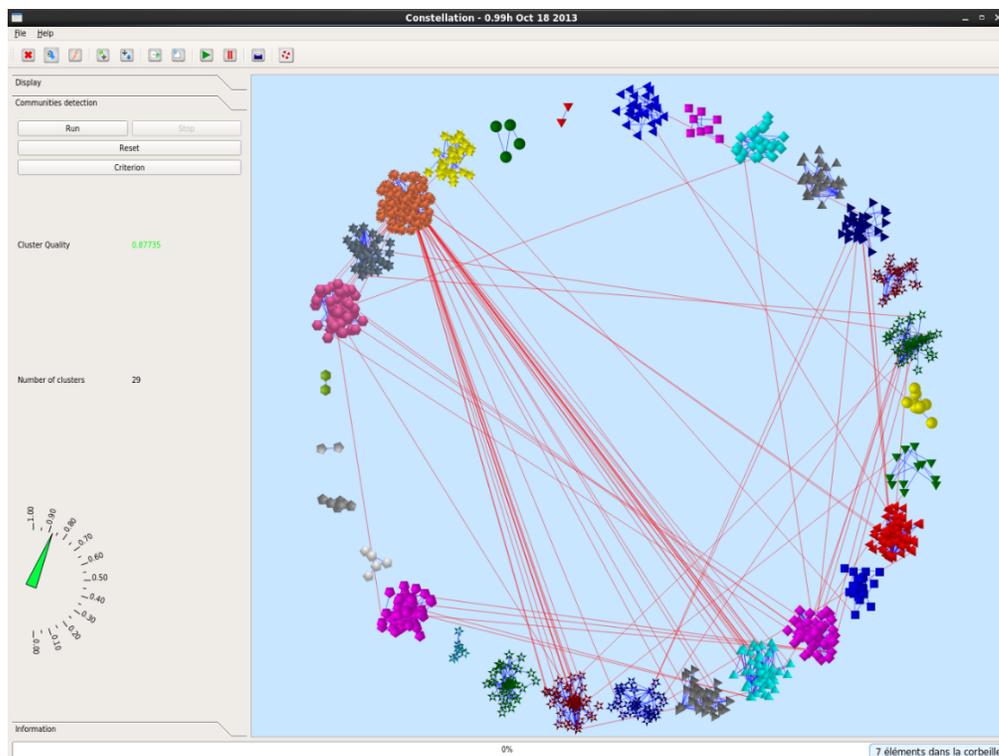


Figure 8 – Graphe des collaborateurs de 2005 à 2009

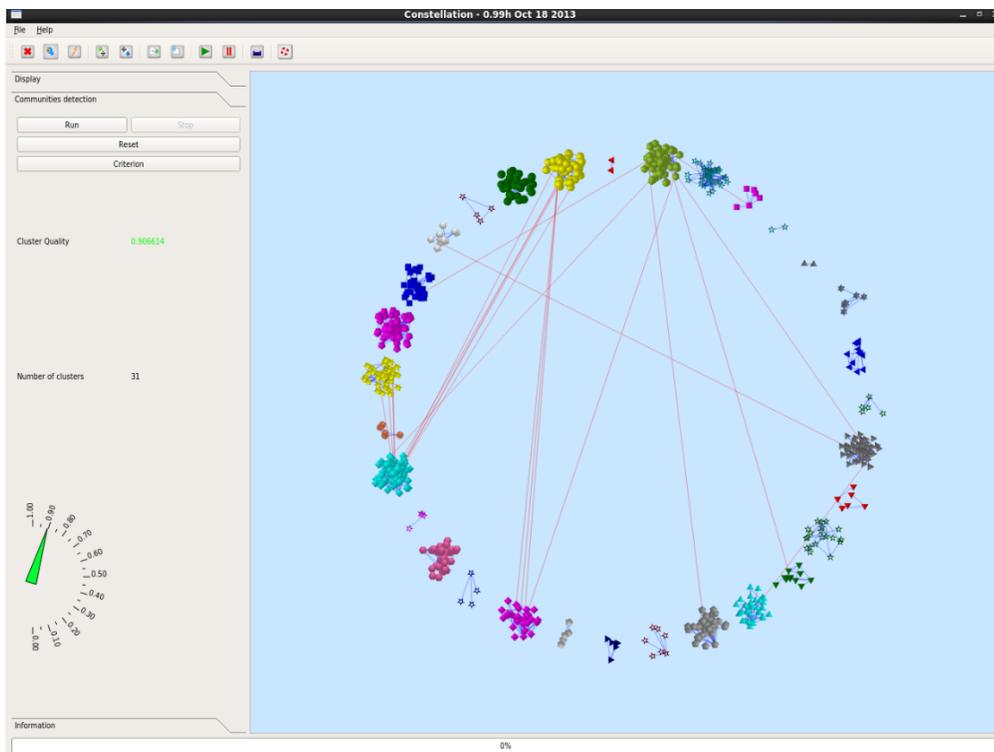


Figure 9 – Graphe des collaborateurs de 2010 à 2013