MATHÉMATIQUES ET SCIENCES HUMAINES

MONIQUE DALUD-VINCENT

Graphes et classification : l'exemple des tables de mobilité sociale

Mathématiques et sciences humaines, tome 147 (1999), p. 47-70

<http://www.numdam.org/item?id=MSH_1999__147__47_0>

Math. Inf. Sci. hum. (37e année, n° 147, *1999*, pp. 47-70)

GRAPHES ET CLASSIFICATION : L'EXEMPLE DES TABLES DE MOBILITE SOCIALE

Monique DALUD-VINCENT[1]

RÉSUMÉ – *L'objectif est de mettre en évidence, à partir d'un tableau de contingence croisant 2 variables utilisant la même nomenclature, une typologie des catégories de cette nomenclature. La recherche de cette typologie est basée sur l'hypothèse selon laquelle il existe des groupes de catégories en fonction des attractions entretenues entre elles ainsi que de leurs enchaînements. On s'appuie sur une modélisation sous forme de graphes et sur une méthode de décomposition des composantes (fortement) connexes.*

MOTS-CLÉS – Classification, tableau de contingence, graphe, composante (fortement) connexe, mobilité sociale.

SUMMARY – Graphs and clustering : the case of social mobility

The aim of this paper is to build a categories typology from a square contingency table containing variables using the same nomenclature. Our assumption is that it exists categories groups according to attractions between categories and sequence of attractions among them. We use a modelling based on graphs and a decomposition method of connected (strong) component.

KEYWORDS – Clustering, square contingency table, graph, connected (strong) component, social mobility.

INTRODUCTION

L'objectif de cet article est de présenter une méthode de classification ayant abouti à l'élaboration du logiciel RESO[2] . Cette méthode est formalisée en Théorie des Graphes. Elle propose une décomposition des composantes (fortement) connexes d'un graphe.

On se propose, du même coup, de montrer comment l'outil peut être appliqué partant d'un tableau de contingence croisant deux variables admettant les mêmes modalités.

Cette présentation s'appuiera sur un exemple concret portant sur des données sociologiques, à savoir les tables de mobilité sociale.

---

[1] Université Lumière Lyon II, Faculté de Sociologie, 5 avenue Pierre Mendès France, C.P. 11, 69676 BRON Cedex., e-mail : Monique.Vincent-Dalud@univ-lyon2.fr

[2] Ce logiciel sera prochainement accessible sur site WEB.

Notre démarche est donc développée en plusieurs parties, les deux premières (présentation des données et modélisation) ayant plus pour objectif de montrer comment et pourquoi un graphe peut être construit sur la base d'un tableau de contingence, les deux autres (méthode et résultats) ayant plus pour but de décrire l'outil RESO et les résultats auxquels il aboutit.

1. LES DONNÉES

L'INSEE, au travers des enquêtes «Formation Qualification Professionnelle» (FQP) , donne régulièrement[3] des résultats dans le domaine dit de la «mobilité sociale» sous la forme de tables donnant le croisement entre l'origine sociale de l'enquêté (repérée le plus souvent par la Profession et Catégorie Socio-professionnelle (PCS) du père) et la position de l'enquêté lui-même (repérée par sa Profession et Catégorie Socio-professionnelle). Ces tables apparaissent donc comme des tableaux de contingence particuliers puisque la même nomenclature est utilisée pour les deux variables croisées.

	Catégorie du fils						
Catégorie du père	Agriculteurs exploitants	Artisans, commerçants, chefs d'entreprise	Cadres, professions intellectuelles supérieures	Professions intermédiaires	Employés	Ouvriers	Ensemble
Agriculteurs exploitants	33,83	8,78	5,04	11,96	6,75	33,63	100
Artisans, commerçants chefs d'entreprise	1,96	29,00	19,57	19,24	7,19	23,04	100
Cadres, professions intellect. supérieures	0,48	9,21	59,77	20,74	5,97	3,84	100
Professions interméd.	0,12	9,96	31,83	31,26	8,85	17,99	100
Employés	0,31	9,73	22,83	31,66	13,94	21,53	100
Ouvriers	1,42	9,84	7,66	21,95	10,22	48,91	100
Inconnue	1,00	10,03	6,84	16,95	12,24	52,94	100
Ensemble	8,37	12,60	15,43	20,73	9,02	33,85	100

Tableau 1 : Catégorie socio-professionnelle (ou dernière catégorie socio-professionnelle d'actif) du fils en fonction de celle du père. Hommes français de naissance, actifs ou anciens actifs, âgés de 40 à 59 ans. Tableau des destinées (d'après Gollac, Laulhé, Soleilhavoup, 1988).

[3] Les enquêtes ont eu lieu en 1953, 1964, 1970, 1977, 1985 et 1993. Les tableaux de l'enquête de 1985 sont rassemblés dans l'ouvrage de M. Gollac, P. Laulhé et J. Soleilhavoup (Gollac, Laulhé, Soleilhavoup, 1988).

A titre d'exemples, nous donnons les résultats de l'enquête de 1985 avec 6 catégories dans les Tableaux 1 et 2 concernant respectivement les enquêtés hommes nés français, actifs ou anciens actifs, de 40-59 ans et les enquêtées femmes françaises de naissance de 40-59 ans.

	Catégorie de la fille						
Catégorie du père	Agriculteurs exploitants	Artisans, commerçants, chefs d'entreprise	Cadres, professions intellectuelles supérieures	Professions intermédiaires	Employés	Ouvriers	Ensemble
Agriculteurs exploitants	34,72	8,49	1,29	9,85	30,87	14,79	100
Artisans, commerçants chefs d'entreprise	2,40	21,36	8,79	21,72	37,51	8,23	100
Cadres, professions intellect. supérieures	0,00	5,99	25,30	35,31	28,34	5,06	100
Professions interméd.	0,48	7,15	10,73	31,24	38,02	12,37	100
Employés	0,90	10,45	4,03	20,33	50,96	13,32	100
Ouvriers	2,25	8,66	2,06	11,03	47,48	28,52	100
Inconnue	3,48	5,14	1,25	11,92	63,60	14,61	100
Ensemble	8,65	10,12	4,96	16,18	41,64	18,46	100

Tableau 2 : Catégorie socio-professionnelle (ou dernière catégorie socio-professionnelle d'active) de la fille en fonction de celle du père. Femmes françaises de naissance, âgés de 40 à 59 ans. Tableau des destinées (d'après Gollac, Laulhé, Soleilhavoup, 1988).

Pour les deux tables données en exemple, le test du Khi-deux amène à refuser l'hypothèse d'indépendance avec un seuil de 1 %.

2. MODÉLISATION SOUS FORME DE GRAPHE

S'agissant d'un tel tableau de données, la Théorie des Graphes[4] permet une modélisation sous la forme suivante :

– L'ensemble des sommets est défini par l'ensemble des PCS de la nomenclature utilisée (par exemple, en 32 catégories pour l'enquête de 1985),

– La relation sur cet ensemble est définie par : une catégorie X est en relation avec une catégorie Y si le profil-ligne dans la case (X,Y) est supérieur au profil-ligne moyen correspondant en colonne Y. Le sociologue met alors l'accent sur les cases marquant une attraction entre une modalité ligne (une origine) et une modalité colonne (une PCS). Cette opération revient à considérer qu'il existe un lien entre une catégorie X et une catégorie Y si l'écart à l'indépendance dans la case (X,Y) est positif[5].

Les tableaux 1 et 2 ci-dessus amènent alors aux relations suivantes (les boucles n'ont pas été représentées ; elles seraient systématiques sauf pour la PCS inconnue où elles n'apparaîtraient pas) :

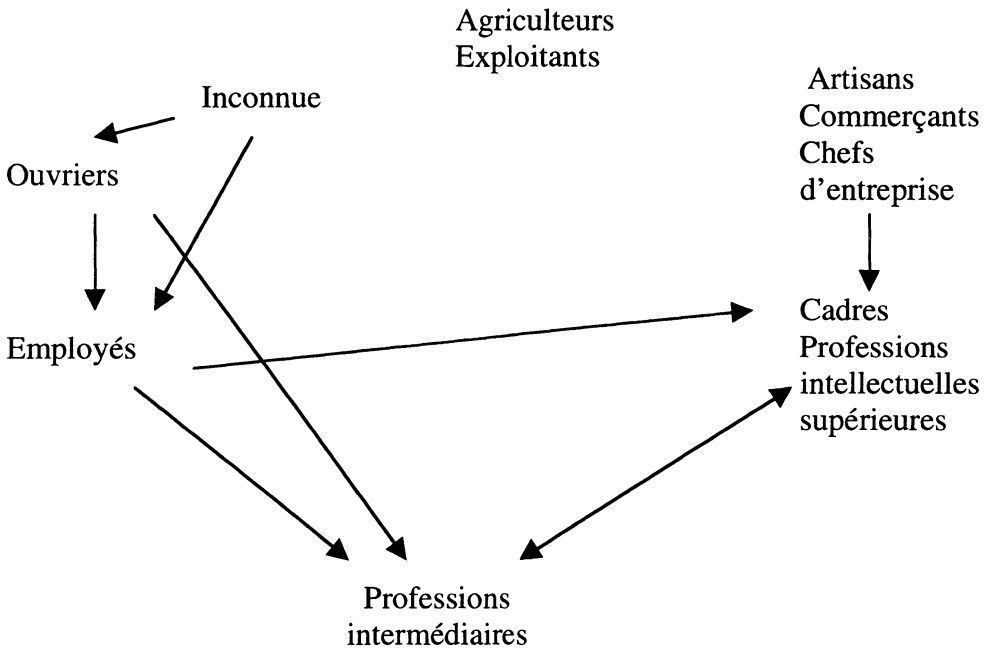


Figure 1 : Graphe des écarts à l'indépendance positifs pour le Tableau 1 (d'après Gollac, Laulhé, Soleilhavoup, 1988)

---

[4] Les notions de graphe, sous-graphe, composante (fortement) connexe, graphe réduit ne sont pas rappelés dans l'article. Le lecteur pourra trouver leur définition dans l'ouvrage de C. Berge (Berge, 1983).

[5] En effet, avec les notations habituelles, $\frac{n_{ij}}{n_{i.}} \rangle \frac{n_{.j}}{N}$ est équivalent à $n_{ij} - \frac{n_{i.}n_{.j}}{N} > 0$ ce qui indique que l'écart à l'indépendance est positif.

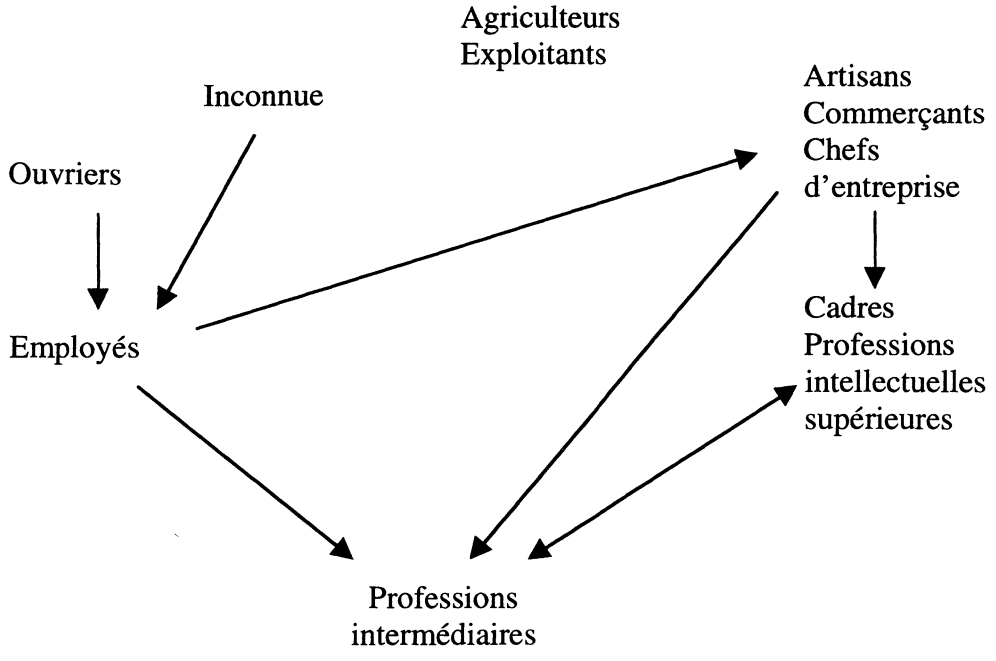


Figure 2 : Graphe des écarts à l'indépendance positifs pour le Tableau 2 (d'après Gollac, Laulhé, Soleilhavoup, 1988)

Ce type de modélisation[6], jamais utilisé à notre connaissance, offre plusieurs avantages :

En premier lieu, sa représentation est facile à lire et met l'accent sur les attractions. Le sociologue peut ainsi voir par exemple, à l'aide de la Figure 1, que les hommes de 40-59 ans sont plus souvent cadres que dans la population masculine de 40-59 ans dans son ensemble lorsque leur origine est artisan, commerçant, chef d'entreprise ou cadre ou profession intermédiaire ou encore employé. La Figure 2 indique que, pour les femmes, l'origine employé ne donne pas lieu à une attraction avec la destinée cadre. Aussi, sur les 2 graphiques on note que la PCS agriculteur ne fait l'objet d'aucune attraction si l'on excepte la boucle : les fils (ou filles) d'agriculteurs deviennent uniquement agriculteurs (agricultrices) en proportion plus importante que la population de référence. Notons que cette modélisation sera d'autant plus pratique que la nomenclature utilisée sera détaillée puisqu'elle évite une lecture par case de la table.

En deuxième lieu, cette modélisation est un bon résumé du contenu de la table en ce sens qu'elle conserve une part importante du Khi-deux pour une densité relativement peu importante du graphe. En effet, si l'on somme les contributions au Khi-deux de l'ensemble des attractions repérées dans la table et que l'on divise cette somme par le Khi-deux total de la table, on obtient 72,54 % du Khi-deux pour le Tableau 1 et 74,67 % du Khi-deux pour le Tableau 2 avec une densité du graphe de 30,61 % pour le premier tableau et 28,57 % pour le deuxième. Cette remarque reste vraie si la nomenclature utilisée est plus détaillée. Par exemple, si l'on choisit une nomenclature

[6] On pourrait proposer de construire le graphe à partir d'un critère plus élaboré comme par exemple le coefficient de reproduction (Goux et Maurin, 1997).

en 32 catégories (d'après Gollac, Laulhé, Soleilhavoup, 1988), les mêmes données amènent respectivement à 79,46 % du Khi-deux avec une densité de 33,88 % et 77,65 % du Khi-deux avec une densité de 32,32 %.

En troisième lieu, cette représentation évite des interprétations erronées, notamment lorsqu'il s'agit de comparer plusieurs tables, car elle tient compte des marges de la table. Ainsi, par exemple, Y. Lemel (1991) propose une représentation sous forme de graphe du tableau des destinées qui est la suivante :

Chaque case représente une catégorie.

«*Les flèches partant d'une case représentent les destins possibles pour des hommes de 40 à 60 ans dont les pères appartiennent (ou appartenaient) à la catégorie socio-professionnelle représentée par la case. (...) Tous les mouvements ne sont pas représentés : n'ont été retenus que ceux qui atteignaient une importance minimum. De plus, l'épaisseur des flèches permet de hiérarchiser les flux par importance. Celle-ci est mesurée par la proportion des personnes concernées dans la case*» (Lemel, 1991, p. 154).

Ce principe, appliqué aux deux exemples précédents, donne les résultats suivants (les boucles n'ont pas été représentées) :

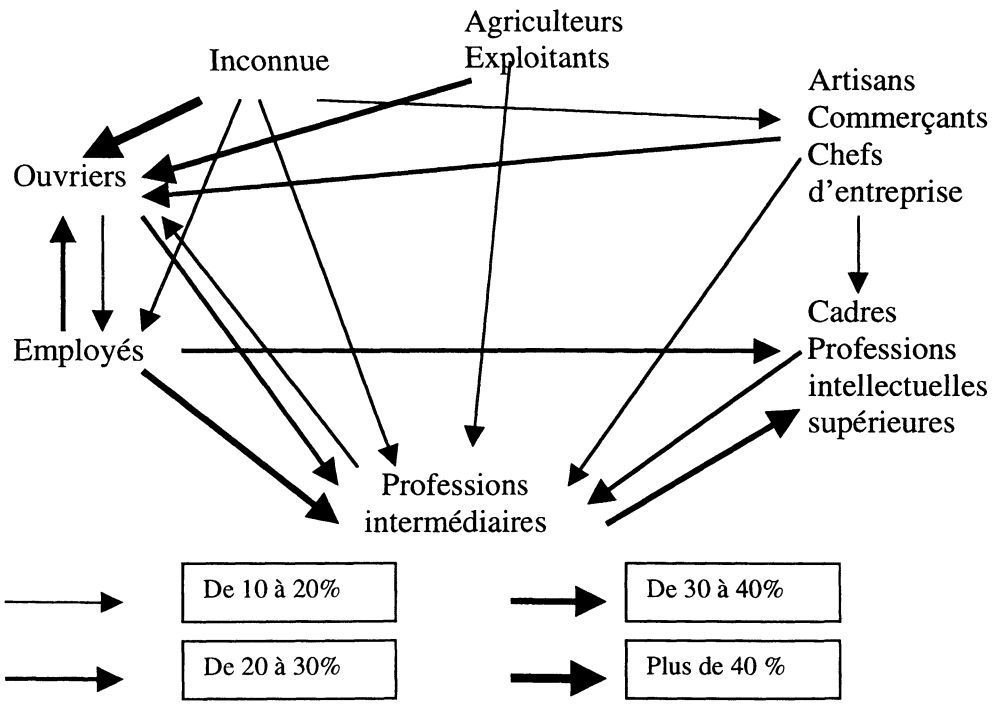


Figure 3 : Les flux de mobilité pour les hommes de 40-59 ans d'après le principe de Y. Lemel et les données FQP de 1985 (d'après le Tableau 1)

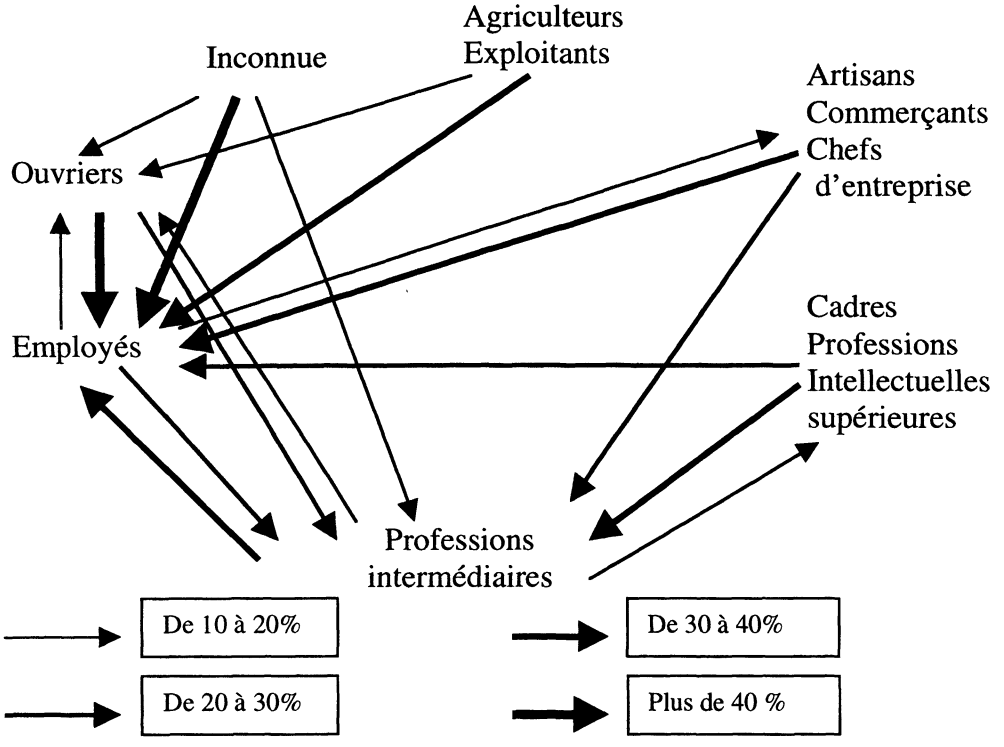


Figure 4 : Les flux de mobilité pour les femmes de 40-59 ans d'après le principe de Y. Lemel et les données FQP de 1985 (d'après le Tableau 2)

A la lecture de ces graphiques, on peut constater, par exemple, que dans le cas des hommes de 40-59 ans 6 PCS origines sur 7 (et 5 sur 7 dans le cas des femmes) se dirigent vers la destinée ouvriers ce qui s'explique par le fait, non visible sur les graphiques, qu'il y a une part importante d'enquêtés relevant de cette catégorie (33,85 % des hommes et 18,46 % des femmes). En réalité, il n'y a pas forcément attraction entre toutes ces origines et la catégorie ouvriers. En fait, le schéma souffre de la non prise en compte du profil-moyen et prend pour seuils des valeurs absolues (i.e. indépendantes du profil-moyen) et non relatives (dans la population enquêtée, il y a en priorité des ouvriers puis des professions intermédiaires,... dans le cas des hommes ; il y a en priorité des employés, des ouvriers, ... dans le cas des femmes). Au contraire, les Figures 1 et 2 montrent bien que, dans le cas des hommes, seules les catégories inconnue et ouvriers dépassent le seuil moyen et sont donc en attraction avec la destinée ouvriers et que, dans le cas des femmes, seule l'origine ouvrière est dans cette situation.

Ce type de représentation conduit donc à des effets pernicieux puisqu'une PCS a d'autant plus de chances de «recevoir une flèche» et de la «recevoir avec une forte intensité» sur le graphique qu'elle contient une forte proportion d'enquêtés. Pourtant, un faible profil-ligne non représenté sur le graphique (car inférieur à 10 %) peut être révélateur d'une attraction. C'est le cas de la case intersection entre l'origine artisans et

la catégorie cadres pour la population féminine (avec 8,79 % contre 4,96 % en profil-moyen). On peut, dans l'autre sens, constater un fort profil-ligne (représenté sur le schéma) sans qu'il y ait attraction entre l'origine et le débouché correspondant à la case. C'est le cas pour les origines agriculteurs, artisans, professions intermédiaires, employés et le débouché ouvriers dans le cas des hommes par exemple.

De la même manière, on rencontre un problème de comparabilité des tables[7] lorsque l'on utilise la représentation proposée par Y. Lemel. En effet, par exemple, sur les Figures 3 et 4, on voit que les employés reçoivent 7 flèches (dont la boucle) pour les femmes et seulement 3 (dont la boucle) pour les hommes. On pourrait croire que les attractions sont plus nombreuses dans le cas de la population féminine. A l'aide des Figures 1 et 2, qui tiennent compte des profil-moyens (il est passé dans ce cas de 41,64 % pour les femmes à 9,02 % pour les hommes), on voit que cette catégorie reçoit autant de flèches dans les deux situations.

En résumé, le type de représentation proposé par Y. Lemel ne permet pas de constater la «fluidité relative» tenant compte de la répartition des enquêtés.

En dernier lieu, remarquons que l'on peut affiner la modélisation que nous avons construite. En effet, il est clair que les attractions ne sont pas toutes de même intensité ; elles n'ont pas toutes contribué de la même façon au Khi-deux. On peut donc proposer de choisir un seuil en dessus duquel on conservera les attractions et en dessous duquel on les éliminera. Pour cela, on peut utiliser les paramètres de position tels que les quartiles, déciles, ... qui permettent de sérier les attractions en fonction de leur contribution au Khi-deux. Les exemples des Tableaux 1 et 2 ne permettent pas de montrer l'intérêt de cette démarche car les attractions sont peu nombreuses. Par contre, cette possibilité reste intéressante si la nomenclature est plus détaillée. Ainsi, avec 32 catégories, sur les mêmes données que celles des Tableaux 1 et 2, nous avons utilisé les déciles de manière à construire 10 graphes. Le premier d'entre eux nommé R0 est celui qui tient compte de toutes les attractions ; le deuxième R1 est le graphe R0 auquel on ôte les 10 % d'attractions les plus faibles au regard de leur contribution au Khi-deux ; le troisième R2 est égal à R0 moins les 20 % d'attractions les plus faibles selon le même critère ; ... R9 est le dernier graphe construit. Il contient, selon la logique suivie, les 10 % d'attractions les plus fortes. Les tableaux suivants indiquent, en définitive, pour chacun des 10 graphes retenus, le nombre de liens, la densité du graphe, la part du Khi-deux conservée :

---

[7] Le problème de comparabilité porte ici non pas sur l'ensemble des conditions de production des données (qui sont supposées comparables dans cet article qui met l'accent sur la méthode ce qui ne veut pas dire qu'elles ne sont pas discutables) mais uniquement sur la modélisation et ce qu'elle fait subir aux données.

Graphe	Nombre de liens	%	densité du graphe	% du Khi-deux conservé	nombre moyen de liens émis ou reçus
R0	347	100 %	33,88 %	79,46 %	10,84
R1	312	89,91 %	30,46 %	79,45 %	9,75
R2	277	79,83 %	27,05 %	79,38 %	8,66
R3	242	69,74 %	23,63 %	79,17 %	7,56
R4	208	59,94 %	20,31 %	78,66 %	6,5
R5	173	49,85 %	16,89 %	77,58 %	5,41
R6	138	39,77 %	13,47 %	75,82 %	4,31
R7	104	29,97 %	10,15 %	73,2 %	3,25
R8	69	19,88 %	6,74 %	68,79 %	2,16
R9	34	9,8 %	3,32 %	60,95 %	1,06

Tableau 3 : Caractéristiques de chaque graphe pour les hommes de 40-59 ans d'après l'enquête FQP de 1985.

Graphe	Nombre de liens	%	densité du graphe	% du Khi-deux conservé	nombre moyen de liens émis ou reçus
R0	331	100 %	32,32 %	77,65 %	10,34
R1	297	89,73 %	29 %	77,62 %	9,28
R2	264	79,76 %	25,78 %	77,52 %	8,25
R3	231	69,79 %	22,56 %	77,25 %	7,22
R4	198	59,82 %	19,33 %	76,67 %	6,19
R5	165	49,85 %	16,11 %	75,63 %	5,15
R6	132	39,88 %	12,89 %	73,74 %	4,12
R7	99	29,91 %	9,67 %	70,74 %	3,09
R8	66	19,94 %	6,44 %	65,36 %	2,06
R9	33	9,97 %	3,22 %	56,05 %	1,03

Tableau 4 : Caractéristiques de chaque graphe pour les femmes de 40-59 ans d'après l'enquête FQP de 1985.

On s'aperçoit que, dans les deux cas, même si le nombre de liens retenus et donc la densité du graphe diminuent fortement en passant de R0 vers R9, la part du Khi-deux conservée, elle, ne diminue pas fortement. Avec 10 % des attractions et une densité du graphe d'environ 3 %, on conserve près de 61 % du Khi-deux pour les hommes et 56 % pour les femmes.

3. LE PROBLÈME ET LA MÉTHODE RESO

Il s'agit, maintenant, de s'interroger sur l'existence de groupes de catégories au sein desquels les attractions s'opèrent. Cette modélisation doit permettre de mettre en évidence une (voire plusieurs) structure(s)[8] des catégories répondant ainsi à l'hypothèse sociologique selon laquelle les différentes catégories n'échangent pas des attractions selon la loi du pur hasard mais qu'au contraire, on doit pouvoir regrouper les catégories en fonction des proximités entretenues entre elles et lisibles au travers de ces échanges.

En résumé, la question est : existe-t-il une typologie des catégories telle qu'à l'intérieur de chaque groupe les liens soient «fluides» [9] ?

Cette modélisation doit permettre aussi une comparaison[10] des résultats obtenus sur la base de plusieurs tables de mobilité (à différentes dates, à des âges différents, ...). En effet, comme le graphe construit tient compte des marges, il rend comparables les différentes structures des catégories mises en évidence.

Du point de vue de la Théorie des Graphes, la question que nous nous posons peut être traduite et adaptée, dans un premier temps, sur la base d'une recherche des sous-graphes pleins[11] maximaux (au sens de l'inclusion). En effet, ces sous-ensembles peuvent permettre d'aborder ce que l'on pourrait nommer la fluidité stricte (chaque catégorie est en attraction avec toutes les autres appartenant au même sous-ensemble qu'elle).

En pratique, cette première tentative[12], appliquée aux données de 1985 classées en 32 catégories, montre qu'il existe 42 sous-graphes pleins maximaux de 2 catégories ou plus (dont 6 contenant 4 catégories, 24 contenant 3 catégories et 12 contenant 2 catégories) pour la population féminine à partir de R0 et 36 sous-graphes pleins maximaux de 2 catégories ou plus dont 6 contenant 4 catégories, 15 contenant 3 catégories et 15 contenant 2 catégories) pour la population masculine en R0 également.

---

[8] Ce mot n'est pas employé au sens que lui donne la théorie dite structuraliste mais au sens moins fort signifiant que l'ensemble des catégories est organisé, structuré.

[9] Certains sociologues dissocient la mobilité dite structurelle de la mobilité dite nette, cette dernière permettant seule de mesurer la " fluidité sociale" c'est-à-dire la mobilité indépendante des évolutions structurelles (voir, entre autres, à ce propos (Merllié, 1994, Cuin,1993, Thelot,1982). Cette distinction fait, à notre sens, problème du point de vue de l'interprétation. Nous rejoignons R. Pohl, J. Soleilhavoup, J. Ben Rezigue qui parlent de " *commodité méthodologique* " et qui écrivent : " *Dans l'enquête FQP, on ne peut distinguer par exemple les fils d'agriculteurs qui ont dû quitter l'exploitation familiale sous la pression des contraintes économiques de ceux qui, quelles que fussent les circonstances et, en raison de leurs préférences et de leurs aptitudes personnelles, auraient de toute manière exercé une autre activité professionnelle que celle de leur père.*" (Pohl, Soleilhavoup, Ben Rezigue, 1982, p. 13). En conséquence, nous employons le terme de fluidité sans pour autant suggérer ce qui relèverait de la liberté de l'individu. Dans notre esprit, la notion d'attraction reflète à la fois les contraintes que subissent les individus et leur liberté.

[10] Dans la mesure où l'on admet la comparabilité des conditions de production des données (méthode de sondage, ...).

[11] Un graphe est dit plein si tous les arcs (dans notre cas, toutes les attractions) possibles existent entre les sommets (dans notre cas les catégories), boucles comprises. Dans notre cas, nous avons utilisé cette notion sans nous soucier des boucles, l'important étant de cerner plutôt les échanges entre catégories distinctes.

[12] Les résultats sont donnés en Annexe 1.

Seules les catégories clergé et inconnue n'appartiennent à aucun de ces sous-ensembles dans le cas des hommes ; dans le cas des femmes, il faut ajouter les ouvriers agricoles. D'un point de vue qualitatif, pour les deux tables étudiées, on peut constater que pour les sous-graphes pleins maximaux contenant 4 catégories les PCS les plus fréquemment présentes sont des catégories de cadres et professions intellectuelles supérieures (notamment les cadres de la fonction publique) et des catégories de professions intermédiaires (toutes sauf le clergé et la catégorie contremaître). Aucune catégorie d'agriculteurs , aucune catégorie d'ouvriers (pour la population féminine, aucune catégorie d'artisans, commerçants également) n'y apparaît. Malgré ces constats, on peut remarquer que les 6 sous-ensembles obtenus ne sont pas identiques pour les deux populations (cf. l'annexe 1).

Le même travail à partir de R5 (soit la moitié des attractions) montre une diminution très nette du nombre de sous-graphes pleins maximaux de 2 catégories ou plus puisqu'on en compte 15 pour la population masculine (dont 6 contenant 3 éléments et 9 en contenant 2) et 13 pour la population féminine (dont 1 contenant 4 éléments et 12 en contenant 2). Les ensembles comptant 3 éléments ou plus contiennent dans les deux cas des catégories de cadres auxquelles on ajoute les instituteurs. Dans le cas des hommes, on trouve également un ensemble formé des catégories d'agriculteurs.

Si l'on pousse l'investigation avec l'étude de R8, on ne repère plus aucun sous-graphe plein maximal de 3 éléments ou plus dans les deux populations.

Nous pourrions, bien sûr, analyser ces résultats plus en détail en reprenant chaque sous-graphe obtenu, ... Mais ce qui nous semble important, du point de vue de la méthode, tient dans les deux remarques suivantes :

- Tout d'abord, l'utilisateur n'obtient pas, dans le cas général et plus particulièrement dans les exemples que nous avons traités, une partition de l'ensemble des catégories. Nous avons obtenu, en effet, des sous-graphes dont l'intersection n'est pas vide.
- Ensuite, la recherche des sous-graphes pleins maximaux portent sur les liens directs entre catégories ; elle ne tient pas compte des liaisons indirectes.

En conséquence, la notion de fluidité stricte peut être affaiblie afin d'obtenir une partition des catégories basée non seulement sur les liens directs mais aussi sur les liens indirects entre catégories. Nous pouvons par exemple proposer la recherche des composantes connexes (ou fortement connexes représentées par le graphe réduit). Remarquons tout de suite que, par définition, tous les éléments d'un même sous-graphe plein maximal appartiennent à la même composante fortement connexe (et donc à la même composante connexe également). Il s'agit donc bien d'affaiblir la notion précédente.

Les données dont nous disposons montrent l'intérêt de ce deuxième traitement. Nous avons vu, en effet, que les sous-graphes pleins maximaux ne comptent pas plus de 4 catégories dans les deux populations étudiées en R0. Mais si l'on construit le graphe réduit pour la population féminine, on s'aperçoit qu'une seule composante fortement connexe non singleton apparaît. Elle contient toutes les catégories sauf les catégories clergé et inconnue qui sont composantes singletons. Jusqu'en R6, la même décomposition est obtenue. Ainsi, si l'on exclut le clergé et la catégorie inconnue, on sait que jusqu'en R6 il existe une suite d'attractions reliant toute catégorie à toute autre.

Par contre l'interprétation sociologique des liens indirects et donc du graphe réduit est délicate. En effet, les attractions sont notées à un instant T. Ainsi, il est bon de souligner que si une origine notée A est en attraction avec une destinée notée B et que l'origine B est en attraction avec une destinée notée C, on ne peut penser qu'il y a eu (ou qu'il y aura) passage d'enquêtés ayant l'origine A vers la destinée C, ni même qu'un fils d'enquêté ayant l'origine A a eu (ou aura) la destinée C. Chaque catégorie est en fait une origine (pour toute attraction qui en sort) et une destinée (pour toute attraction qui se dirige vers elle) et ce au même moment. Remarquons d'ailleurs que faire l'hypothèse de la stabilité du graphe des attractions dans le temps reviendrait, du point de vue de la sociologie, à donner à cette formalisation une valeur prédictive. On pourrait être tenter de calculer ainsi le nombre de générations qu'il est nécessaire de passer pour aboutir, par attractions successives, à une PCS donnée (par exemple cadre supérieur) connaissant l'origine d'un individu (par exemple ouvrier).

Ceci étant, la recherche des composantes fortement connexes peut être intéressante dans le sens où elle part du principe que s'il existe une suite d'attractions dans les deux sens (aller et retour) entre deux catégories, ces deux catégories sont plus proches entre elles que ne le sont deux autres n'admettant pas une telle suite au moins dans un sens. Autrement dit, deux catégories sont proches s'il existe une suite d'attractions dans les deux sens entre elles. Elles ne le sont pas si le «passage» de l'une à l'autre nécessite le changement d'au moins une répulsion en une attraction (c'est-à-dire qu'un profil-ligne parvienne à dépasser le profil-moyen correspondant alors qu'il est plus petit que lui).

Cette recherche peut apparaître d'autant plus intéressante lorsqu'on est plus exigeant vis-à-vis des attractions retenues (étude de R1, R2, ...). On suppose alors que deux catégories sont proches s'il existe entre elles, dans les deux sens, une suite d'attractions dépassant un certain seuil. Elles ne le sont pas si le «passage» de l'une à l'autre nécessite le changement d'au moins une répulsion en une attraction ou le «passage» par une ou plusieurs attraction(s) plus faible(s) que le seuil fixé.

Nous pourrions là encore nous arrêter plus longuement sur les résultats issus d'un tel traitement. Nous insisterons plus particulièrement sur la remarque suivante pour progresser concernant la méthode. Comme nous l'avons exposé plus haut, la table portant sur la population féminine ne fait état que d'une seule composante fortement connexe si l'on excepte les deux singletons (clergé et catégorie inconnue) et ce jusqu'en R6. On s'aperçoit alors que cette composante ne présente pas pour autant la même structure interne en passant de R0 vers R6. Par exemple, en R0 (mais aussi en R1 et R2), on note la position particulière des agriculteurs sur grande exploitation qui sont un passage obligé pour «passer» (par attractions successives) d'une quelconque catégorie (autre qu'agriculteurs) vers les autres catégories d'agriculteurs. On s'aperçoit aussi, à partir de R4, que toutes les catégories d'une même composante ne tiennent pas la même place : certaines semblent fortement liées (notamment celles qui appartenaient aux sous-graphes pleins maximaux) et d'autres semblent plus en périphéries, moins liées. Du point de vue de la sociologie, ces remarques reposent sur l'hypothèse selon laquelle une composante n'est pas un groupe homogène du point de vue des attractions entretenues entre les catégories et que même si l'on peut «accéder» à toute catégorie partant d'une catégorie donnée, certains accès sont plus difficiles que d'autres car ils nécessitent le passage par une catégorie donnée laquelle peut voir ses attractions reçues et/ou émises changer dans le temps et ainsi compromettre la forte connexité.

Ainsi l'investigation peut être poussée plus loin à l'intérieur de chaque composante. La décomposition *en* composantes peut donc être complétée par une décomposition *des* composantes en tenant compte d'un critère de fragilité des liens. Cet objectif est réalisé dans l'outil RESO. Il est basé sur la notion de point d'articulation en Théorie des Graphes, notion que nous avons précisée dans le cadre de notre méthode, et celle de «point fragile». Nous définissons ces notions de la manière suivante :

Soit G = (X,U) un graphe connexe (respectivement fortement connexe). Nous rappelons que $b \in X$ est point d'articulation de G si $G_{X-\{b\}}$, sous-graphe de G engendré par $X-\{b\}$, n'est pas connexe (respectivement fortement connexe).

1. Soit $b \in X$ avec b point d'articulation de G. Soit k entier naturel non nul.

b est *point d'articulation d'ordre* k *de* G[13] si et seulement si la plus petite des composantes connexes (respectivement fortement connexes) de $G_{X-\{b\}}$ a pour cardinal k.

L'ensemble des points d'articulation d'ordre k de G sera noté PAk(G) et l'ensemble de tous les points d'articulation de G sera noté PA(G).

2. Soit $c \in X$. c est *point fragile de* G si et seulement si il existe $b \in X-\{c\}$ tel que b point d'articulation d'ordre 1 de G et tel que $G_{\{c\}}$, sous-graphe de G engendré par le singleton {c}, composante connexe (respectivement fortement connexe) de $G_{X-\{b\}}$.

L'ensemble des points fragiles de G sera noté PF(G).

On peut dès lors décrire l'algorithme du logiciel RESO de la manière suivante[14] :

- décomposer le graphe G en *CC*[15]
- *classer* séparément[16] toutes les *CC* singletons
- **Tant que** l'ensemble des *CC* non singletons n'est pas vide faire :
  - considérer l'une des *CC*
  - retirer cette *CC* de l'ensemble des *CC*
  - rechercher PA(*CC*), PA1(*CC*) et PF(*CC*)
  - **si** PF(*CC*) est non vide
  - **alors** : *classer* ensemble les éléments de PF(*CC*)
  - **sinon** : - **si** le nombre de points d'articulation d'ordre 1 déjà repérés, non classés et présents dans cette *CC* est strictement positif
    - **alors** : *classer* ensemble ces points d'articulation d'ordre 1

---

[13] Dans le cadre de cette application, nous avons utilisé plus particulièrement les points d'articulation d'ordre 1 car ils mettent en évidence l'isolement d'une catégorie. L'utilisation des points d'articulation d'ordre 2, 3, ... posent inévitablement le problème du seuil au-delà duquel l'utilisateur arrête l'investigation. Mais la méthode pourrait bien sûr être adaptée en fonction des objectifs poursuivis.

[14] Le lecteur désirant plus de détails concernant l'algorithme pourra consulter notre thèse de Doctorat (Dalud-Vincent, 1994) ou un article paru dans la revue Social Networks (Dalud-Vincent, Forse, Auray, 1994).

[15] L'abréviation CC désigne, selon le choix a priori de l'utilisateur, soit le terme Composante Connexe soit le terme Composante Fortement Connexe.

[16] Dans la version originale, l'algorithme classe ensemble les composantes singletons.

- **sinon** : - **si** le nombre de points d'articulation déjà repérés, non classés et présents dans cette *CC* est strictement positif
  - **alors** : *classer* ensemble ces points d'articulation
  - **sinon** : *classer* ensemble les points restants de la *CC*
- considérer le sous-graphe obtenu après retrait des sommets classés
- décomposer en *CC* les sommets restants de la *CC*
- **si** le nombre de *CC* singletons est strictement positif
- **alors** : *classer* ensemble ces *CC* singletons
- inclure chaque *CC* non singleton dans l'ensemble des *CC*

- **Fin Tant que**

Après avoir repéré les différentes composantes, l'algorithme identifie les différentes «périphéries» de chaque composante en progressant vers le «centre» de la composante. Chaque «périphérie» est une classe définie à partir d'un critère de fragilité de la position des sommets dans la composante.

Par exemple, à partir de la Figure 2 précédente et sur la base des composantes connexes, RESO repère deux composantes dont l'une est singleton (il s'agit de la catégorie agriculteur que le logiciel classe séparément). RESO recherche alors les points fragiles et les points d'articulation de la composante contenant les 6 autres catégories. Le graphe suivant met en évidence ces sommets particuliers :

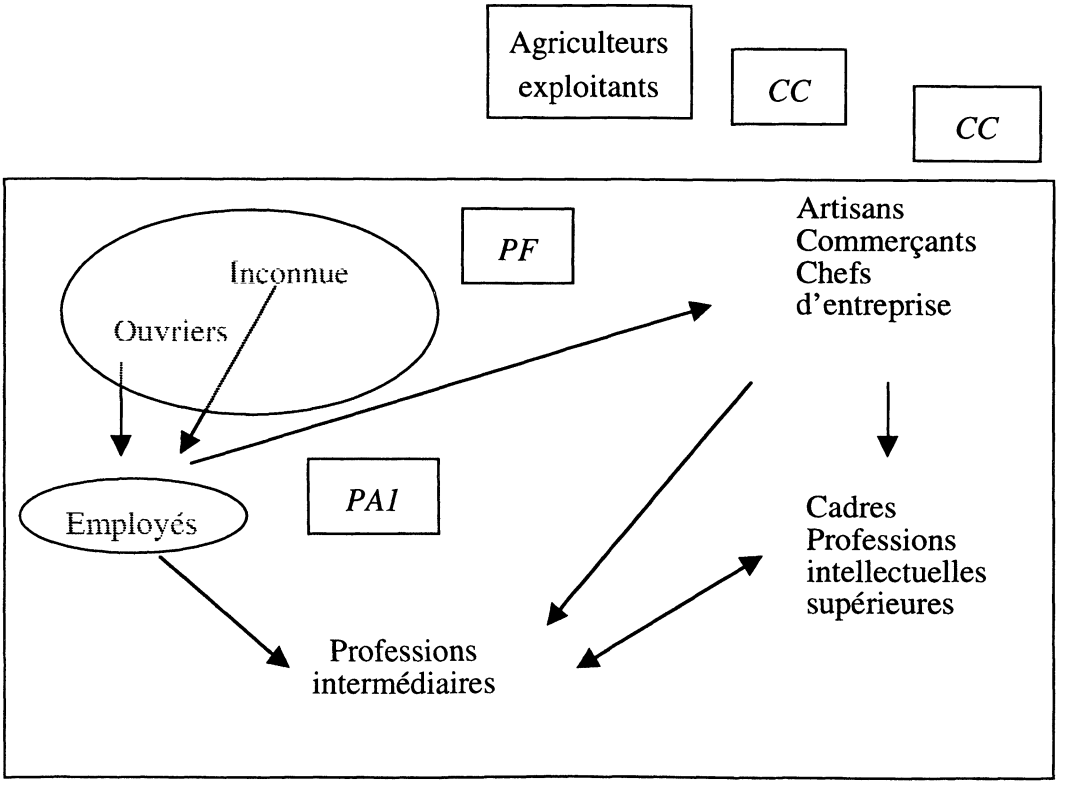


Figure 5 : Composantes connexes (CC), points fragiles (PF) et points d'articulation d'ordre 1(PA1) sur l'exemple de la Figure 2.

En effet, la catégorie ouvrier est point fragile puisqu'il suffit de retirer les employés du graphe pour que cette catégorie se retrouve isolée (i.e. composante singleton). De ce fait, la catégorie employé est dite point d'articulation d'ordre 1. L'algorithme classe donc ensemble les catégories inconnue et ouvrier et les retire du graphe. Une seule composante connexe apparaît alors qui est représentée par la figure suivante :

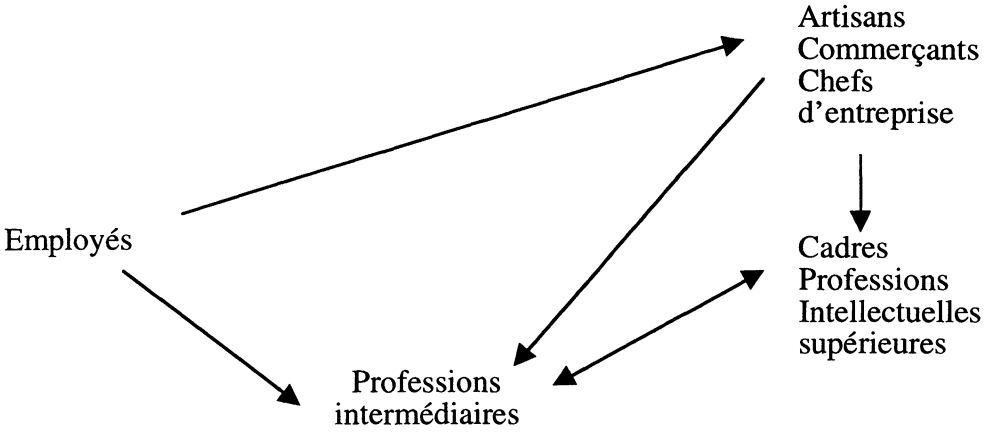


Figure 6 : Composante connexe résultant du retrait des points fragiles

L'algorithme ne repère alors ni point fragile, ni point d'articulation. Il classe donc ensemble les catégories ayant déjà été repérées comme points d'articulation d'ordre 1. Il s'agit de la catégorie employé. RESO construit donc le sous-graphe correspondant au retrait de cette catégorie et met en évidence une seule composante connexe non singleton. Aucun point fragile, aucun point d'articulation n'apparaît. L'algorithme classe ensemble les catégories restantes, cette partie centrale étant non décomposable.

On aboutit ainsi à une décomposition en 4 classes avec d'une part les agriculteurs (composante singleton) et d'autre part, l'ensemble des autres catégories séparées en 3 classes (une classe formée des catégories ouvrier et inconnue formant une périphérie ; une classe charnière formée des employés et une classe plus centrale formée des autres catégories).

4. RÉSULTATS

L'application de la méthode aux données de 1985[17] (Gollac, Laulhé, Soleilhavoup, 1988) croisant la PCS du père avec la PCS de l'enquêté a permis de mettre en évidence des typologies différentes selon notamment le sexe de l'enquêté.

En effet, nous avons vu que, pour les femmes, une seule composante fortement connexe non singleton apparaissait jusqu'en R6. Ce n'est pas le cas pour la population masculine qui présente, de R0 à R4, 2 composantes non singletons : le groupe des agriculteurs forme une composante (non décomposable par RESO jusqu'en R8).

De plus, les décompositions des différentes composantes ne sont pas comparables. Par exemple, en R4, pour la population masculine la composante la plus importante présente 4 classes dont 2 classes de points fragiles (contenant le clergé et les ouvriers de type industriel) et une classe de point d'articulation d'ordre 1 (contenant les professeurs et les ouvriers non qualifiés de type artisanal). Le reste de la composante n'est pas décomposable; il compte 22 catégories (cf. Figure 7).

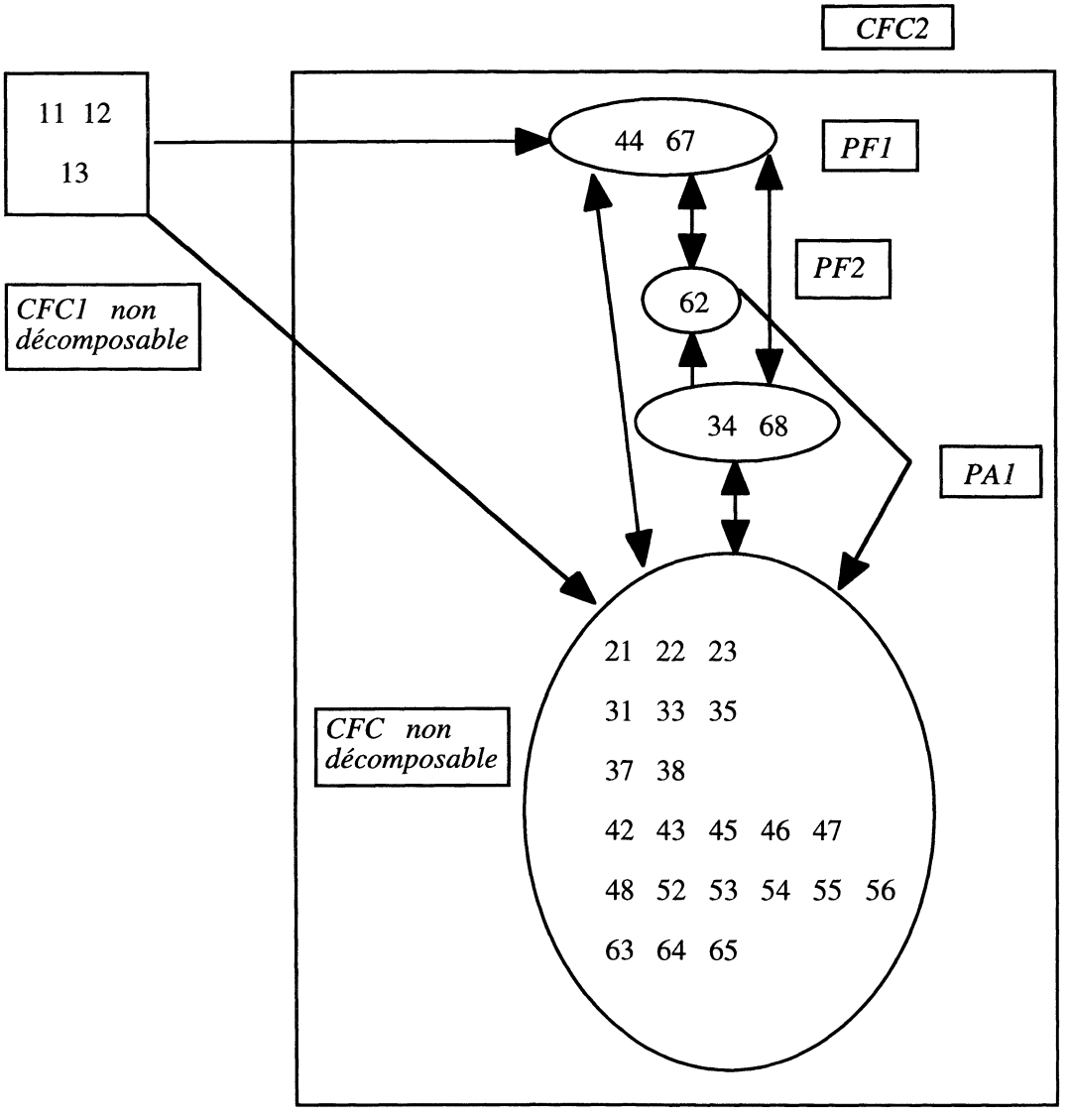
Au même niveau, la population féminine présente une décomposition de son unique composante non singleton (contenant 30 catégories) en 23 classes. Le découpage est beaucoup plus fin que dans le cas des hommes[18]. Il montre une certaine «hiérarchie» des catégories avec en premières périphéries les catégories d'agriculteurs, puis les ouvriers agricoles, les ouvriers de type artisanal, les ouvriers de type industriel, les ouvriers qualifiés de la manutention et chauffeurs. En positions intermédiaires, on trouve les artisans, certaines professions intermédiaires, les employés. Pour finir, le «centre» de la composante est constituée des cadres (sauf les cadres de la fonction publique et les cadres administratifs et commerciaux d'entreprise classés points d'articulation dans la décomposition) et des professions intermédiaires (instituteurs, techniciens, professions intermédiaires de la santé et du travail social).

On peut donc retenir à ce niveau que, l'agriculture étant mise à part, la «fluidité» semble plus importante dans le cas de la population masculine. La décomposition par RESO dépasse-là le constat auquel on est amené par la recherche des composantes qui met seulement en évidence la constitution du groupe des agriculteurs dans le cas de la population masculine et l'isolement des PCS ouvriers agricoles et inconnue dans le cas des hommes et des PCS clergé et inconnue dans le cas des femmes.

---

[17] Nous avons traité les données de 1985 plutôt que celles de 1993 parce qu'elles reposent sur la nomenclature en 32 postes alors qu'en 1993 le découpage n'est pas aussi fin.

[18] Les résultats, ne pouvant faire l'objet d'un graphique suffisamment lisible, sont donnés en Annexe 2.

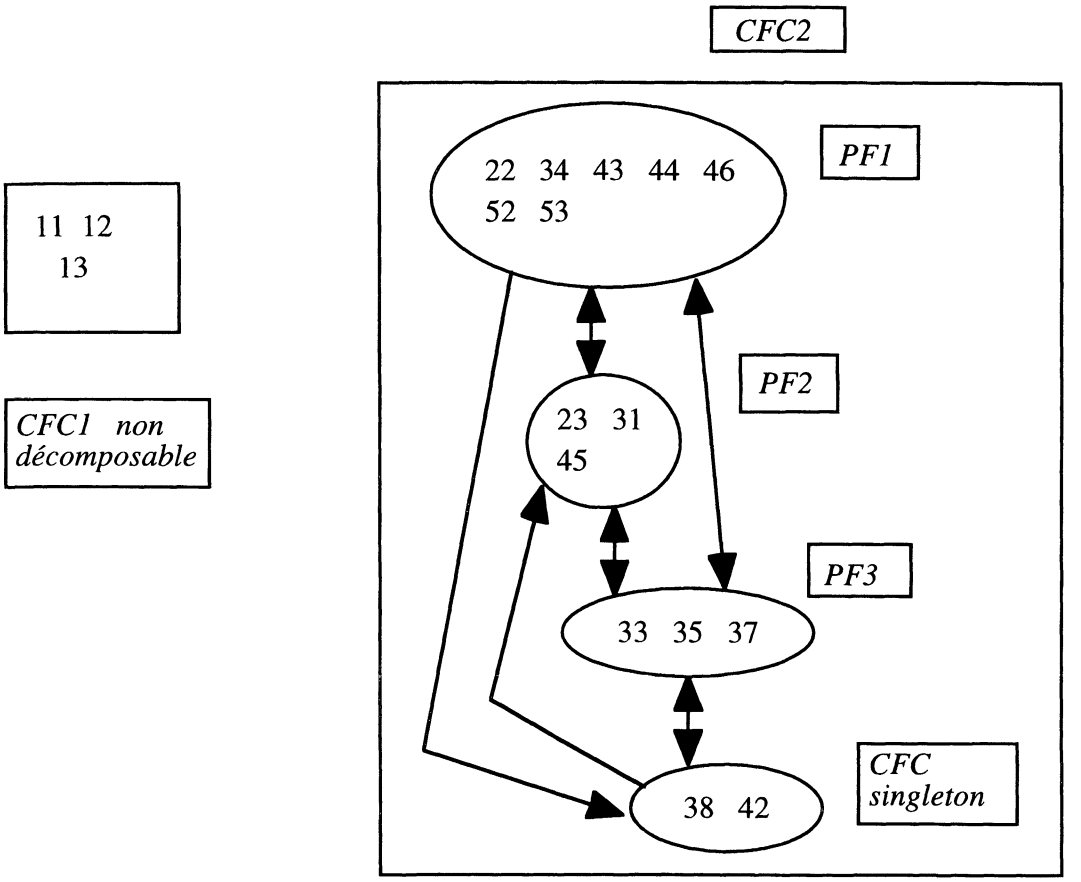


CFC singleton : 70 et 69

Figure 7 : Décomposition de R4 par RESO avec la forte connexité pour les hommes de 40-59 ans[19] .

[19] Nous utilisons les notations suivantes pour toutes les représentations : CFC pour composante fortement connexe ; CFC1, CFC2, ... pour différencier les différentes composantes d'un même graphe ; PA1 pour point d'articulation d'ordre 1 ; PF pour point fragile ; PF1, PF2, ... pour différencier, dans la décomposition d'une composante, le premier ensemble de points fragiles du deuxième puis du troisième, ...Les catégories sont numérotées selon la nomenclature des PCS en 32 postes (cf. Annexe 1). Les composantes fortement connexes non singletons sont représentées par des rectangles et les classes obtenues par RESO à l'intérieur de chaque composante sont représentées par des ellipses. Une flèche entre deux classes distinctes indique qu'il existe au moins une attraction entre une catégorie appartenant à la classe d'où est issue la flèche et une catégorie appartenant à la classe où se dirige la flèche.

L'étude de R7 est également intéressante (cf. Figures 8 et 9). En effet, concernant la population masculine 14 composantes singletons apparaissent contre seulement 4 pour la population féminine. Dans le cas des hommes, il s'agit de toutes les catégories d'ouvriers augmentées entre autres de catégories d'employés, de professions intermédiaires. Ainsi, à ce niveau, la recherche des composantes montre déjà une différence notable dévoilant une «fluidité» plus importante pour la population féminine.

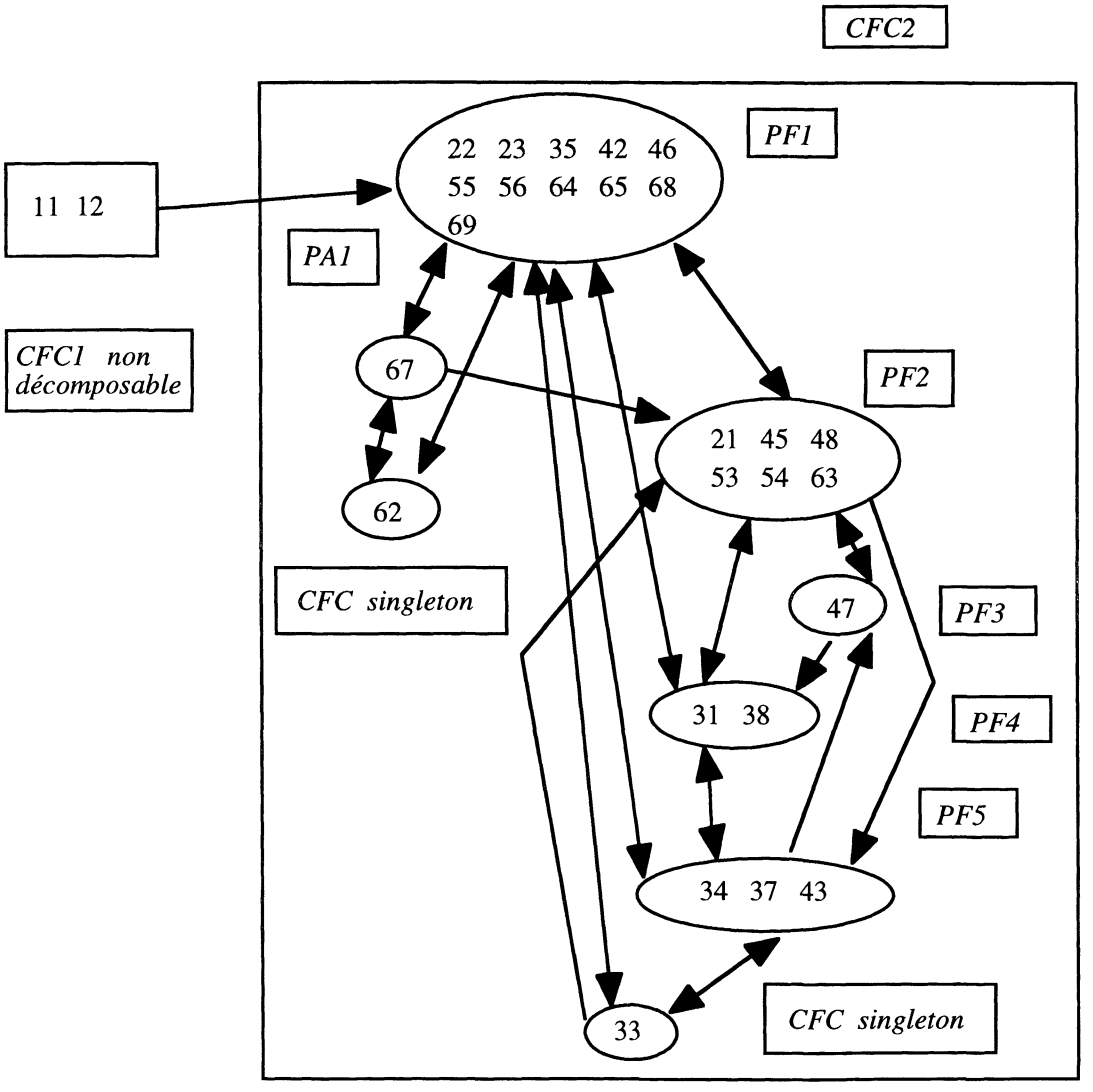


CFC singleton : 21, 47, 48, 54, 55, 56, 62, 63, 64, 65, 67, 68, 69, 70.

Figure 8 : Décomposition de R7 par RESO avec la forte connexité pour les hommes de 40-59 ans.

On compte dans les deux tables 2 composantes non singletons avec dans les 2 cas une composante d'agriculteurs. Dans le cas des hommes, la plus importante des composantes (contenant 15 catégories) comprend toutes les catégories de cadres, des professions intermédiaires (toutes sauf techniciens et contremaîtres), les commerçants, chefs d'entreprise et 2 catégorie d'employés; les cadres et instituteurs gardent une position centrale dans la décomposition. Pour les femmes, la décomposition de la composante la plus importante (contenant 26 catégories) met en évidence un groupe

important de points fragiles (11 catégories dont des ouvriers, des artisans, des professions intermédiaires). On voit ensuite une première composante comprenant les ouvriers de type industriel et une deuxième composante comprenant les cadres (tous sauf les professions de l'information des arts et des spectacles), des professions intermédiaires essentiellement. Il y a donc scission dans la décomposition entre les catégories d'ouvriers et les catégories de cadres.



CFC singleton : 13, 44, 52, 70.

Figure 9 : Décomposition de R7 par RESO avec la forte connexité pour les femmes de 40-59 ans.

Pour conclure, on peut retenir que l'outil RESO est porteur de résultats dans le sens où il permet de s'intéresser à la répartition des attractions entre catégories, l'hypothèse étant que cette répartition dénote une structure des catégories selon un schéma de type centre/périphérie pour chaque contexte mis en évidence et caractérisé par une certaine «fluidité». RESO permet notamment :

- d'évaluer l'importance de la «fluidité» (le nombre de composantes singletons et de composantes non singletons, la délimitation et le nombre des classes obtenues sont des indicateurs de cette importance),
- de dégager une (ou plusieurs) typologie(s) des catégories en mettant en évidence des groupes (comme par exemple agriculteurs/non agriculteurs) mais également des sous-ensembles dans ces groupes (par exemple la distinction entre les catégories d'ouvriers plus périphériques et les catégories de cadres plus au centre),
- de dégager des différences plus fines allant au-delà de l'élaboration d'une typologie grâce à la caractérisation des classes rendue possible par les notions de points fragiles (traduisant une certaine vulnérabilité), de points d'articulation (points charnières), de parties non décomposables (non vulnérables), ...

La représentation graphique des résultats (voire la matrice d'adjacence donnant les attractions entre classes) permet, quant à elle, de visualiser les liens entre les différents groupes obtenus. C'est ainsi que l'on aura pu remarquer l'isolement des agriculteurs en Figure 8 par exemple.

Pour finir, nous pourrions revenir sur tel ou tel résultat ou discuter de la comparabilité préalable des tables compte tenu des conditions de production. Nous avons délibérément mis l'accent sur la méthode RESO afin de montrer en quoi elle peut être un outil pour le sociologue. Utilisé comme procédure de comparaison, RESO semble être relativement bien adapté.

ANNEXE 1

NOMENCLATURE DES CATÉGORIES SOCIO-PROFESSIONNELLES

1 - Agriculteur exploitant
- 11- Agriculteur sur petite exploitation
- 12- Agriculteur sur moyenne exploitation
- 13- Agriculteur sur grande exploitation

2 - Artisan, commerçant, chef d'entreprise
- 21- Artisan
- 22- Commerçant et assimilé
- 23- Chef d'entreprise de 10 salariés ou plus

3 - Cadre, profession intellectuelle supérieure
- 31- Profession libérale
- 33- Cadre de la fonction publique
- 34- Professeur, profession scientifique
- 35- Profession de l'information, des arts et des spectacles
- 37- Cadre administratif et commercial d'entreprise
- 38- Ingénieur et cadre technique d'entreprise

4 - Profession intermédiaire
- 42- Instituteur et assimilé
- 43- Profession intermédiaire de la santé et du travail social
- 44- Clergé, religieux
- 45- Profession intermédiaire administrative de la fonction publique
- 46- Profession intermédiaire administrative et commerciale des entreprises
- 47- Technicien
- 48- Contremaître, agent de maîtrise

5 - Employé
- 52- Employé civil et agent de service de la fonction publique
- 53- Policier et militaire
- 54- Employé administratif d'entreprise
- 55- Employé de commerce
- 56- Personnel des services directs aux particuliers

6 - Ouvrier
- 62- Ouvrier qualifié de type industriel
- 63- Ouvrier qualifié de type artisanal
- 64- Chauffeur
- 65- Ouvrier qualifié de la manutention, du magasinage et du transport
- 67- Ouvrier non qualifié de type industriel
- 68- Ouvrier non qualifié de type artisanal
- 69- Ouvrier agricole

Inconnue (codé 70)

SOUS-GRAPHES PLEINS MAXIMAUX DE 3 CATEGORIES OU PLUS

Population masculine avec 32 catégories (d'après (Gollac, Laulhé, Soleilhavoup, 1988) et la nomenclature ci-dessus).

Pour R0 :

33-34-42-45	48-55-65
33-42-46-43	22-23-31
31-33-35-38	22-31-38
31-33-34-35	22-43-46
34-45-47-53	62-63-65
23-31-34-37	45-55-53
45-47-54	63-64-65
37-43-46	63-64-68
33-38-42	11-12-13
31-37-38	21-63-64
43-46-47	

Pour R5 :

31-33-38	33-38-42
31-35-38	31-37-38
31-33-34	11-12-13

Pour R8 :

Aucun sous-graphe plein maximal de 3 catégories ou plus n'est mis en évidence.

2. Population féminine avec 32 catégories (d'après (Gollac, Laulhé, Soleilhavoup, 1988) et la nomenclature ci-dessus).

Pour R0 :

33-45-46-54
33-34-37-42
31-33-37-43
31-33-37-46
31-33-34-37
33-34-38-42
52-54-47
48-54-63
33-48-54
33-34-45
33-47-54
22-34-55
34-37-55
22-23-31
23-31-46
31-46-54
22-31-34
31-34-35
34-35-42
21-54-47
31-46-47
46-47-54
52-53-65
21-52-55
11-12-13
33-38-43
33-38-47
56-62-67
56-67-68
62-65-67

Pour R5 :

33-34-37-42

Pour R8 :

Aucun sous-graphe plein maximal de 3 catégories ou plus n'est mis en évidence.

ANNEXE 2

Résultats obtenus par RESO avec la forte connexité pour la population féminine - Etude de R4

- CFC singleton : 44-70
- Toutes les autres PCS forment une seule CFC se décomposant comme suit :
  - PF1 : 11
  - PF2 : 12
  - PA1 : 13
  - PF3 : 69
  - PA1 : 63
  - PF4 : 68
  - PA1: 67
  - PF5 : 56
  - PA1 : 62
  - PF6 : 65
  - PF7 : 55-64
  - PA1 : 33-21-45
  - PF8 : 48
  - PA1 : 23

- PF9 : 22-53
- PF10 : 46
- PA1 : 52-37
- PF11 : 54
- PA1 : 42
- PF12 : 35-47
- PF13 : 34
- PF14: 31-38
- CFC singleton : 43

La matrice d'adjacence suivante indique par la valeur 1 s'il y a au moins une attraction entre une catégorie de la classe lue en lignes et une catégorie de la classe lue en colonne, pour tous les couples de classes obtenues par RESO. Les classes ont été numérotées selon leur ordre d'apparition.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
Classe 1 : 44	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
Classe 2 : 70	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
Classe 3 : 11	0	0	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Classe 4 : 12	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Classe 5 : 13	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Classe 6 : 69	0	0	0	0	0	1	0	0	0	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0
Classe 7 : 63	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
Classe 8 : 68	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
Classe 9 : 67	0	0	0	0	0	0	0	1	1	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0
Classe 10 : 56	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
Classe 11 : 62	0	0	0	0	0	0	0	0	1	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0
Classe 12 : 65	0	0	0	0	0	0	0	0	0	0	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0
Classe 13 : 55 64	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
Classe 14 : 33 21 45	0	0	0	0	0	0	1	0	0	0	0	0	1	1	1	0	1	0	1	0	1	1	1	1	1
Classe 15 : 48	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	1	0
Classe 16 : 23	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	1	0	0	0	1	1	0
Classe 17 : 22 53	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0	1	0	0	0	1	0	1	1	1
Classe 18 : 46	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	1	0	1	1	0	0	0	0
Classe 19 : 52 37	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	1	0	1	1	0	1
Classe 20 : 54	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	1	0	0	0	0
Classe 21 : 42	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0
Classe 22 : 35 47	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0	1	1
Classe 23 : 34	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	1	1	0
Classe 24 : 31 38	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	1	0	1	0	1	1	1
Classe 25 : 43	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	1

BIBLIOGRAPHIE

BERGE, C., *Graphes et hypergraphes*, troisième édition, Gauthier-Villars, Paris, 1983.

CUIN, C.H., *Les sociologues et la mobilité sociale,* PUF, Paris, 1993.

DALUD-VINCENT, M., *Modèle prétopologique pour une méthodologie d'analyse de réseaux : concepts et algorithmes,* Thèse de Doctorat, Université Lyon I, 1994.

DALUD-VINCENT, M., FORSE, M., AURAY, J.P., An algorithm for finding the structure of social groups, *Social Networks*, 16, (1994), 137-162.

GOLLAC, M., LAULHE, P., SOLEILHAVOUP, J., Mobilité sociale. Enquête formation qualification professionnelle de 1985, *Les collections de l'INSEE*, série D,126, 1988.

GOUX, D., MAURIN, E., Destinées sociales : le rôle de l'école et du milieu d'origine, *Économie et Statistique*, 306, (1997), 13-25.

LEMEL, Y., *Stratification et mobilité sociale,* A.Colin, Paris, 1991.

MERLLIE, D., *Les enquêtes de mobilité sociale,* PUF, Paris, 1994.

POHL, R., SOLEILHAVOUP, J., BEN REZIGUE, J., Formation, mobilité sociale, salaires, *Les collections de l'INSEE*, série D, 93, 1982.

THELOT, C., *Tel père, tel fils ? Origine familiale et position sociale,* Dunod, Paris, 1982.