

**« MEANS-END CHAIN MODEL » AND AUTOMATIC NETWORK  
DRAWING AS COMMERCIAL AWARENESS TOOLS.**

**Eric Boutin \*, Jean-Marc Ferrandi\*\* , Pierre. Valette Florence\*\*\***

**L'ANALYSE DES CHAINAGES COGNITIFS ET LA CONSTRUCTION  
AUTOMATIQUE DE RESEAUX COMME OUTILS DE VEILLE  
COMMERCIALE.**

\* Dr. Eric Boutin, Le Pont, IUT, TC, BP 132, 83957 La Garde Cedex. France  
boutin@univ-tln.fr

\*\*Dr. Jean Marc Ferrandi, Centre de recherche de l'ESA Grenoble France  
centrerecherch@demeter.fr

\*\*\* Pr. Pierre Valette Florence, Centre de recherche de ESA Grenoble France.  
valette@esa.grenet.fr

**Keywords** network, data analysis, marketing.

**Résumé**

L'objet de cette recherche appliquée est de découvrir les aspirations et les pratiques relatives à la consommation d'énergie. Ces attentes en matière de service énergétique ont été appréhendées à travers la méthode des chaînages cognitifs et déterminées grâce à une technique originale de traitement des données: l'analyse de réseaux. La conclusion dresse le bilan de cette étude, ses apports et limites respectives et dégage un certain nombre de voies de recherche futures.

**Summary**

The aim of this applied research is to discover aspirations and practices linked with energy consumption. These expectations as regards energy services have been apprehended through the « Means-End Chain Model » and determined thanks to network analysis . The conclusion gives an overview of this study and points at various research directions in the future.

# **L'ANALYSE DES CHAINAGES COGNITIFS ET LA CONSTRUCTION AUTOMATIQUE DE RESEAUX COMME OUTILS DE VEILLE COMMERCIALE.**

**Eric Boutin \*, Jean-Marc Ferrandi\*\* , Pierre. Valette Florence\*\*\***

« MEANS-END CHAIN MODEL » AND AUTOMATIC NETWORK DRAWING AS COMMERCIAL AWARENESS TOOLS.

\* Dr. Eric Boutin, Le Pont, IUT, TC, BP 132, 83957 La Garde Cedex. FRANCE  
boutin@univ-tln.fr

\*\*Dr. Jean Marc Ferrandi, ESA Grenoble

\*\*\* Pr. Pierre Valette Florence, ESA Grenoble. valette@esa.grenet.fr

L'objet de ce travail est de porter à la connaissance de la communauté académique une pratique de veille commerciale originale. Cette originalité réside en deux points:

- Le recours à la méthode des chaînages cognitifs comme mode de recueil de l'information.

- L'utilisation de l'analyse en terme de réseaux, comme outil de traitement et de visualisation des données collectées.

Cette méthode a été utilisée par une société de distribution d'énergie d'une ville du Sud de la France qui s'interrogeait sur les attentes et les aspirations de ses clients en matière de consommation d'énergie. L'objectif était d'identifier, au delà de la seule satisfaction d'un besoin primaire, se chauffer, des pratiques de consommation d'énergie spécifiques.

Les choix méthodologiques seront présentés en début de cet article. Les résultats principaux qui découlent des analyses effectuées seront ensuite exprimés. A titre de synthèse, la conclusion dressera le bilan de cette étude, ses apports et limites respectives et dégagera les voies de recherche futures.

## **I. METHODOLOGIE EMPLOYEE**

### **A- L'analyse selon les chaînages cognitifs.**

Introduite en 1988 par Reynolds et Gutman, [GUT 82],[REY et alii 88], l'analyse selon les chaînages cognitifs a connu depuis d'amples développements, tant conceptuels, que méthodologiques. Le lecteur intéressé trouvera dans Valette-Florence [VAL 94] un panorama assez complet des caractéristiques propres à cette démarche. Nous présenterons brièvement ci-dessous ses principales spécificités.

#### **1. Caractéristiques principales de la méthode.**

Tout individu, lorsqu'il consomme un produit (ou un service) cherche à satisfaire des besoins immédiats, parfois primaires, comme la faim ou la soif, directement liés aux caractéristiques propres au produit, mais aussi un ensemble de conséquences pratiques ou plus personnelles qu'il va en retirer (eg: le confort), ainsi que des valeurs plus motrices dans sa vie, telles que par exemple, une vie aisée et confortable ou la sécurité familiale.

Reconstituer ces liens entre attributs, conséquences et valeurs, au niveau d'un même segment de clientèle, s'appelle un chaînage cognitif. Celui-ci représente l'ensemble des liens existants entre certains attributs, conséquences et valeurs, liens qui sont partagés conjointement par plusieurs individus.

En d'autres termes, un chaînage cognitif est la résultante, à un niveau agrégé, des aspirations personnelles (ou attentes) partagées par un même groupe de consommateurs. Ces aspirations individuelles, collectées de manière qualitative et en face à face, s'appellent des chaînes individuelles faisant explicitement le lien entre un attribut d'un produit et les conséquences d'ordre fonctionnel ou psychologique, ainsi que les valeurs qui en découlent. Concrètement la phase d'enquête met à jour, pour chaque individu plusieurs chaînes de type:  $A \rightarrow C_1 \rightarrow C_2 \rightarrow V_1 \rightarrow V_2$ , où A représente l'attribut choisi,  $C_1$  et  $C_2$  deux conséquences qui en découlent, et  $V_1$  et  $V_2$  deux valeurs directement reliées aux deux conséquences précédentes.

Comme le recueil qualitatif de l'information met directement en jeu des processus d'ordre cognitif, le terme de chaînages cognitifs s'est tout naturellement imposé. En résumé, la démarche procède en deux étapes successives:

- tout d'abord, à un niveau individuel, l'expression de plusieurs chaînes individuelles (toutes du type  $A \rightarrow C \rightarrow V$ ) au moyen d'un protocole spécifique de recueil d'information présenté plus loin,

- ensuite, une série d'analyses basées sur la construction automatique de réseaux. Il s'agit d'une démarche originale, présentée plus loin qui a fait l'objet de récents développements.

## **2. Les caractéristiques de la collecte des données.**

Comme nous l'avons souligné dans le paragraphe précédent, l'analyse selon les chaînages cognitifs est une démarche en essence qualitative, bien que les traitements statistiques soient de nature quantitative. De ce fait, la collecte se déroule en deux étapes successives:

- la première, basée sur une réunion de groupe qualitative, est nécessaire pour valider les éléments hiérarchiques (attributs - conséquences - valeurs) qui seront utilisés dans la deuxième phase,

- la seconde est réalisée au moyen d'entretiens individuels en face à face, d'une trentaine de minutes environ, utilisant l'information obtenue à l'étape précédente.

L'originalité de la collecte des données réside dans l'utilisation d'un jeu de « cartes » préétablies sur lesquelles sont notées en clair les différents éléments constituant les listes d'attributs, conséquences et valeurs. L'enquêteur demande tout d'abord au répondant de classer en plusieurs piles (deux ou trois) les attributs suivant qu'il les perçoit comme « important », « moyennement important » ou « pas important » pour caractériser sa pratique de consommation quotidienne. L'enquêteur lui demande alors de sélectionner dans la pile des attributs importants celui qu'il considère comme l'étant le plus. Cet attribut va servir de point « d'ancrage » pour la réalisation de la suite de l'entretien. On demande alors au répondant de sélectionner parmi la pile des conséquences celle(s) (une ou deux, rarement plus) qui sont les plus importantes pour lui et qui découlent directement de l'attribut (ou caractéristique) qu'il vient de sélectionner. Ce ou ces cartes sont donc sélectionnées et mises devant l'enquêté, par exemple  $A_2 - C_1 - C_6$ . On recommence ensuite la même opération avec les valeurs pour aboutir à la première chaîne individuelle du répondant, par exemple  $AC_1 - C_6 - V_2 - V_8$ .

On recommence après cela la même opération avec le deuxième attribut jugé le plus important (dans la première pile des attributs jugés « importants ») et ainsi de suite jusqu'à ce que le répondant ne puisse (ou ne désire) plus évoquer des chaînes individuelles supplémentaires. L'entretien se termine alors avec un court recueil d'informations plus classiques, concernant quelques variables socio-démographiques simples, ainsi que des renseignements caractérisant le

chauffage actuel du répondant et son degré de satisfaction correspondant. Soulignons aussi que lors de l'entretien par cartes, à chaque niveau hiérarchique (attributs, conséquences, valeurs) des cartes vierges sont laissées à la disposition de l'enquêté au cas où celui-ci désirerait noter un élément, non disponible dans la liste, et qu'il jugerait très important pour lui. Cette procédure, qui n'est pratiquement jamais utilisée, comme ici, permet d'éviter de passer à côté d'éléments qui ne seraient pas apparus lors de la phase qualitative préliminaire.

Au total, l'enquête effectuée sur 79 personnes a conduit à l'obtention de 184 chaînes individuelles, faisant les liens entre les attributs recherchés, les conséquences qu'on en retire et les valeurs qu'ils satisfont (soit une moyenne de 2,3 chaînes par personne).

## B. Les réseaux: un outil original d'analyse des chaînages cognitifs.

Notre objectif est de présenter un nouvel outil d'analyse de données: l'analyse de réseaux et d'asseoir sa pertinence [BOUT et alii 96 1]. Cependant, avant d'appliquer cette méthode dans un cadre concret, le comportement en matière d'énergie, nous préciserons ses spécificités par rapport aux techniques classiques d'analyse de données.

### 1. Les spécificités de l'analyse de réseau.

L'analyse en termes de réseaux présente trois spécificités: le point de départ, le point d'arrivée et les présupposés de la méthode.

#### 1- Le point de départ de la méthode.

Le point de départ de la méthode est, dans le cas présent, une succession de n chaînes individuelles exprimées par m individus, chacun d'entre eux ayant explicité une ou plusieurs chaînes.

Raisonnons pour simplifier sur cinq chaînes.

Individu 1 chaîne 1: Prix; Economie; Maîtrise de la chaleur; Vie confortable aisée.  
 Individu 1 chaîne 2: Plaisir; Confort; Sécurité; Sécurité familiale.  
 Individu 2 chaîne 1: Indispensable; Maîtrise de la chaleur; Qualité de la vie; Vie confortable aisée.  
 Individu 2 chaîne 2: Température constante; Tranquillité; Indépendance; Confort; Sécurité familiale.  
 Individu 2 chaîne 3: Puissant; Maîtrise de la chaleur; plaisir.

Tableau 1: Exemples de chaînes.

Table 1: Examples of chains.

Cette information peut être retranscrite sous forme de tableau binaire comme l'illustre la matrice 1. Cette matrice établit la présence ou l'absence de chaque variable pour chacune des chaînes exprimées. La matrice 1 préserve toute la richesse de l'information contenue dans les chaînes exprimées.

	Prix	Economie	Maîtrise chaleur	Vie confortabl	Plaisir	Confort	Sécurité	Sécu. familial	Indispensable	Qualité de vie	Temp. cste	Tranquillité	Indépendant	Puissant
Ind1 ch.1	1	1	1	1	0	0	0	0	0	0	0	0	0	0

Ind1 ch.2	0	0	0	0	1	1	1	1	0	0	0	0	0	0
Ind2 ch.1	0	0	1	1	0	0	0	0	1	1	0	0	0	0
Ind2 ch.2	0	0	0	0	0	1	0	1	0	0	1	1	1	0
Ind2 ch.3	0	0	1	0	1	0	0	0	0	0	0	0	0	1

*Matrice 1: Matrice visualisant les chaînes exprimées*  
*Matrix 1: Matrix visualizing expressed chains*

Dans les techniques classiques d'analyse de données, l'information contenue dans la matrice 1 est manipulée mathématiquement pour donner lieu à une matrice qui exprime pour chaque couple de variables (ou chaque couple de chaînes) la distance existant entre ces deux variables ou ces deux chaînes. Si on retient par exemple l'indice de Jaccard pour apprécier l'intensité de la liaison entre les variables prises deux à deux, on obtient, en application de la formule ci dessous un indice qui mesure l'intensité de la relation entre les variables « prix » et « économie ».

		Variable Prix	
		Présence	Absence
Variable Economie	Présence	N <sub>A</sub>	N <sub>B</sub>
	Absence	N <sub>C</sub>	N <sub>D</sub>

N<sub>A</sub> désigne le nombre de chaînes dans lesquelles « Prix » et « Economie » sont présents. Ici, N<sub>A</sub> =1

N<sub>B</sub> désigne le nombre de chaînes pour lesquelles « Prix » est un descripteur de la chaîne mais pas « Economie ».

En suivant le même principe on trouve N<sub>C</sub>=2 et N<sub>D</sub> =2

l'indice de Jaccard apprécie l'association entre les deux variables par l'expression:

$\frac{N_A}{N_A + N_B + N_C}$  ce qui donne un indice de 0.33 dans l'exemple considéré entre les descripteurs

« Prix » et « Economie ». De proche en proche, il est possible de construire la matrice de Jaccard associé au problème.

Ces matrices de distance garantissent la projection des variables ou des chaînes dans un espace réduit. Toutefois, le recours à de telles matrices de distance introduit un biais en ce sens que ce n'est plus l'information brute de départ qui est manipulée mais un construit mathématique.

Dans l'approche réseau, on privilégie au contraire une démarche plus pragmatique en raisonnant tout le long de la chaîne de traitement sur l'information brute initiale sans la transformer. L'information brute est alors traduite sous forme matricielle de deux façons différentes mais complémentaires.

- Si nous centrons l'analyse sur les chaînes exprimées par les consommateurs, la matrice résultante va exprimer le nombre de variables que partagent ces chaînes prises deux à deux. En reprenant l'exemple précédent, nous obtenons la matrice présentée dans le tableau 2 et connue sous le nom de «matrice de Condorcet ».

	Individu 1 chaîne 1	Individu 1 chaîne 2	Individu 2 chaîne 1	Individu 2 chaîne 2	Individu 2 chaîne 3
Individu 1 chaîne 1	4	0	2	0	1
Individu 1 chaîne 2	0	4	0	2	0
Individu 2	2	0	4	0	1

chaîne 1					
Individu 2 chaîne 2	0	2	0	5	1
Individu 2 chaîne 3	1	0	1	1	3

Tableau 2: Matrice de Condorcet réalisée à partir des cinq chaînes.

Table 2: Condorcet matrix realized up five chains.

Ce tableau nous fournit donc une information relative à la proximité des chaînes prises deux à deux en fonction du nombre d'éléments qu'elles partagent. Par exemple, la chaîne 1 de l'individu 1 et la chaîne 1 de l'individu 2 ont deux éléments en commun.

L'information présentée ici est plus pauvre que celle proposée dans le tableau 1, parce que nous ne pouvons plus préciser quels sont les éléments partagés par les chaînes prises deux à deux.

- Si, en revanche, nous centrons l'analyse sur les variables, nous exprimons le nombre de chaînes dans lesquelles deux variables sont présentes. En reprenant l'exemple précédent, nous obtenons la matrice présentée dans le tableau 3 et connue sous le nom de « matrice de Burt ».

	Prix	Economie	Maîtrise chaleur	Vie confortabl	Plaisir	Confort	Sécurité	Sécu. familial	Indispensable	Qualité de vie	Temp. cste	Tranquillité	Indépendant	Puissant
Prix	1	1	1	1	0	0	0	0	0	0	0	0	0	0
Economie	1	1	1	0	0	0	0	0	0	0	0	0	0	0
Maîtrise de la chaleur	1	1	1	2	0	0	0	0	1	1	0	0	0	1
Vie confortable	1	1	②	1	0	0	0	0	1	1	0	0	0	0
plaisir	0	0	0	0	1	1	1	1	0	0	0	0	0	1
Confort	0	0	0	0	1	1	1	2	0	0	1	1	1	0
Sécurité	0	0	0	0	1	1	1	1	0	0	1	1	1	0
Sécurité familiale	0	0	0	0	1	1	2	1	0	0	0	0	0	0
Indispensable	0	0	1	1	0	0	0	0	1	1	0	0	0	0
Qualité de la vie	0	0	1	1	0	0	0	0	1	1	0	0	0	0
Température constante	0	0	0	0	0	1	0	1	0	0	1	1	1	0
Tranquillité	0	0	0	0	0	1	0	1	0	0	1	1	1	0
Indépendance	0	0	0	0	0	1	0	1	0	0	1	1	1	0
Puissant	0	0	1	0	1	0	0	0	0	0	0	0	0	1

Tableau 3: Tableau de Burt réalisé à partir des cinq chaînes.

Table 3: Burt matrix realized up five chains.

Les variables « Vie confortable » et « Maîtrise de la chaleur », par exemple, sont co-présentes dans deux chaînes.

Cette matrice nous fournit une information sur l'intensité de la relation qui existe entre les variables sur la base du nombre de chaînes dans lesquelles elles sont présentes ensemble. L'information obtenue dans ce tableau est plus pauvre que celle du tableau 1, parce que nous ne pouvons pas préciser quelles sont les deux chaînes dans lesquelles les variables « Vie confortable » et « Maîtrise de la chaleur » sont co-présentes.

Les matrices de Condorcet et de Burt retransmettent en fait l'intégralité de l'information contenue dans les données brutes et doivent être comprises comme les deux facettes d'une même réalité. En termes mathématiques, la seconde analyse correspond au dual de la première. L'analyse que nous allons développer va consister à considérer successivement chacune de ces matrices. Ces « aller-retour » entre matrice de Condorcet et matrice de Burt garantiront la restitution de l'intégralité de l'information contenue dans les chaînes exprimées. Il n'y a pas de perte d'information entre l'information brute et l'information matricielle qui est manipulée mathématiquement.

## 2- Le point d'arrivée de la méthode.

L'objectif de la méthode consiste à construire des réseaux à partir des matrices élaborées lors de l'étape précédente. Ceux-ci sont susceptibles d'être interprétés très simplement [BOUT et alii 96 2], comme le montrent les réseaux ci-dessous élaborés à partir des deux matrices de l'exemple précédent.

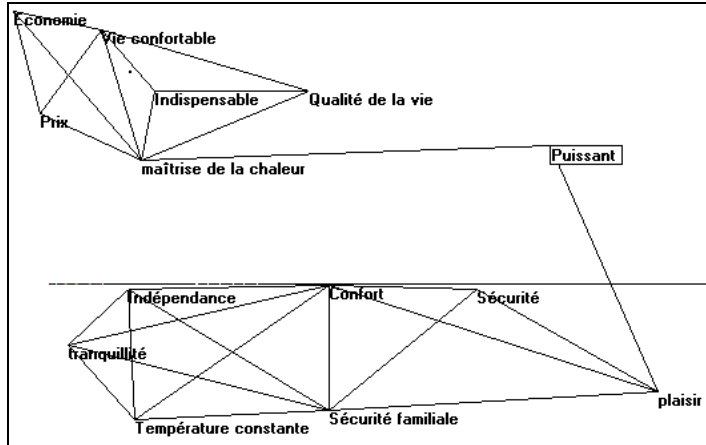


Figure 1: Réseau des relations entre variables  
Figure 1: Relationships network between variables

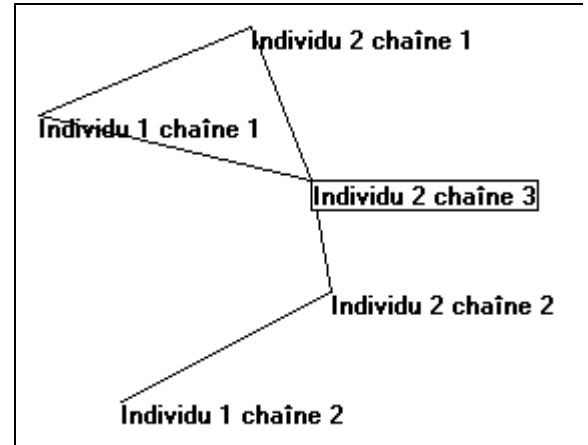


Figure 2: Réseau des relations entre individus  
Figure 2: Relationships network between people

Considérons le réseau présenté dans la figure 1. En son sein, chaque sommet correspond à une variable et chaque lien entre deux sommets indique l'existence d'une co-présence entre les deux variables dans au moins une des chaînes. Cette figure est la transcription visuelle rigoureuse de la matrice de Burt. Chaque élément présent dans la première colonne de la matrice est représenté en tant que sommet et chaque cellule différente de zéro à l'intersection de deux variables se traduit par un lien entre ces deux variables. Le raisonnement est identique pour la construction du réseau figure 2, à partir de la matrice de Condorcet.

La figure 1 fait apparaître un certain nombre d'éléments caractéristiques. Les variables sont regroupées au sein de sous-groupes. On remarque que l'attribut « puissant » joue un rôle particulier en tant qu'interface ou pivot entre deux d'entre eux. Une telle variable est appelée isthme [AHO et alii 87].. Si nous retirions cette variable de la matrice initiale, les deux sous-groupes seraient déjoints.

Le processus par lequel est généré automatiquement ce genre de réseau dépasse l'objet de cette communication. Simplement, le passage d'une matrice à une représentation cartographique utilise certains algorithmes tels que les algorithmes génétiques [BICK et alii 94] ou le recuit simulé [BOUT et alii 96 3]. Le but est d'optimiser le positionnement des sommets sur un graphe compte tenu de certaines contraintes d'esthétisme. Ce type de recherches fait l'objet de développements au sein du laboratoire « Le Pont » de l'Université de Toulon et du Var.

Nous pouvons mettre en avant, à ce stade, trois qualités de cette méthode d'analyse de données.

- Tout d'abord, cette approche donne lieu à une représentation graphique visuelle qui peut être interprétée sans formation préalable aux techniques d'analyse de données. Cette qualité est fondamentale et constitue un point fort de la méthode en comparaison des analyses de données traditionnelles. Ces dernières tirent souvent leur complexité du fait qu'elles transforment les matrices de départ en matrices de distance entre les variables pour pouvoir



effectuer des projections en espaces réduits. Or, le choix de la métrique conditionne le résultat final. Ceci fait des techniques d'analyse de données classiques des analyses entre les mains d'experts souvent mal comprises de l'utilisateur néophyte.

- Cette technique présente également l'avantage d'une grande souplesse. Elle peut être conduite sur toute matrice, pourvu que celle-ci soit carrée et symétrique.

- Enfin, cette méthode présente l'avantage non négligeable de l'autodétermination des sous-groupes sans intervention préalable humaine.

### **3- Les hypothèses de la méthode.**

L'exemple précédent est séduisant mais trop limité pour être suffisamment convaincant quant à l'utilité de la méthode. Il semble en effet difficile de comparer les techniques d'analyses de données classiques qui synthétisent une information massive pouvant comporter des milliers d'enregistrements avec une analyse illustrée seulement à partir de cinq chaînes.

Si le réseau présenté figure 1 est si simple, c'est parce que les associations entre les variables sont suffisamment peu nombreuses pour garantir sa lisibilité. Toutefois, si d'autres individus avaient émis des chaînages supplémentaires, de nouvelles associations entre les chaînes auraient été exprimées et la représentation simultanée de ce genre de relations sur un réseau aurait donné lieu à un graphe surchargé voire inextricable. La réalité d'un phénomène complexe et nuancé ne peut pas être a priori représentée par un réseau clair et limpide. Partant de ce constat, le réseau doit être simplifié pour gagner en lisibilité. Cette simplification va consister à supprimer du réseau certains liens entre variables. Le but de ces suppressions est d'accepter de perdre de l'information pour gagner en visualisation et en significativité. Toutefois, une telle simplification du réseau obéit à une logique qui peut être clairement énoncée et échappe ainsi à toute forme d'arbitraire.

Lorsqu'on analyse une information massive, quelle qu'en soit la forme, il est en effet possible de la décomposer en trois composantes qualifiées respectivement d'information triviale, d'information intéressante et de bruit au sens statistique du terme.

- L'information triviale se compose des associations de variables les plus fréquentes. Celles-ci constituent le dénominateur commun de l'ensemble considéré. En tant que tel, ce genre d'association apporte une information très générale. Si nous enlevons du réseau de tels liens, non seulement nous éclaircissons le réseau, mais nous lui donnons aussi une sensibilité dominante correspondant au sens des sommets ôtés.

- A l'opposé, l'information qualifiée de bruit est constituée de l'ensemble des associations entre variables qui sont les moins fréquemment obtenues. Ces associations correspondent soit à une information émergente, soit à une information relative ici à une chaîne atypique. Le fait de supprimer de telles relations du réseau va accroître sa lisibilité.

- Le reste des associations entre les variables est qualifié d'information utile.

Des analyses entreprises à partir du concept d'entropie [LHE et alii 95] ont permis de dégager des seuils permettant de séparer automatiquement ces trois types d'informations. Ainsi dispose-t-on, lorsque le réseau est trop fourni, de critères clairs permettant de le simplifier.

## **II. ILLUSTRATION DE LA DEMARCHE:**

Nous proposons d'illustrer la technique à partir d'un jeu de 184 chaînes exprimant les réponses énoncées par 79 individus. Chaque chaîne associe au maximum cinq variables correspondant à des associations attributs - conséquences - valeurs, faites par les personnes interrogées. Ces chaînes se rapportent à une étude relative au comportement de la population grenobloise en matière de consommation d'énergie.

### 1- Le point de départ de la méthode.

Considérons dans un premier temps la matrice de Condorcet associée à l'application. Cette matrice a été obtenue automatiquement à l'aide du logiciel Dataview développé au Centre de Recherche et de Rétrospective de Marseille [ROS 93]. Cette matrice 184-184 exprime le nombre de variables que partagent deux chaînes. Une extraction des neuf premières chaînes de cette matrice est présentée figure 3 pour illustrer la démarche.

Les éléments qui figurent en première ligne et première colonne correspondent pour le chiffre des unités au numéro de la chaîne exprimée et pour le reste au numéro de l'individu interrogé.

	11	12	21	22	23	24	31	32	33
11	4	0	2	0	1	1	1	1	3
12	0	4	0	2	0	0	1	1	0
21	2	0	4	0	1	1	3	1	1
22	0	2	0	5	1	0	1	4	0
23	1	0	1	1	3	0	0	1	0
24	1	0	1	0	0	4	1	1	1
31	1	1	3	1	0	1	4	2	1
32	1	1	1	4	1	1	2	5	1
33	3	0	1	0	0	1	1	1	3

Figure 3: Extrait de la matrice de Condorcet associée au problème.

Figure 3: Extract from Condorcet matrix linked to the problem.

On voit ainsi par exemple que la chaîne 3 de l'individu 3 et la chaîne 1 de l'individu 1 ont trois éléments en commun.

### 2- Le réseau obtenu à partir de la matrice de Condorcet.

Comme la représentation de l'ensemble des relations entre les chaînes est confuse, le choix a été fait de ne visualiser que les liens associés aux chaînes qui partagent au moins quatre variables en commun. Le chiffre 4 correspond au nombre de degré de liberté du système, chaque chaîne n'étant associée qu'à cinq variables maximum. Ce seuil relativement élevé permet de limiter la variété du système. Le réseau obtenu est présenté figure 4.

Ce réseau, bien que complexe et embrouillé permet de faire ressortir clairement plusieurs groupes de chaînes indépendants. Nous pouvons distinguer les groupes à structure forte: A, B et C, celui à petites structures situé en bas du réseau et les structures isolées amassées au coin sud-est. Chaque groupe identifié lors de cette phase correspond à des chaînes voisines, qui partagent au moins quatre des cinq variables les définissant pour les quatre premiers, et moins de quatre pour la structure isolée.

Chaque groupe ainsi identifié a un poids proportionnel au nombre de chaînes qui le caractérisent. Par conséquent, le groupe B défini par cinquante chaînes aura plus de poids que

le groupe C défini par trente chaînes. En exprimant ces poids en pourcentages, nous pouvons répartir la quantité d'information du réseau entre les diverses composantes de la façon suivante:

- Le groupe A: 19.03 %
- Le groupe B: 27.17 %
- Le groupe C: 16.3 %
- Le groupe des isolés: 22.28 %
- Le groupe des petits: 15.22 %.

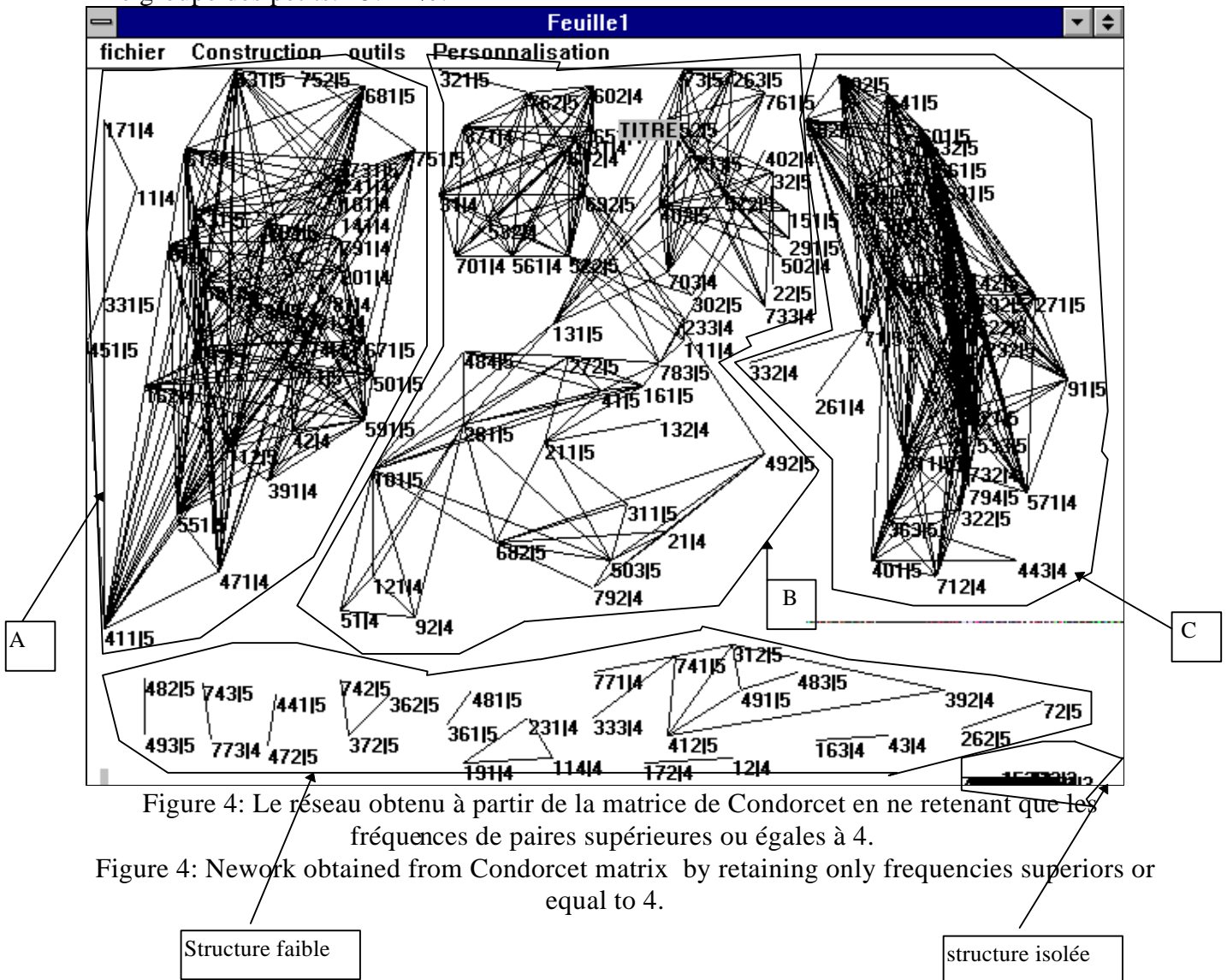


Figure 4: Le réseau obtenu à partir de la matrice de Condorcet en ne retenant que les fréquences de paires supérieures ou égales à 4.

Figure 4: Network obtained from Condorcet matrix by retaining only frequencies superior or equal to 4.

L'objectif est maintenant de caractériser chacun de ces groupes.

Pour chacun d'entre eux, il est possible de construire un réseau des variables en reconstituant la matrice de Burt associée. Le réseau résultant met en évidence les relations entre les variables partagées par les individus du groupe. Nous allons illustrer cette démarche en analysant le groupe C.

### 3- Du réseau des chaînes au réseau des variables.

Le groupe C est composé de trente chaînes présentées dans l'annexe 2.

Par définition, chacune de ces chaînes comporte au minimum quatre variables en commun avec toutes les autres. Nous pouvons construire à partir de cette liste une matrice variable - variable

qui associe à chaque couple de variables le nombre de références dans lesquelles elles sont co-présentes. Cette matrice de Burt est présentée dans la figure 5.

	3	20	38	25	39	35	28	21	34	36	33	22	17
Individuel (3)	30	29	27	21	13	11	4	3	2	1	1	1	1
Indépendance / immeuble (20)	29	29	26	21	12	11	3	3	2	1	1	0	1
Indépendance (38)	27	26	27	19	10	11	3	3	2	1	0	1	1
Tranquillité (25)	21	21	19	21	10	7	0	0	2	1	1	0	0
Vie aisée, confortable (39)	13	12	10	10	13	0	3	0	0	0	1	1	0
Liberté (35)	11	11	11	7	0	11	0	2	0	0	0	0	1
Maîtrise de la chaleur (28)	4	3	3	0	3	0	4	0	0	0	0	1	0
Economie (21)	3	3	3	0	0	2	0	3	0	0	0	0	0
Sécurité familiale (34)	2	2	2	2	0	0	0	0	2	0	0	0	0
Qualité de l'environnement (36)	1	1	1	1	0	0	0	0	0	1	0	0	0
Plaisir (33)	1	1	0	1	1	0	0	0	0	0	1	0	0
Confort (22)	1	0	1	0	1	0	1	0	0	0	0	1	0
Qualité de la vie (17)	1	1	1	0	0	1	0	0	0	0	0	0	1

Figure 5: Matrice de Burt associée au groupe C.  
Figure 5: Burt matrix associated with group C

On remarque que le groupe C est parfaitement défini à partir de treize variables: un attribut (chauffage individuel), six conséquences (qualité de la vie, indépendance par rapport aux autres habitants de l'immeuble, économie, confort et maîtrise de la chaleur) et six valeurs (plaisir, sécurité familiale, tranquillité, liberté, indépendance, et vie aisée et confortable). Nous avons choisi d'éliminer un lien entre les variables lorsque celui-ci avait une fréquence inférieure à deux. Supprimer ces liens a pour but, comme nous l'avons vu précédemment, d'éliminer du réseau le bruit statistique et donc de clarifier le réseau.

Après suppression des liens de fréquence inférieure à deux, nous obtenons le réseau suivant:

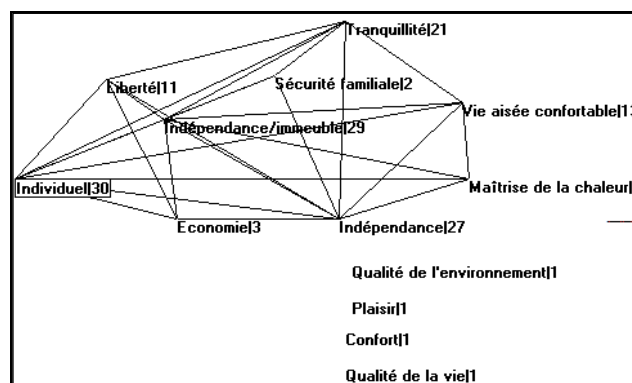


Figure 6: Réseau associé au groupe C.  
Figure 6: Network associated with group C

Dans ce réseau, il reste neuf variables en interrelations. Parmi ces dernières, trois d'entre elles sont en relation avec l'ensemble des huit autres: chauffage individuel (3), indépendance par rapport aux autres habitants de l'immeuble (20) et indépendance (38). De telles variables jouent un double rôle. D'une part, elles permettent de donner la sensibilité dominante du groupe: l'indépendance. D'autre part, elles gênent par leur présence sur le réseau la mise en évidence de phénomènes plus fins. La figure suivante correspond au réseau de base auquel ont été soustraites ces trois variables.

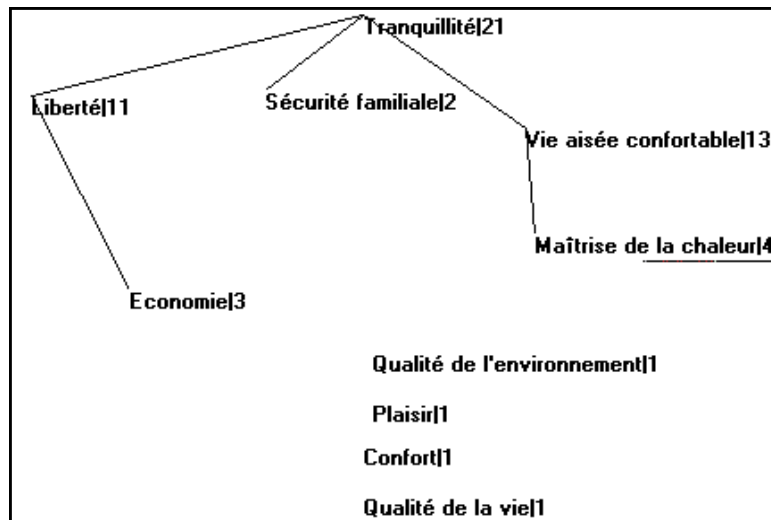


Figure 7: Réseau associé au groupe C en enlevant les isthmes.

Figure 7: Network associated with group C obtained by removing isthmuses.

La variable 25 (tranquillité) peut être considérée comme un isthme du système. Si nous l'enlevons, nous faisons apparaître deux associations indépendantes:

- 21, 35 correspondant aux variables Economie, liberté
- 28, 39 correspondant aux variables maîtrise de la chaleur, vie aisée et confortable.

Ces deux associations correspondent à deux sous-orientations du groupe C.

Une analyse similaire peut être conduite sur chacune des grandes structures identifiées à la figure 4. Ceci permet de mettre en évidence trois dominantes par rapport au chauffage :

- Groupe A: l'économie
- Groupe B: le confort utilitariste
- Groupe C: l'indépendance.

Dans le cas présent nous avons considéré les isthmes comme les éléments permettant de déterminer la sensibilité de chacun des groupes. Ceci n'est plus le cas pour les groupes des petites structures et des isolés. Ces groupes étant divisés en plusieurs segments, le retrait des isthmes va nous permettre de les segmenter. Les isthmes sont alors envisagés comme des éléments perturbants qui viennent brouiller l'information. Leur retrait permet de segmenter les chaînes émises par la population.

L'analyse des petits groupes situés en bas du réseau de la figure 4 permet de faire ressortir des comportements mineurs. Ceux ci correspondent à quatre dominantes: efficace, écologique, « pépère » et pratique.

L'analyse des isolés situés au sud-est de la figure 4 permet de retrouver l'ensemble des structures énoncées précédemment. Ces chaînes sont constituées de trois ou quatre éléments. Pour cette raison, elles n'avaient pu être intégrées dans les structures précédentes. Comme précédemment, nous avons choisi de ne retenir sur la matrice de Burt que les éléments ayant une fréquence d'occurrence supérieure à deux. Ceci nous permet d'éliminer le bruit au sens

statistique du terme. Grâce à ceci et suite au retrait des isthmes par itérations successives, nous avons pu identifier sept sous-groupes: les propres, les « température constante », les économes, les individualistes, les qualitatifs, les sécuritaires et les **pirates**.

### Conclusion.

L'analyse des réseaux fait ressortir cinq groupes de consommateurs en matière de comportement énergétique. Les trois premiers indiquent les tendances générales de la population: confort, indépendance et économie. Le groupe des petites structures a permis l'identification de comportements émergents comme l'écologie, tandis que celui des isolés permettait de retrouver de façon plus étroite les tendances dégagées préalablement dans les trois premiers groupes. La synthèse des résultats est présentée en Annexe 3.

L'étude réalisée accuse quelques limites. Tout d'abord, cette démarche est centrée tantôt sur les variables en perdant l'information relative aux chaînes, tantôt sur les chaînes en perdant l'information relative aux variables. De plus, nous n'avons pas pu déterminer pour chaque groupe une chaîne type attribut - conséquences - valeurs qui serait représentative du groupe considéré. En effet, la méthode ne distingue pas lors de la construction du réseau les attributs des conséquences et des valeurs. Rien ne garantit a priori que les variables ôtées appartiennent à ces trois catégories.

Diverses voies de recherche semblent se dégager. Il serait intéressant de comparer cette méthode avec l'approche spatiale, l'analyse canonique généralisée, développée par Valette-Florence. Ceci permettrait d'asseoir la validité conjointe des deux méthodes et de vérifier qu'elles aboutissent bien à des résultats semblables en matière de segmentation. De plus, il serait utile d'intégrer à l'étude entreprise des informations complémentaires comme les variables socio-démographiques pour avoir une vision plus complète de la réalité. Enfin, il conviendrait de passer de la notion de chaîne à la notion d'individu, un individu ayant pu citer plusieurs chaînes. On pourrait ainsi établir une typologie des individus et les distinguer selon que les chaînes qu'ils ont exprimées appartiennent aux mêmes groupes ou à des groupes distincts.

Nous aurions ainsi trois degrés de finesse dans l'analyse: le premier au niveau des variables, le deuxième au niveau des chaînes et le dernier au niveau des individus.

### Annexe 1: Liste des attributs - conséquences - valeurs et occurrence correspondante.

Attributs	Conséquences	Valeurs
- Propre	- Qualité de vie	- Plaisir
- Pas de stockage	- Gain de place	- Sécurité familiale
- Individuel	- Pas de problème	- Liberté
- Collectif	d'approvisionnement	- Qualité de
- Puissant (chauffe rapidement)	- Indépendance par rapport aux autres habitants de l'immeuble	l'environnement
- Pas de danger (fiable)	- Economie	- Propreté
- Pas de coupure (disponible)	- Confort (bien-être)	- Indépendance
- Pratique (facile à utiliser)	- Gain de temps	- Vie aisée, confortable
- Indispensable	- Sécurité	- Un monde de beauté
- Température constante	- Tranquillité	
	- Préservation de l'environnement	

- Préserve un taux idéal d'humidité de l'air - Prix	- Pas d'odeur, de fumée - Maîtrise de la chaleur	
--	---	--

### Annexe 2. Les chaînages exprimés par les membres du groupe C.

N° de chaîne	Variable 1	Variable 2	Variable 3	Variable 4	Variable 5
71	individuelle	indépendance / l'immeuble	maîtrise de la chaleur	vie aisée confortable	indépendance
91	individuelle	indépendance / l'immeuble	tranquillité	indépendance	sécurité familiale
113	individuelle	indépendance / l'immeuble	tranquillité	liberté	indépendance
142	individuelle	indépendance / l'immeuble	tranquillité	indépendance	liberté
192	individuelle	indépendance / l'immeuble	tranquillité	liberté	indépendance
222	individuelle	indépendance / l'immeuble	tranquillité	liberté	indépendance
232	individuelle	indépendance / l'immeuble	tranquillité	liberté	indépendance
261	individuelle	indépendance / l'immeuble	maîtrise de la chaleur	vie aisée confortable	
271	individuelle	indépendance / l'immeuble	tranquillité	indépendance	sécurité familiale
301	individuelle	confort	maîtrise de la chaleur	indépendance	vie aisée confortable
322	individuelle	qualité de vie	indépendance	liberté	indépendance
332	individuelle	indépendance / l'immeuble	maîtrise de la chaleur	indépendance	
363	individuelle	indépendance / l'immeuble	économie	liberté	indépendance
401	individuelle	indépendance / l'immeuble	économie	liberté	indépendance
421	individuelle	indépendance / l'immeuble	tranquillité	qualité de l'environnement	indépendance
443	individuelle	indépendance / l'immeuble	économie	indépendance	
521	individuelle	tranquillité	indépendance	indépendance	vie aisée confortable
533	individuelle	indépendance / l'immeuble	tranquillité	liberté	indépendance
541	individuelle	indépendance / l'immeuble	tranquillité	indépendance	vie aisée confortable
571	individuelle	indépendance / l'immeuble	tranquillité	indépendance	

592	individue 1	tranquillité	indépendance	indépendance	vie aisée confortable
601	individue 1	indépendance / l'immeuble	tranquillité	indépendance	vie aisée confortable
611	individue 1	indépendance / l'immeuble	tranquillité	liberté	indépendance
632	individue 1	tranquillité	indépendance	indépendance	vie aisée confortable
661	individue 1	indépendance / l'immeuble	tranquillité	indépendance	vie aisée confortable
691	individue 1	tranquillité	indépendance	indépendance	vie aisée confortable
702	individue 1	indépendance / l'immeuble	tranquillité	indépendance	vie aisée confortable
712	individue 1	indépendance / l'immeuble		liberté	indépendance
732	individue 1	indépendance / l'immeuble	tranquillité	vie aisée confortable	
794	individue 1	indépendance / l'immeuble	tranquillité	plaisir	vie aisée confortable



## **Bibliographie:**

- AHO A., HOPCRAFT J. et ULLMAN J.(1987), *Structure des données et algorithmes*, Interéditions, France.
- BICKING F., FONTEIX C.,CORRIOU J.P. et MARC I. (1994), « Global Optimization by Artificial Life: A New Technique Using Genetic Population Evolution », *R.A.I.R.O Recherche Opérationnelle*, 28, 23-36.
- BOUTIN E., FERRANDI J.M. et VALETTE-FLORENCE P. (1996-1), Les réseaux comme outil d'analyse des chaînages cognitifs: une illustration expérimentale, *Papier de Recherche 96-06*, CERAG Ecole Supérieure des Affaires, Grenoble.
- BOUTIN E, DUMAS P, ROSTAING H. et QUONIAM L. (1996-2), Les réseaux comme outil d'analyse en bibliométrie. Un cas d'application: les réseaux d'auteurs, *Les cahiers de la documentation Belge*, Vol 50, N°1, 1996, p 3-13
- BOUTIN E. et FERRANDI J.M. (1996-3),« *La construction automatique de réseaux sociaux: étude exploratoire*», *Actes du Colloque National de Recherches en IUT en Mathématiques, Statistiques, Informatique*», Clermont-Ferrand.
- GUTMAN J. (1982), «A Means-End Chain Model Based on Categorization Processes», *Journal of Marketing Research* 46, 1, 60-72.
- LHEN J., LAFOUGE T.,ELSKENS Y., QUONIAM L. et DOU H. (1995), La statistique des lois de Zipf», *Acte du colloque: «Les systèmes d'information élaborés*», Ile Rousse.
- REYNOLDS T.J. et GUTMAN J. (1988), «Laddering Theory, Method, Analysis and Interpretation », *Journal of Advertising Research* 28, 1, 11-31.
- ROSTAING H.,(1993),*Veille technologique et bibliométrie: concepts, outils et applications* Thèse: Aix Marseille III, 353p.
- VALETTE-FLORENCE P. (1994), «Introduction à l'analyse des chaînages cognitifs», *Recherche et Applications en Marketing*9, 1, 93-117.