

Auteur : Jérôme Baray

Les limitations du modèle P-médian

Le modèle p-médian est uniquement axé, comme nous l'avons précisé, sur les notions de distance et de demande dans le cas du modèle p-médian pondéré. Mais, même si le propos n'est pas fondamentalement de critiquer le modèle lui-même, il est évident que le consommateur ne se fondera pas seulement sur le seul critère de distance pour choisir son point de vente parmi un éventail de concurrents. Les autres paramètres jouant sur la décision des clients peuvent être formulés sous forme de cinq principes qui guideront la sélection des emplacements commerciaux^{1 2}:

- *le principe d'interception*: le point de vente a d'autant plus de chance de capter la clientèle que le passage de consommateurs à proximité, mesuré le plus souvent en terme de trafic piétonnier ou routier, est important.

- *le principe d'attraction cumulative*: le regroupement de commerces dans un environnement géographique proche d'activités similaires crée souvent une synergie d'attractivité alors supérieure à la somme des attractivités individuelles des commerces.

- *le principe de compatibilité*: de même, le regroupement d'activités complémentaires permet assez souvent de parvenir à ce même effet synergique des attractivités.

- *le principe de suréquipement*: une trop grande concentration de points de vente a tendance à avoir un effet répulsif sur la clientèle alors soumis à l'inconfort de la congestion du trafic.

- *le principe d'accessibilité*: l'accessibilité au point de vente, c'est-à-dire la facilité d'entrer, de pénétrer, de traverser et de sortir du site commercial doit être le plus facile possible. Une

¹ DELOZIER M.W. et LEWISON D.M. (1986) *Retailing*, 2nd ed., Merill.

² CLIQUET G. (1992) *Le Management Stratégique des Points de Vente*, p.187-191, Ed. Sirey.

bonne signalisation, la qualité et le nombre de voies d'accès, les facilités de parking sont autant de paramètres à privilégier dans le choix d'un emplacement convenable.

Une solution que nous proposons pour intégrer tous ces critères seraient d'introduire dans le problème p-médian un poids évaluant chaque site potentiel. Rien n'empêche en effet de mesurer et de quantifier pour chaque emplacement potentiel des points de vente les différents critères mentionnés tels que l'ont fait Lewison et Delozier³ et Jallais, Orsony et Fady⁴. Puis il convient d'attribuer une note globale à chaque site potentiel, note résultant d'une pondération de ces critères. A titre d'exemple, il suffirait d'attribuer des notes par exemple de 0 (emplacement vraiment médiocre) à 100 (emplacement excellent) pour chaque critère, de les pondérer avec des coefficients positifs de manière à quantifier leur importance et d'obtenir une note globale n_j . Cette note introduite dans la fonction objectif avec un signe négatif (ayant tendance à minimiser avantageusement la fonction) au même niveau que les distances d_{ij} servirait de pondération à chaque site potentiel d'activité. La distance d_{ij} en tant que critère impliquant à la fois le commerce et le client considéré devrait alors être normalisée (sur la même échelle de notes) de manière à obtenir une homogénéité au sein de la fonction objectif. La formulation mathématique du p-médian deviendrait alors:

$$\text{Minimiser } \sum_i \sum_j a_i (d_{ij} - n_j) x_{ij} \quad \text{représente la nouvelle fonction objectif,} \quad (1)$$

$$\text{avec } \sum_i x_{ij} = 1, \quad \forall i, \quad \text{assure que tous les clients sont assignés à une activité et une seule,} \quad (2)$$

$$\text{et } d_{ij} - n_j < 0, \quad \forall j \quad \text{assure que toute activité possède un minimum d'attractivité,} \quad (2)'$$

$$x_{ij} \leq y_j, \quad \forall i, j \quad \text{empêche d'assigner un client à une activité si elle n'est pas ouverte,} \quad (3)$$

$$\sum_j y_j = p, \quad \text{le nombre total d'activités est } p, \quad (4)$$

$$x_{ij}, y_j \in \{0,1\}, \quad \forall i, j \quad \text{nature binaire des variables } x_{ij}, y_j \quad (5)$$

³ DELOZIER M.W. et LEWISON D.M. (1986) *Retailing*, 2nd ed., Merill.

⁴ JALLAIS J., ORSONY J. et FADY A. (1994) *Marketing du Commerce de Détail*, Vuibert, Paris.

où

a_i : est la demande au nœud i ,

n_j : la note positive attribuée au site potentiel du nœud j ,

d_{ij} : la distance du nœud i au nœud j ,

p : le nombre d'activités à localiser,

$x_{i,j} = 1$, si le nœud i est assigné à l'activité j et 0 autrement,

$y_j = 1$, si l'activité j est ouverte et 0 autrement.

On observe que logiquement dans la fonction objectif, plus l'activité au nœud j est attractive (note n_j importante), meilleure est la fonction objectif (valeur plus faible) et que l'attractivité est un réducteur de distance, la valeur $(d_{ij} - n_j)$ pouvant être qualifiée de "distance psychologique" qui est la distance ressentie effectivement par les consommateurs et non plus seulement la distance routière ou temporelle^{5 6}. Un autre avantage de cette nouvelle formulation que l'on pourrait qualifier de "modèle p -médian généralisé" est qu'elle est soluble par les méthodes traditionnelles de résolution du p -médian puisque la forme de la fonction objectif est similaire (forme linéaire).

2.3.2 La complexité des méthodes de résolution du p -médian

L'autre reproche déjà évoqué au niveau du modèle p -MP concerne cette fois ses méthodes existantes de résolution. Celles-ci sont très lourdes et même souvent impossibles à mettre en œuvre si l'on considère un réseau de plusieurs centaines de milliers de nœuds. Quelle tâche herculéenne que d'imaginer vouloir trouver une solution plus ou moins optimale à un réseau comprenant l'ensemble des clients potentiels d'une future grande surface ! Un autre problème

⁵ HUBBARD R. (1978) A Review of Selected Factors Conditioning Consumer Travel Behavior, *Journal of Consumer Research*, vol. 5, june, p.1-21.

⁶ VOLLE P. (1999) Du Marketing des Points de Vente à Celui des Sites Marchands : Spécificités, Opportunités et Questions de Recherche, *Cahier n°276*, Centre de Recherche DMSP, juin 1999.

du même ordre, mais cette fois dans le domaine de la gestion informatique serait de considérer le réseau de communication Internet et les différents micro-ordinateurs, serveurs et réseaux Intranet le composant (soit des millions d'ordinateurs à travers le monde): sur quels serveurs (les nœuds d'activités) serait-il préférable de localiser les données pour que celles-ci soient acheminées le plus rapidement possible à destination des autres ordinateurs (les clients) connaissant les probabilités de requête (les demandes) ?

Le problème classique du p-médian a pourtant fait couler beaucoup d'encre, en témoigne le nombre d'articles où ce modèle est cité (en excluant les auto-citations):

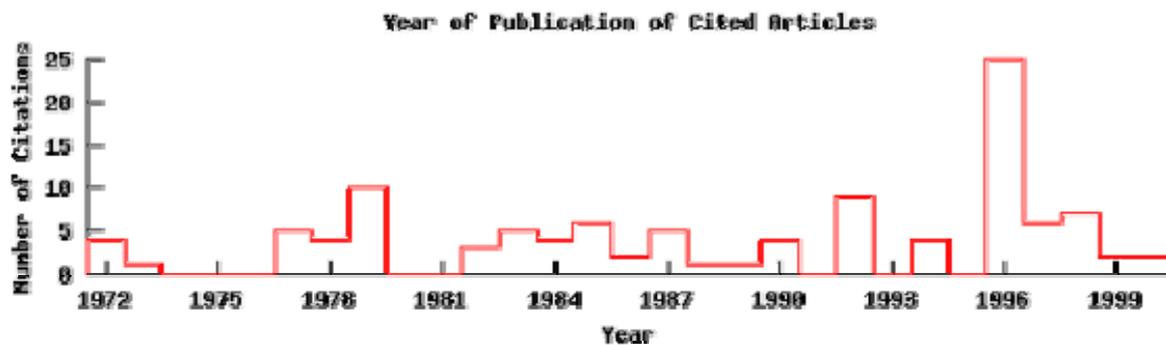


Fig. 2.13 - Nombres d'articles traitant du p-médian par année - source: *The Scientific Literature Digital Library*

Les 225 articles cités sur 30 ans ne s'attaquent cependant qu'à l'affinage de la résolution du p-médian à savoir essentiellement améliorer l'optimalité des solutions à travers de nouvelles heuristiques d'amélioration (et rarement des algorithmes de résolution). Les autres publications de recherche traitent principalement des applications du p-MP en adaptant le modèle à certains cas particuliers comme l'efficacité des moteurs de recherche⁷, le coloriage

⁷ CRASWELL N. et BAILEY P. (1999) *Is it fair to evaluate Web systems using TREC ad hoc*, Department of Computer Science - CSIRO, The Australian National University Pub.

de graphes à l'aide du p-MP et d'un algorithme génétique⁸ ou encore la localisation de simulateurs de conduite de char de l'armée américaine au plus près des casernes⁹.

2.3.3 L'analyse typologique : une approche pour simplifier la problématique du p-MP

Peu d'articles abordent vraiment la partie fondamentale, mais ô combien stratégique de la formulation ou de la modélisation même du problème p-MP. Si elle est mieux établie, elle permettra de simplifier le problème et de parvenir à une solution plus rapidement ou à une meilleure solution en utilisant un algorithme plus lent mais plus performant capable de traiter les modèles intégrant un très grand nombre de nœuds. Cependant, il est à noter qu'un problème assez étudié et à la mode est celui de l'analyse typologique (cluster analysis en anglais). Déjà évoqué par Aristote, cette dernière consiste à partitionner un ensemble d'entités en sous-ensembles bien séparés et homogènes selon un certain nombre de paramètres. Dans le domaine de l'identification de classes, on peut recenser dans la littérature deux approches qui se sont développées, la première étant celle de la reconnaissance déterministe et de décision par sélection de distance; la seconde est statistique car elle a pour base des probabilités acquises lors d'une phase d'apprentissage et une décision par sélection de pénalisations encourues. En particulier, la classification séquentielle utilisée par de nombreux chercheurs dans les procédures de reconnaissance de forme se rattache à cette dernière approche^{10 11 12 13}

⁸ RIBEIRO FILHO G. et NOGUEIRA LORENA L.A. (2000) *Improvements On Constructive Genetic Approaches To Graph Coloring*, Sao José Dos Campos, Spain.

⁹ MURTY K.G. et DJANG P.A. (2000) The US Army National Guard's Mobile Training Simulator Location and Routing Problem, *US Army Training and Doctrine Analysis Command*, Dept. of IOE, University of Michigan.

¹⁰ DUDA R.O. et HART P.E. (1973) *Pattern Classification and Scene Analysis* Editions J. Wiley, New York, N.Y.

¹¹ TOU J.T. et GONZALEZ R.C. (1974) *Pattern Recognition Principles*, Edition Addison-Wesley, Reading, Mass.

¹² FU K.S. (1968) *Sequential Methods in Pattern Recognition and Machine Learning*, Editions Academic Press, New York, N.Y.

¹³ PAVLIDIS T. (1977) *Structural Pattern Recognition*, Editions Springer-Verlag, New York, N.Y.

¹⁴ ¹⁵ ¹⁶. Il existe également des méthodes de segmentation fondées sur un algorithme génétique associé à un modèle de régression¹⁷.

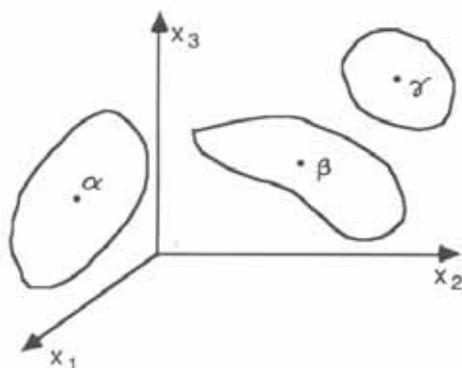


Fig. 2.14 - Un exemple de 3 classes α , β et γ identifiées par les variables x_1 , x_2 , x_3

Si l'analyse typologique intéresse principalement le domaine des mathématiques statistiques, elle recense également des applications dans le domaine de la localisation et du problème p-médian¹⁸ (voir son utilisation au § 2.4.2.4. dans la délimitation des zones de chalandise). Un exemple récent est représenté par la décomposition de la Suisse ou plutôt des 2863 principales communes suisses en 23 cellules associées chacune à un centre urbain¹⁹ selon diverses méthodes comparées: p-médian, méthode Alt²⁰, algorithme de décomposition²¹. Le problème est en premier lieu d'identifier les 23 centres urbains et d'y affecter ensuite des communes selon le critère du plus proche voisin en terme de distance à vol d'oiseau. Selon notre point de vue, cette analyse typologique géographique peut conduire en quelque sorte à une simplification du problème p-médian, dans le sens où, plutôt que prendre comme nœuds du réseau l'ensemble des 2863 communes au sein desquelles on rechercherait une ou plusieurs

¹⁴ PATRICK E.A. (1972) *Fundamentals of Pattern Recognition*, Edition Prentice-Hall, Englewood Cliffs, N.J.

¹⁵ POSTAIRE J.G. (1987) *De l'Image à la Décision*, Editions Dunod-Informatique Bordas, Paris.

¹⁶ CHEN C.H. (1973) *Statistical Pattern Recognition*, Editions Spartan Books, Rochelle Park, N.J.

¹⁷ AURIFEILLE J.M. (2000) A Bio-mimetic Approach to Marketing Segmentation : Principles and Comparative Analysis, *European Journal of Economic and Social Systems* 14 N°1, p. 93-108.

¹⁸ ROGER P. (1983) *Description du Comportement Spatial du Consommateur*, Thèse de Doctorat, Lille.

¹⁹ TAILLARD E.D. (1996) Heuristic Methods for Large Centroid Clustering Problems, *Technical Report Idsia 96-96*, Idsia.

²⁰ COOPER L.G. (1963) Location-Allocation Problems, *Operations Research* 11, 331-341.

²¹ TAILLARD E.D. (1996) Heuristic Methods for Large Centroid Clustering Problems, *Technical Report Idsia 96-96*.

localisations optimales, on pourrait plus aisément considérer les 23 centres urbains comme étant les nœuds d'un réseau p-médian simplifié: l'étape d'analyse typologique constituerait une étape amont destinée à constituer ce réseau p-médian simplifié (avec comme nœuds les 23 centres urbains) préalablement à sa résolution.

Mais, cette manière de simplifier le problème p-médian comporterait de nombreuses limitations dont celle en premier lieu de devoir spécifier à l'avance en combien de classes (et donc de nœuds du futur réseau p-médian simplifié) il conviendrait de partitionner l'ensemble des entités de départ: ainsi, dans le problème précédent, il a été fixé d'entrée de jeu de partitionner le territoire suisse en un nombre de 23 cellules associées à 23 centres urbains sans que ce nombre ait été ultérieurement justifié.



Fig. 2.15- Partition des communes suisses en 23 cellules représentées chacune par un centre urbain ²²

²² TAILLARD E.D. (1996) Heuristic Methods for Large Centroid Clustering Problems, *Technical Report Idsia 96-96*, Idsia.

Les exemples personnels suivants illustrent le principe de l'analyse typologique utilisée dans le cadre du p-médian et l'erreur dans la recherche d'une localisation optimale qui peut résulter d'un nombre prédéfini de partitions. D'abord, un cas où l'approche de l'analyse typologique fonctionne bien: considérons 76 villes de même taille réparties sur un territoire, chaque ville étant représentée par un cercle :

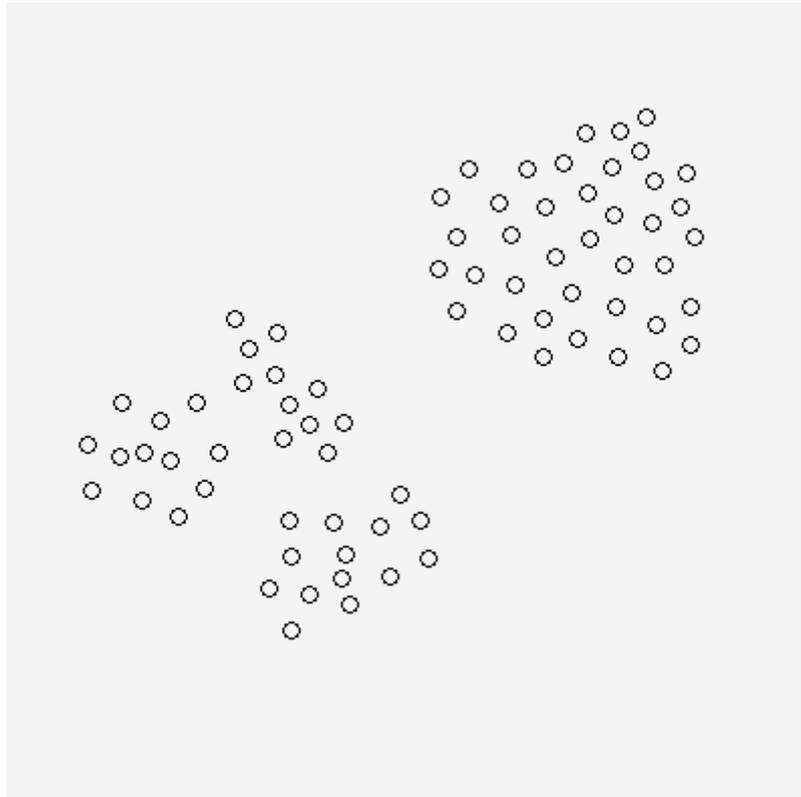


Fig. 2.16 - Les 76 villes représentées chacune par un cercle

Si l'on décide de partitionner l'espace géographique en deux par la méthode des k-plus proches voisins ($k=2$) par exemple, on aura deux régions A et B avec leurs centres respectifs, les nœuds a et b:

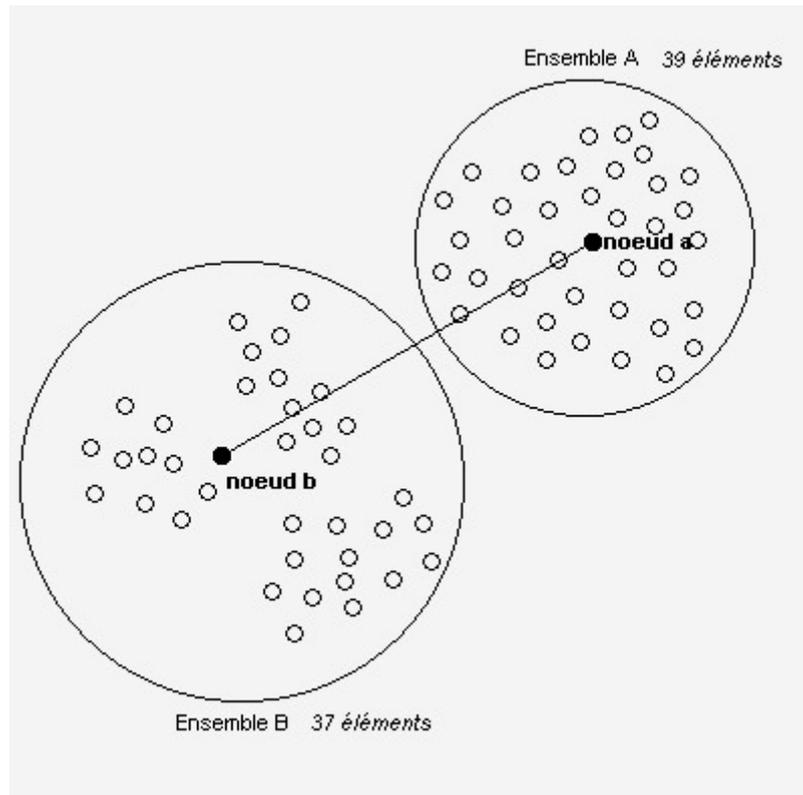


Fig. 2.17 - Partition des 76 villes en 2 ensembles

Prenons comme poids des nœuds a et b de ce réseau pondéré, l'importance de la population représentée respectivement par le nombre de villes des deux ensembles soit 39 villes et 37 villes. Alors la localisation optimale d'une activité unique selon le problème 1-médian sera choisie au nœud a car il a un coefficient de pondération plus important. Mais, si l'on choisit un nombre de partitions égal à 4, on obtient alors la répartition suivante avec chacun des nœuds reliés aux nœuds de ceux des autres ensembles les plus proches:

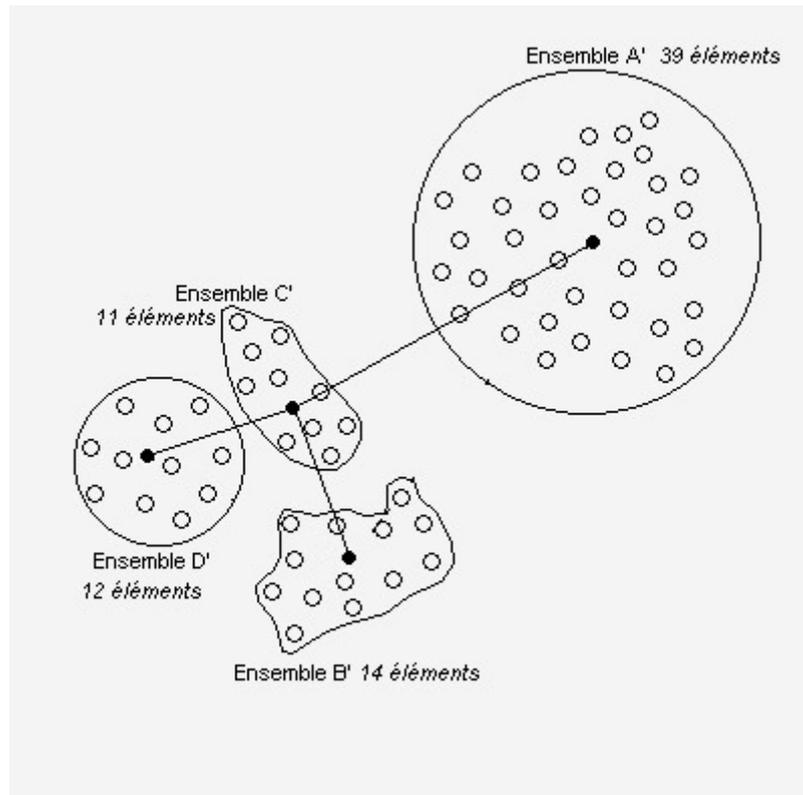


Fig. 2.18 - Partition des 76 villes en 4 ensembles

Le centre optimal est alors situé sur le nœud de l'ensemble A' qui correspond bien au nœud a de l'ensemble A précédemment. Si on considère maintenant le contre-exemple suivant avec 3 partitions de 17 villes, la droite verticale passant par O étant la droite médiane du segment ab:

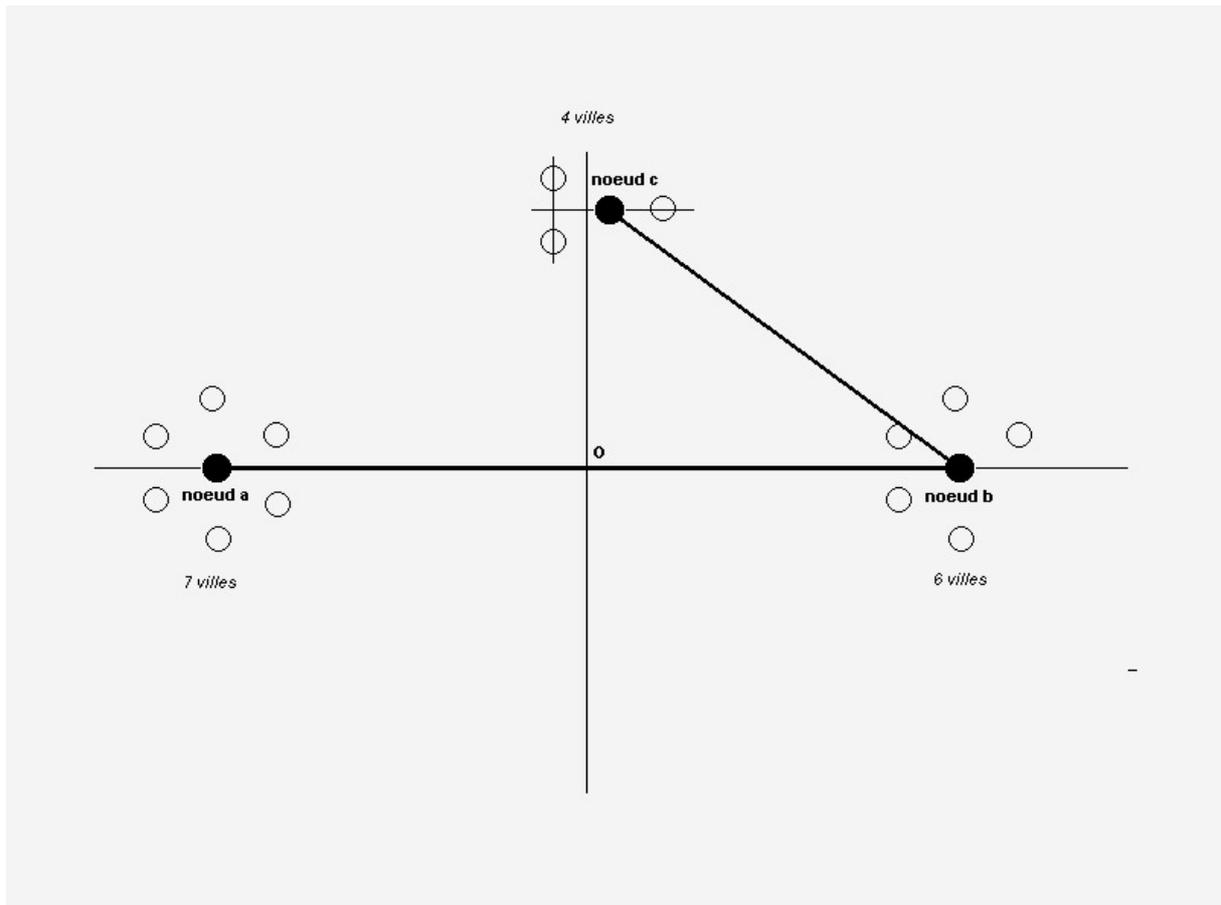


Fig. 2.19 - Partition de 17 villes en 3 ensembles

Selon le modèle 1-médian, le centre optimal est situé au nœud b. En considérant une partition en deux ensembles, le centre optimal est cette fois situé au nœud a puisque que sa partition a une pondération plus élevée en nombre de villes (9 villes):

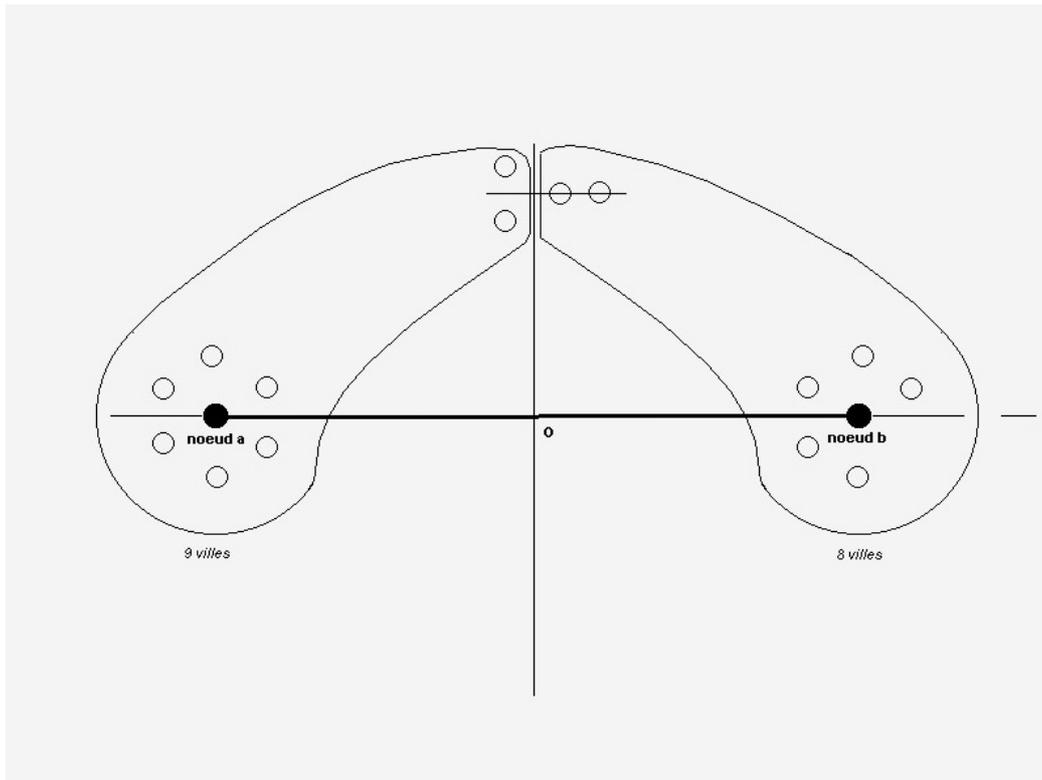


Fig. 2.20 - Partition des 17 villes en 2 ensembles

De plus, on peut se poser la question de savoir si certaines villes, ou paquets de villes, très éloignées valent ou non la peine d'être prises en compte dans l'étude de localisation: quel que soit leur potentiel de clientèle même faible, leur éloignement a tendance à attirer vers elles le (ou les) site(s) à implanter sans que cela ne vaille peut-être la peine. Dans l'exemple précédent, la partition constituée des quatre villes du nœud c est d'une importance moindre devant les deux autres partitions des nœuds a (7 villes) et b (6 villes). Si le problème était de livrer un produit ou un service à partir du nœud a à un ensemble de clients, faudrait-il intégrer les clients associés au nœud c qui, compte tenu du faible potentiel et du coût de déplacement, engendrerait une faible rentabilité ? La réponse à cette question n'est pas fournie par l'analyse typologique et ses méthodes de partition qui traitent le problème d'une manière brute et mathématique.

Outre le fait de fournir des résultats rationnels, le développement de méthodes plus performantes dans la recherche de localisations optimales doit conduire à notre sens, à

améliorer la capacité du p-médian à traiter des réseaux de plus en plus importants afin de modéliser de manière plus précise le monde réel et pouvoir prendre en compte par exemple l'ensemble des clients d'une grande surface, le réseau Internet et ses nombreux ordinateurs, la totalité des entrepôts d'un transporteur,...