7e Conférence Internationale de MOdélisation et SIMulation - MOSIM'08 - du 31 mars au 2 avril 2008 – Paris- France
« Modélisation, Optimisation et Simulation des Systèmes : Communication, Coopération et Coordination.».

RESOLUTION DU PROBLEME DE CLUSTERING VIA UNE METHODE HYBRIDE

Youssef MASMOUDI, Habib CHABCHOUB

Laboratoire LOGIQ, Institut Supérieur de Gestion Industrielle de Sfax, Route Mharza km 1.5, Boite postale 954, 3018 Sfax, Tunisie
youssef_m_tn@yahoo.fr,
habib.chabchoub@fsegs.rnu.tn

Saïd HANAFI

Laboratoire L.A.M.I.H équipe ROI, Université de Valenciennes et du Hainaut - Cambrésis (UVHC) – 59393 Valenciennes France
said.hanafi@univ-valenciennes.fr

RESUME : *Le problème de clustering consiste à rassembler des éléments dans des classes cohérentes. Ce problème a été traité par plusieurs méthodes de résolutions, nous citons par exemple, des méthodes mathématiques, des méthodes statistiques et des méthodes informatiques. Le problème de clustering a plusieurs applications dans divers contextes, comme par exemple dans les problèmes de datamining, bioinformatiques, finance etc. Dans ce papier nous considérons le problème de clustering qui consiste à minimiser la distance entre les éléments qui composent le problème avec les centres des clusters. Il s'agit donc de déterminer les bons centres de clusters.*
Premièrement, nous proposons une formulation du problème de clustering sous forme d'un programme linéaire avec des variables mixtes. Le modèle mathématique proposé est implémenté avec le langage C++ en utilisant la bibliothèque commerciale CPLEX dans sa version 10.0.
Dans la deuxième partie de ce papier, l'heuristique k-means pour le problème de clustering va être proposé. Afin d'améliorer le temps de résolution du problème de clustering avec le modèle mathématique que nous proposons, et pour mieux initialiser les paramètres du programme mathématique proposé, nous avons opté pour initialiser CPLEX avec une solution réalisable en utilisant l'heuristique k-means, ce qui sera présenté dans la partie suivante. Enfin, dans la dernière partie de ce papier, une validation de l'approche proposée sera présentée avec des résultats expérimentaux.

MOTS-CLES : *Classification, Clustering, CPLEX, k-means, Programmation mathématique, approche hybride*

1. INTRODUCTION

Le problème de clustering consiste à partitionner un ensemble de m points dans un espace de dimension n dans k groupes (clusters) tout en optimisant un critère donné. On considère le problème de clustering qui consiste à déterminer k centres des clusters qui minimisent la somme des distances entre chaque point et le centre le plus proche (P. S. Bradley et *al.*). Plusieurs méthodes existent dans la littérature selon la distance utilisée. Par exemple l'algorithme k-median (P. S. Bradley et *al.*) utilise la distance induite par la norme 1 (resp. la norme 2). Pour les deux normes 1 et 2, le problème de clustering est NP-difficile. Le problème de clustering apparaît dans plusieurs applications (Alan H. Filding, Richard O. Duda et *al.* et Saman K. Halgamuge et *al.*) :

- Marketing : segmentation du marché en découvrant des groupes de clients distincts à partir de bases de données d'achats.
- Environnement : identification des zones terrestres similaires (en termes d'utilisation) dans une base de données d'observation de la terre.
- Assurance: identification de groupes d'assurés distincts associés à un nombre important de déclarations.
- Planification de villes : identification de groupes d'habitations suivant le type d'habitation, valeur, localisation géographique.
- Médecine : Localisation de tumeurs dans le cerveau.

Plusieurs méthodes ont été proposées dans la littérature pour résoudre le problème de clustering dans différents contextes, notamment celle de O. L. Mangasarian qui sera présentée dans le paragraphe suivant.

2. MODELE MATHEMATIQUE POUR LE PROBLEME DE CLUSTERING

2.1. Un modèle bilinéaire pour le problème de clustering

Etant donnés un ensemble de m points $\{A^1, A^2, \ldots, A^n\}$ dans l'espace réel de dimension n (A^i dans R^n pour $i = 1, \ldots, m$) et un entier fixe k de clusters, le problème de clustering consiste à déterminer k centres dans R^n $\{x^1, \ldots, x^k\}$ tels que la somme des distances entre chaque point et le plus proche centre soit minimisée. Plus formellement, le problème de clustering est :

(P1)

$$\min\left\{\sum_{i=1}^{m}\min\left\{\left\|A^i - x^l\right\| \;: l=1,...,k\right\} \;: x^1,...,x^k \text{ } k \text{ centres}\right\}$$

où ||.|| est une norme arbitraire dans R^n. Notons que P1 est un problème d'optimisation difficile puisque la fonction à minimiser n'est ni convexe ni concave.
Bradley et *al.* ont proposé un programme bilinéaire pour la norme L1 en se basant sur le lemme suivant :
Lemme 1 : Soient a un vecteur dans R^k, e est un vecteur de composantes égales à 1 avec la dimension appropriée et e^l le lème vecteur unité pour l= 1, …., k, on a

$$\min\{a^T e^l : l=1, \ldots, k\} = \min\{a^T y : e^T y = 1, y \geq 0\} \qquad (1)$$

Le lemme 1 permet d'éliminer le minimum dans la somme de la fonction objective en introduisant des variables de sélection y_i dans R^k. Donc le problème P1 peut être reformulé comme suit :

$$\text{(P2)} \qquad \min\left\{\sum_{i=1}^{m}\sum_{l=1}^{k} y_{il}\left\|A^i - x^l\right\| \;: e^T y_i = 1, y_i \geq 0\right\}$$

Maintenant en prenant la norme L1 et en introduisant les variables artificielles z_{il} dans R^n qui encadrent la valeur absolue de la différence $A^i - x^l$, le problème P2 peut être reformulé comme suit :

$$\text{(P3)} \begin{cases} \min & \sum_{i=1}^{m}\sum_{l=1}^{k} y_{il}(e^T z_{il}) & \\ s.c. & -z_{il} \leq A^i - x^l \leq z_{il} & \\ & \sum_{l=1}^{k} y_{il} = 1 & i=1,...,m \\ & y_{il} \geq 0 & i=1,...,m; l=1,...,k \end{cases}$$

2.2 Linéarisation

Dans cette section nous proposons un programme linéaire en variables mixtes pour le problème de clustering. Le minimum dans le problème peut être remplacé par les variables de sélection ai et les variables binaires y_{il} sont introduites pour exprimer que ai est le minimum des $\| A^i - x^l \|$. Ainsi le problème P1 peut être reformulé comme suit :

(Q1)

$$\begin{cases} \min & \sum_{l=1}^{m} \alpha_i & \\ s.c. & \alpha_i \leq \left\|A^i - x^l\right\| \leq \alpha_i + M(1-y_{il}) & i=1,...,m; l=1,...,k \\ & \sum_{l=1}^{k} y_{il} = 1 & i=1,...,m \\ & y_{il} \in \{0,1\} & i=1,...,m; l=1,...,k \\ & \alpha_i \geq 0 & i=1,...,m \end{cases}$$

où M est est un paramètre assez grand. Pour trouvez sa bonne valeur, il a été sujet à des variations de valeur avec différentes instances. La constatation qui a été faite est que ce paramètre dépend de la grandeur de l'instance et de l'espacement entre ses objets. De ce fait, sa détermination sera faite avec l'heuristique k-means dont la démarche sera expliquée dans les paragraphes suivants.
Comme précédemment, en prenant la norme L1 et en introduisant les variables artificielles z_{il} dans R^n qui encadrent la valeur absolue de la différence $A^i - x^l$, le problème Q1 peut être reformulé comme suit :

(Q2)

$$\begin{cases} \min & \sum_{l=1}^{m} \alpha_i & \\ s.c. & \alpha_i \leq ez_{il} \leq \alpha_i + M(1-y_{il}) & i=1,...,m; l=1,...,k \\ & \sum_{l=1}^{k} y_{il} = 1 & i=1,...,m \\ & -z_{il} \leq A^i - x^l \leq z_{il} & i=1,...,m; l=1,...,k \\ & y_{il} \in \{0,1\} & i=1,...,m \\ & z_{il}, \alpha_i \geq 0 & i=1,...,m; l=1,...,k \end{cases}$$

3. L'HEURISTIQUE K-MEANS POUR LE PROBLEME DE CLUSTERING

L'heuristique k-means est une méthode de clustering pour grouper des objets en se basant sur des attributs dans *k* groupes (avec *k* un nombre positif). C'est une technique de classification qui permet d'organiser un ensemble de données en classes cohérentes et homogènes. Elle s'applique, sur n'importe quel type de données : tableau de contingence, tableau de distances, etc. Cette méthode, largement utilisée dans la littérature, a été appliquée surtout dans le domaine de classification de documents et de données et la segmentation d'images. Le groupement est donné par la minimisation de la distance entre les objets et les centres des clusters concernés (JOHN A. Hartigan).

Son principe est le suivant :

1. choix d'une métrique pour le calcul des distances (euclidienne,…).
2. définition d'un nombre k de classes sur un ensemble d'échantillons.
3. initialisation aléatoire des k centres de gravité (centroïdes).
4. affectation de chaque échantillon à son centre le plus proche suivant la métrique choisie.
5. calcul des nouveaux centres suivant les affectations effectuées à l'étape précédente.
6. répétition des étapes 4 et 5 jusqu'à ce que la position des centres n'évolue plus.

Cette méthode de classification, qui a donné des résultats de bonne qualité et qui a prouvé une robustesse dans plusieurs contextes, présente aussi des faiblesses qui proviennent principalement de l'initialisation aléatoire des premiers centroïdes, ce qui affectera par la suite la

solution finale de cette heuristique. Autre faiblesse de cette méthode c'est qu'elle ne tient pas en compte de l'équilibrage des objets à grouper (Y. Masmoudi et *al.*).

Dans le contexte de ce travail, cette heuristique a été implémentée pour la résolution du problème de clustering. En effet, k-means fournit une bonne solution qui sera utilisé pour deux buts se rapportant à la résolution exacte du problème formulé selon le modèle (Q2). Le premier but est d'initialiser le solveur CPLEX avec une bonne solution, ce qui va lui permettre de gagner énormément sur le temps de résolution. Ainsi, le résultat donné par k-means va affecter les valeurs des T_{il} . Le second but est de déterminer les valeurs du paramètre *M* comme sera expliqué dans le paragraphe suivant.

4. HYBRIDATION HEURISTIQUE - METHODE EXACTE

Pour expérimenter l'approche qui a été présentée dans les paragraphes précédents, plus précisément le modèle (Q2), l'implémentation de ce modèle a été faite avec le langage C++ en utilisant la bibliothèque commerciale de CPLEX pour sa résolution. Pour la valeur du paramètre *M*, son initialisation été faite au départ avec des valeurs assez diverses. Ce qui a été remarqué est qu'il dépond de la grandeur de l'instance et de l'espacement des objets à classifier.

Pour ce fait, l'heuristique k-means a été utilisé pour trouver la valeur du paramètre *M*. En considérant les clusters fournis par k-means, *M* a été déterminé comme la valeur de la plus grande distance entre un point d'un cluster et son centre de gravité. Ainsi nous obtenons *k* valeurs de *M* noté par M_l .

Quand aux valeurs des T_{il} , ils étaient initialisé selon la classification donnée par k-means. Par la suite, la démarche d'optimisation avec CPLEX est lancée. Une fois les valeurs des variables du problème furent déterminées, l'opération d'affectation des objets à classifier commence. Cette affectation se fait selon la règle du centre de classe le plus proche. Pour ce fait, pour chaque point est déterminé son eD_{il} et trouver le minimum par rapport aux centres des classes, ce qui revient à trouver $\min\left\{\sum_{j}^{n} D_{ilj}\right\}$.

5. VALIDATION DE LA METHODE

L'approche proposée est résumé dans l'algorithme 1. Cet algorithme est implémenté avec le langage C++ sous une machine équipé d'un processeur Intel Core 2 Duo T5200 1.60GHz et d'un mémoire de 1024MB, en utilisant la bibliothèque CPLEX dans sa version 10.0. Plusieurs benchmark ont servis pour le test de cet algorithme, notamment des instances de la littérature et des instances générées qui représentent des formes géométriques. Une instance de la littérature utilisé s'appelle « butterfly » (Bezdek, J. C.) qui est disponible à l'adresse suivante http://www.xplore-stat.de/data/butterfly.dat. Cette instance est composée de 15 point à deux dimensions comme le montre la figure 1. Avec un nombre de cluster égal à deux, l'optimalité sera représenté comme l'illustration de la figure 2.

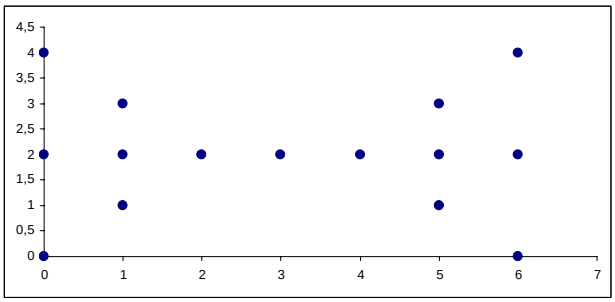


Figure 1. Représentation de l'instance « butterfly »

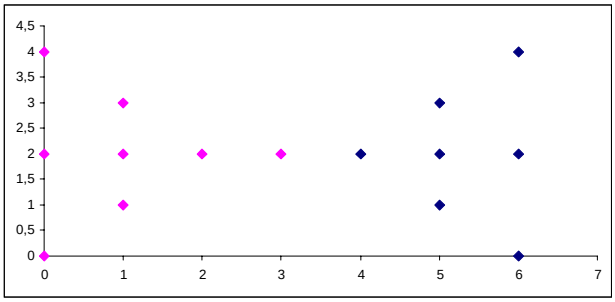
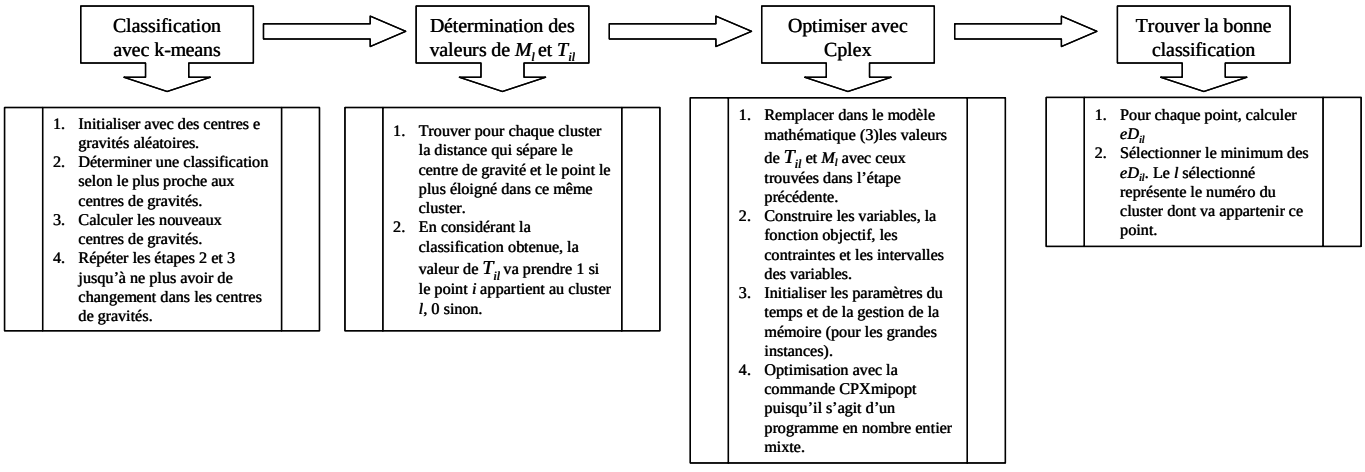


Figure 2. Optimalité avec k=2



Algorithme 1. Méthode hybride pour le problème de classification

Vu que c'est une petite instance, le rôle de k-means va s'arrêter dans la détermination des valeurs des M_l. L'initialisation avec une classification initiale fournit par k-means sera ignoré. Avec les valeurs de M_l et les coordonnées des points, un programme mathématique sera généré et résolut automatiquement par CPLEX. Les valeurs de M_l et les résultats fournis par CPLEX sont données dans le tableau 1.

	Valeur de M	*Points contenus dans le cluster*
Cluster 1	5.385165	1, 2, 3, 4, 5, 6, 7, 8
Cluster 2	2.123724	9, 10, 11, 12, 13, 14, 15

Tableau 1. Résultat de « butterfly »

La classification obtenue est la même que celle de la figure 2, ce qui valide l'approche proposée. Le temps d'exécution est de 0,141000 secondes et la valeur de la fonction objectif est 30,38137947.

Les autres instances de la littérature qui sont utilisés pour le test de notre algorithme sont de taille beaucoup plus importante que celle évoqué précédemment.
Une instance de taille 138 qui concerne la méningite et une autre qui concerne les iris de taille 150 sont téléchargeable à partir de http://stats.math.uni-augsburg.de/Klimt/down.html
Vu que ce sont des instances de grande taille, k-means va servir pour déterminer les valeurs de M_l et aussi pour fournir une classification initiale. De même, avec les valeurs de M_l et les mesures des différents points, un programme mathématique sera généré automatiquement par CPLEX. Pour le benchmark de l'iris, la valeur de *k* est égale à 3. Le tableau 2 résume les valeurs de M_l et les résultats de classification. Le temps d'exécution est limité à 5400 secondes, et la valeur de la fonction objectif est 331,7851269. Quand au benchmark de la méningite, la valeur de *k* est 2. Le tableau 3 résume les valeurs de M_l et les résultats de classification. Le temps d'exécution est limité à 7200 secondes, et la valeur de la fonction objectif est 715577,3316.

	Valeur de M	*Points contenus dans le cluster*
Cluster 1	3.412511	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, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 58, 61, 80, 94, 99
Cluster 2	1.248031	54, 56, 60, 62, 63, 65, 67, 68, 70, 72, 81, 82, 83, 85, 89, 90, 91, 93, 95, 96, 97, 100
Cluster 3	5.031328	51, 52, 53, 55, 57, 59, 64, 66, 69, 71, 73, 74, 75, 76, 77, 78, 79, 84, 86, 87, 88, 92, 98, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148, 149, 150

Tableau 2. Résultat de l' « iris »

	Valeur de M	*Points contenus dans le cluster*
Cluster 1	19366.768392	2, 3, 4, 7, 8, 9, 10, 11, 12, 13, 14, 16, 19, 20, 21, 22, 25, 26, 28, 29, 31, 32, 33, 34, 35, 36, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 59, 60, 61, 62, 63, 64, 65, 67, 68, 69, 70, 71, 72, 73, 75, 76, 78, 79, 81, 82, 83, 84, 85, 86, 87, 89, 90, 91, 93, 94, 95, 98, 99, 100, 101, 102, 104, 105, 106, 109, 111, 113, 114, 115, 116, 118, 119, 120, 121, 122, 123, 124, 125, 126, 128, 129, 131, 132, 133, 134, 135, 136, 137
Cluster 2	40372.225546	1, 5, 6, 15, 17, 18, 23, 24, 27, 30, 37, 58, 66, 74, 77, 80, 88, 92, 96, 97, 103, 107, 108, 110, 112, 117, 127, 130, 138

Tableau 3. Résultat de la « méningite »

D'autres instances ont été générées manuellement. Ces instances représentent des formes géométriques : losange, parallélogramme, rectangle et trapèze (voir Annexe I). Les valeurs de k qui sont utilisé sont 2, 3 et 4. Les tableaux 4, 5, 6 et 7 regroupent les résultats des tests. Le temps d'exécution a été limité au maximum à 5400 secondes.

6. CONCLUSION ET PERSPECTIVE

Ce papier présente un programme linéaire en variables mixtes pour le problème de clustering que nous avons implémenté avec CPLEX pour trouver les bons centres des clusters et affecter tous les points aux clusters correspondants. L'heuristique k-means a servi pour initialiser les différents paramètres du programme mathématique à résoudre. Ce travail a été validé avec une instance de la littérature.
Etant validée, l'application de l'approche proposée sera très intéressante dans les différents types de problèmes de classification. L'un des domaines les plus intéressants de son application est la classification des tumeurs cancéreuses.

Losange (25 points)

Nombre de cluster	k=2		k=3		k=4	
Temps (s)	75.45400238		5400.375		5400.140137	
Fonction Objectif	45		38		33.26627385	
	Valeur de M	*Points contenus dans le cluster*	*Valeur de M*	*Points contenus dans le cluster*	*Valeur de M*	*Points contenus dans le cluster*
Cluster 1	2.236068	3, 4, 7, 8, 9, 13, 14, 15, 16, 20, 21, 24	4.258615	2, 5, 6, 10, 11, 12, 17, 18, 22	3.605551	22, 23, 24, 25
Cluster 2	3.866391	1, 2, 5, 6, 10, 11, 12, 17, 18, 19, 22, 23, 25	1.832491	1, 3, 7, 13, 19, 23, 25	4.714286	4, 8, 9, 14, 15, 16, 20, 21
Cluster 3			3.833259	4, 8, 9, 14, 15, 16, 20, 21, 24	1.577909	1, 3, 7, 13, 19
Cluster 4					3.374743	2, 5, 6, 10, 11, 12, 17, 18

Tableau 4. Résultat du « Losange »

Parallélogramme (30 points)

Nombre de cluster	k=2		k=3		k=4	
Temps (s)	199.371994		5392.890137		5392.100098	
Fonction Objectif	63.80485012		53.07765074		50.48696492	
	Valeur de M	*Points contenus dans le cluster*	*Valeur de M*	*Points contenus dans le cluster*	*Valeur de M*	*Points contenus dans le cluster*
Cluster 1	3.065050	1, 2, 3, 4, 5, 6, 7, 8, 9, 11, 12, 13, 16, 17, 21	8.181612	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12	7.997070	1, 2, 3, 4, 5, 6, 7, 8, 9
Cluster 2	7.060236	10, 14, 15, 18, 19, 20, 22, 23, 24, 25, 26, 27, 28, 29, 30	2.213594	13, 16, 17, 18, 21	5.003124	11, 12, 16, 17, 18, 21
Cluster 3			5.265158	14, 15, 19, 20, 22, 23, 24, 25, 26, 27, 28, 29, 30	2.042753	10, 13, 14, 15
Cluster 4					5.403702	19, 20, 22, 23, 24, 25, 26, 27, 28, 29, 30

Tableau 5. Résultat du « Parallélogramme »

Rectangle (30 points)

Nombre de cluster	k=2		k=3		k=4	
Temps (s)	65.875		5400.625		1869.765991	
Fonction Objectif	59.81966011		50.93844719		52.22949017	
	Valeur de M	*Points contenus dans le cluster*	*Valeur de M*	*Points contenus dans le cluster*	*Valeur de M*	*Points contenus dans le cluster*
Cluster 1	4.472136	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15	3.201562	18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30	3.922723	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 13
Cluster 2	2.236068	16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30	2.061553	11, 12, 13, 14, 15, 16, 17	1.118034	18, 23, 28
Cluster 3			4.924429	1, 2, 3, 4, 5, 6, 7, 8, 9, 10	3.162278	14, 15, 19, 20, 24, 25, 29, 30
Cluster 4					5.175060	11, 12, 16, 17, 21, 22, 26, 27

Tableau 6. Résultat du « Rectangle »

Trapèze (23 points)

Nombre de cluster	k=2		k=3		k=4	
Temps (s)	5.888000011		5400.024902		1169.140991	
Fonction Objectif	42.64589803		34.80391925		31.21919413	
	Valeur de M	*Points contenus dans le cluster*	*Valeur de M*	*Points contenus dans le cluster*	*Valeur de M*	*Points contenus dans le cluster*
Cluster 1	1.677051	1, 3, 4, 8, 9, 13, 14, 18, 19, 21, 22, 23	1.839216	1, 3, 4, 8, 9, 13, 14	4.307090	5, 6, 10, 11, 15, 16
Cluster 2	4.294182	2, 5, 6, 7, 10, 11, 12, 15, 16, 17, 20	4.791574	15, 16, 17, 18, 19, 20, 21, 22, 23	1.343710	12, 13, 14
Cluster 3			4.449260	2, 5, 6, 7, 10, 11, 12	4.000000	17, 18, 19, 20, 21, 22, 23
Cluster 4					4.770744	1, 2, 3, 4, 7, 8, 9

Tableau 7. Résultat du « Trapèze »

REFERENCES

Alan H. Filding, 2007. Cluster and classification techniques for the Biosciences, Cambridge University Press.

Bezdek, J. C., 1981. Pattern Recognition with Fuzzy Objective Function Algorithms, Plenum Press, New York.

ILOG CPLEX 10.0 distribution, including Concert Technology for C++

JOHN A. Hartigan, 1975. Clustering Algorithms. Wiley series in probability and mathematical statistics.

K-Means Clustering Tutorial (mai 2007)
http://people.revoledu.com/kardi/tutorial/kMean/index.html

O. L. Mangasarian, Mathematical Programming in Data Mining, Data Mining and Knowledge Discovery, 1(2), 1997, 183-201.

P. S. Bradley, O. L. Mangasarien, and W. N. Street, Clustering via concave minimisation, Advances in Neural Information Processing Systems 9, MIT Press, Cambridge, MA 1997.

Richard O. Duda, Peter E. Hart, David G. Stork. Pattern Classification (Second Edition), Wiley Interscience.

Saman K. Halgamuge, Lipo Wang, 2005. Classification and clustering for knowledge discovery, Spring.

Y. Masmoudi et H. Chabchoub, 2007. Une méthode d'équilibrage des clusters issus de k-means. Conférence scientifique conjointe FRANCORO /ROADEF2007, du 20 au 23 février 2007, Grenoble, France, pp. 309-310.

ANNEXE I

Losange (25 points)

1	-3	0
2	-2	-1
3	-2	0
4	-2	1
5	-1	-2
6	-1	-1
7	-1	0
8	-1	1
9	-1	2
10	0	-3
11	0	-2
12	0	-1
13	0	0
14	0	1
15	0	2
16	0	3
17	1	-2
18	1	-1
19	1	0
20	1	1
21	1	2
22	2	-1
23	2	0
24	2	1
25	3	0

Parallélogramme (30 points)

1	0	0
2	1	0
3	1	1
4	2	0
5	2	1
6	2	2

7	3	0
8	3	1
9	3	2
10	3	3
11	4	0
12	4	1
13	4	2
14	4	3
15	4	4
16	5	0
17	5	1
18	5	2
19	5	3
20	5	4
21	6	1
22	6	2
23	6	3
24	6	4
25	7	2
26	7	3
27	7	4
28	8	3
29	8	4
30	9	4

Rectangle (30 points)

1	0	0
2	0	1
3	0	2
4	0	3
5	0	4
6	1	0
7	1	1
8	1	2
9	1	3
10	1	4
11	2	0
12	2	1
13	2	2
14	2	3
15	2	4
16	3	0
17	3	1
18	3	2
19	3	3
20	3	4
21	4	0
22	4	1
23	4	2
24	4	3
25	4	4
26	5	0
27	5	1
28	5	2
29	5	3
30	5	4

Trapèze (23 points)

1	0	4
2	1	2
3	1	3
4	1	4
5	2	0
6	2	1
7	2	2
8	2	3
9	2	4
10	3	0
11	3	1
12	3	2
13	3	3
14	3	4
15	4	0
16	4	1
17	4	2
18	4	3
19	4	4
20	5	2
21	5	3
22	5	4
23	6	4