École des Ponts
ParisTech

Entrepôts, bin-packing et sac-à-dos

CERMICS,
ENPC

A. Parmentier

12 décembre 2018

Miniprojet

Bilan sur

- séances en petite classe
- séance en autonomie

Prochaines séances en amphi

- 1h30 cours
- 1h intervenant industriel

Prochaine séance en autonomie : 16 janvier

- sur l'ordonnancement
- sous forme d'exercices de révision
- sujet une semaine avant pour que vous puissiez le travailler
- venez poser des questions

Sommaire de la partie

1. Sac à dos

2. Bin Packing

3. Positionnement d'entrepôts

4. Exercices

PROBLÈME DU SAC-À-DOS

Données. Entiers positifs ou nuls n, $w_1, \ldots, w_n$ et W, et des réels positifs ou nuls $c_1, \ldots, c_n$.

Tâche. Trouver un sous-ensemble $S \subseteq \{1, \ldots, n\}$ tel que $\sum_{j \in S} w_j \leq W$ et $\sum_{j \in S} c_j$ est maximum.

Applications (et généralisations en plusieurs dimensions)

1. Gestion de portefeuilles / investissements
2. Génération de clefs en cryptographie
3. Choix de questions à traiter dans un exam noté sur plus de 20.
4. Technologie embarquée.
5. Etc.

Complexité

Théorème

Le problème du sac-à-dos est **NP***-difficile.*

Programmation dynamique

Quelle fonction valeur pourrait être utilisée pour une approche par programmation dynamique ?

Programmation dynamique

$x(j, W')$ = valeur maximale d'un sac de capacité W', en se limitant aux j premiers objets.

Equation (programmation dynamique)

$$x(j, W') = \max(x(j-1, W'), x(j-1, W' - w_j) + c_j).$$

Complexité : $O(nW)$.

Question

Pourquoi l'algorithme n'est il pas polynomial ?

Programmation dynamique

$x(j, W')$ = valeur maximale d'un sac de capacité W', en se limitant aux j premiers objets.

Equation (programmation dynamique)

$$x(j, W') = \max(x(j-1, W'), x(j-1, W' - w_j) + c_j).$$

Complexité : $O(nW)$.

Question

Pourquoi l'algorithme n'est il pas polynomial ?

Ce n'est pas polynomial car W nécessite $\log_2(W)$ bits pour être codés. C'est un algorithme **pseudo-polynomial**.

Formulation sous forme de programme linéaire en nombres entiers

Donner un programme linéaire en nombres entiers pour ce problème.

Formulation sous forme de programme linéaire en nombres entiers

Donner un programme linéaire en nombres entiers pour ce problème.

$$
\begin{aligned}
\max \quad & \sum_{j=1}^{n} c_j x_j \\
\text{s.c.} \quad & \sum_{j=1}^{n} w_j x_j \leq W \\
& x_j \in \{0, 1\} \qquad \text{pour tout } j \in \{1, \ldots, n\}
\end{aligned}
$$

Calcul de la relaxation linéaire

La borne supérieure

$$\begin{aligned} \max \quad & \sum_{j=1}^{n} c_j x_j \\ \text{s.c.} \quad & \sum_{j=1}^{n} w_j x_j \leq W \\ & 0 \leq x_j \leq 1 \qquad \text{pour tout } j \in \{1, \ldots, n\} \end{aligned}$$

peut facilement se calculer avec un solver.

Exercice

Donner un algorithme très simple qui calcule la solution du relâché continu.

Solution du relâché continu

On suppose que $\sum_{j=1}^{n} w_j > W$ et que pour tout j, $w_j \leq W$, sinon trivial !

- Classer les objets de façon à ce que

$$\frac{c_1}{w_1} \geq \frac{c_2}{w_2} \geq \ldots \geq \frac{c_n}{w_n}.$$

- Poser $x_1 = x_2 = \cdots = x_{\bar{j}} := 1$ avec $\bar{j}$ le plus grand entier tel que $\sum_{j=1}^{\bar{j}} w_j \leq W$.
- Poser $x_{\bar{j}+1} := \frac{W - \sum_{j=1}^{\bar{j}} w_j}{w_{\bar{j}+1}}$.
- Poser $x_{\bar{j}+2} = \ldots = x_n := 0$.

Lemme

Un tel x est solution optimale de la relaxation.

Heuristique avec garantie

Algorithme fournissant une solution valant toujours au moins la moitié de la valeur optimale :

On suppose qu $w_j \leq W$ pour tout j.

On reprend l'indexation par $\frac{c_j}{w_j}$ et le $\bar{j}$ du relâché linéaire.

- $\sum_{j=1}^{\bar{j}+1} c_j$ est une borne supérieure de la valeur optimale OPT.
- $\{1, \ldots, \bar{j}\}$ et $\{\bar{j}+1\}$ sont deux solutions réalisables, l'une des deux étant par conséquent au moins de valeur $\geq \frac{1}{2} \sum_{j=1}^{\bar{j}+1} c_j$, i.e. de valeur $\geq \frac{1}{2} OPT$.

Sommaire de la partie

1. Sac à dos

2. Bin Packing

3. Positionnement d'entrepôts

4. Exercices

Formalisation du bin-packing

BIN-PACKING

Données. Entiers positifs ou nuls $a_1, \ldots, a_n$ et W.

Tâche. Trouver un entier naturel k et une fonction $f : [n] \to [k]$ avec $\sum_{i:\, f(i)=j} a_i \leq W$ pour tout $j \in [k]$ tels que k soit minimum.

Interprétation : conteneurs de taille W, objets de tailles a_i. Mettre tous les objets dans un nombre minimum de conteneurs (f indique l'affectation des objets aux conteneurs).

Applications (et généralisation en plusieurs dimensions)

Application 1. Stockage optimal (entrepôts), transport optimal (camions) ; souvent, il y a des contraintes de forme ou de volume en plus (bin-packing 2D ou 3D)

Application 2. Découpe : on a des poutres de longueur fixée, on a une commande pour des longueurs $\boldsymbol{a_1}, \ldots, \boldsymbol{a_n}$; minimiser le nombre de poutres utilisées.

Application 3. Copier un disque dur avec le moins de clés USB possible.

Complexité

Théorème

BIN-PACKING est **NP***-difficile.*

Un cas particulier facile

Exercice

Supposons qu'un instance $a_1, \ldots, a_n$ du bin-packing soit telle que $\frac{a_i}{W} > \frac{1}{3}$ pour tout $i = 1, \ldots, n$.

Donner un algorithme polynomial pour résoudre ce cas particulier.

Un cas particulier facile

Exercice

Supposons qu'un instance $a_1, \ldots, a_n$ du bin-packing soit telle que $\frac{a_i}{W} > \frac{1}{3}$ pour tout $i = 1, \ldots, n$.

Donner un algorithme polynomial pour résoudre ce cas particulier.

Couplage biparti

Borne inférieure

BIN-PACKING

Données. Entiers positifs ou nuls $a_1, \ldots, a_n$ et W.

Tâche. Trouver un entier naturel k et une fonction $f : [n] \to [k]$ avec $\sum_{i:\, f(i)=j} a_i \leq W$ pour tout $j \in [k]$ tels que k soit minimum.

Donner une borne inférieure simple à la valeur de BIN-PACKING

BIN-PACKING

Données. Entiers positifs ou nuls $a_1, \ldots, a_n$ et W.

Tâche. Trouver un entier naturel k et une fonction $f : [n] \to [k]$ avec $\sum_{i:\, f(i)=j} a_i \leq W$ pour tout $j \in [k]$ tels que k soit minimum.

Donner une borne inférieure simple à la valeur de BIN-PACKING

$$\left\lceil \sum_{i=1^n} \frac{a_i}{W} \right\rceil \leq OPT$$

Formulation programme linéaire

Donner un PLNE modélisant le problème

Formulation programme linéaire

Donner un PLNE modélisant le problème

On suppose que l'on a k conteneurs disponibles.

$$
\begin{aligned}
\min \quad & \sum_{j=1}^{k} z_j \\
\text{s.c.} \quad & \sum_{j=1}^{k} y_{ji} = 1 && i = 1, \dots, n \\
& \sum_{i=1}^{n} a_i y_{ji} \leq W z_j && j = 1, \dots, k \\
& y_{ji}, z_j \in \{0, 1\} && i = 1, \dots, n;\ j = 1, \dots, k
\end{aligned}
$$

où $z_j = 1$ si le conteneur j est utilisé et $y_{ji} = 1$ si l'objet i est mis dans le conteneur j.

Heuristiques

1. NEXT-FIT : **2**-approximation.
2. FIRST-FIT : $\frac{17}{10}$-approximation.
3. FIRST-FIT DECREASING : $\frac{3}{2}$-approximation.

Peut-on faire mieux ?

Heuristiques

1. NEXT-FIT : **2**-approximation.
2. FIRST-FIT : $\frac{17}{10}$-approximation.
3. FIRST-FIT DECREASING : $\frac{3}{2}$-approximation.

Théorème

A moins que $\boldsymbol{P} = \boldsymbol{NP}$ il n'existe pas d'algorithme polynomial garantissant une approximation $\rho < \frac{3}{2}$

Exercice

1. Prouver que l'existence d'une solution à 2 bin est NP-complet en le réduisant depuis le problème partition.
2. En déduire le théorème.

NEXT-FIT

Je prends les objets les uns après les autres. Dès que l'objet $\boldsymbol{i}$ ne peut pas entrer dans le conteneur courante, je passe à un nouveau conteneur.

Théorème

NEXT-FIT fournit une solution $\boldsymbol{SOL}$ *telle que*

$$\boldsymbol{SOL \leq 2OPT - 1.}$$

Sur quoi repose la preuve ?

NEXT-FIT

Je prends les objets les uns après les autres. Dès que l'objet i ne peut pas entrer dans le conteneur courante, je passe à un nouveau conteneur.

Théorème

NEXT-FIT fournit une solution $\mathbf{SOL}$ *telle que*

$$\mathbf{SOL} \leq 2\mathbf{OPT} - 1.$$

Proof.

$$\sum_{i:\ f(i)\in\{2j-1,2j\}} a_i > W \quad \text{donne} \quad \sum_i a_i > \left\lceil \frac{\mathbf{SOL}}{2} \right\rceil W$$

et la borne inférieure $\left\lceil \sum_{i=1^n} \frac{a_i}{W} \right\rceil \leq \mathbf{OPT}$ donne le résultat.

NEXT-FIT

Je prends les objets les uns après les autres. Dès que l'objet i ne peut pas entrer dans le conteneur courante, je passe à un nouveau conteneur.

Théorème

NEXT-FIT fournit une solution $\mathbf{SOL}$ *telle que*

$$\mathbf{SOL} \leq 2\mathbf{OPT} - 1.$$

Exercice : petits objets

Montrer que si l'instance I satisfait $\frac{a_i}{W} \leq \gamma$ pour tout i, alors

$$\mathbf{NF}(I) \leq \left\lceil \frac{\sum_i a_i}{W(1-\gamma)} \right\rceil \leq \left\lceil \frac{\mathbf{OPT}(I)}{1-\gamma} \right\rceil$$

FIRST-FIT

Je prends les objets les uns après les autres. Je mets l'objet $\boldsymbol{i}$ dans le conteneur de plus petit rang où il peut entrer.

Théorème

FIRST-FIT fournit une solution $\boldsymbol{SOL}$ *telle que*

$$\boldsymbol{SOL} \leq \left\lceil \frac{\mathbf{17}}{\mathbf{10}} \boldsymbol{OPT} \right\rceil .$$

FIRST-FIT DECREASING

Je trie d'abord les objets par a_i décroissant. Puis j'applique FIRST-FIT.

Théorème

FIRST-FIT DECREASING fournit une solution **SOL** *telle que*

$$SOL \leq \frac{3}{2} OPT.$$

Quel avantage pratique ont NEXT-FIT et FIRST-FIT sur FIRST-FIT DECREASING ?

Quel avantage pratique ont NEXT-FIT et FIRST-FIT sur FIRST-FIT DECREASING ?

L'avantage de NEXT-FIT et de FIRST-FIT sur FIRST-FIT DECREASING est qu'ils peuvent fonctionner on-line.

Exemple de Recherche locale

On résout le problème de manière indirecte : on cherche à savoir s'il existe une solution en $\boldsymbol{k}$ conteneurs.

- Choisir $\boldsymbol{W'} \geq \boldsymbol{W}$, une capacité commune, pour laquelle il existe une solution réalisable.
- Appliquer une recherche locale pour minimiser $\boldsymbol{W'}$: passage d'une solution à une solution voisine en changeant l'affectation d'un objet.
- Si à la fin $\boldsymbol{W'} \leq \boldsymbol{W}$: solution réalisable au problème de départ en $\boldsymbol{k}$ conteneurs.

On répète l'algorithme en faisant varier le nombre $\boldsymbol{k}$.

Problème ouvert

Il existe encore de nombreuses instances non résolues : par exemple :

$\boldsymbol{W = 150}$
$\boldsymbol{n = 120}$
$\boldsymbol{a_i} =$ 100 22 25 51 95 58 97 30 79 23 53 80 20 65 64 21 26 100 81 98 70 85 92 97 86 71 91 29 63 34 67 23 33 89 94 47 100 37 40 58 73 39 49 79 54 57 98 69 67 49 38 34 96 27 92 82 69 45 69 20 75 97 51 70 29 91 98 77 48 45 43 61 36 82 89 94 26 35 58 58 57 46 44 91 49 52 65 42 33 60 37 57 91 52 95 84 72 75 89 81 67 74 87 60 32 76 85 59 62 39 64 52 88 45 29 88 85 54 40 57

Meilleure solution connue : $\boldsymbol{k = 51}$

Sommaire de la partie

1. Sac à dos

2. Bin Packing

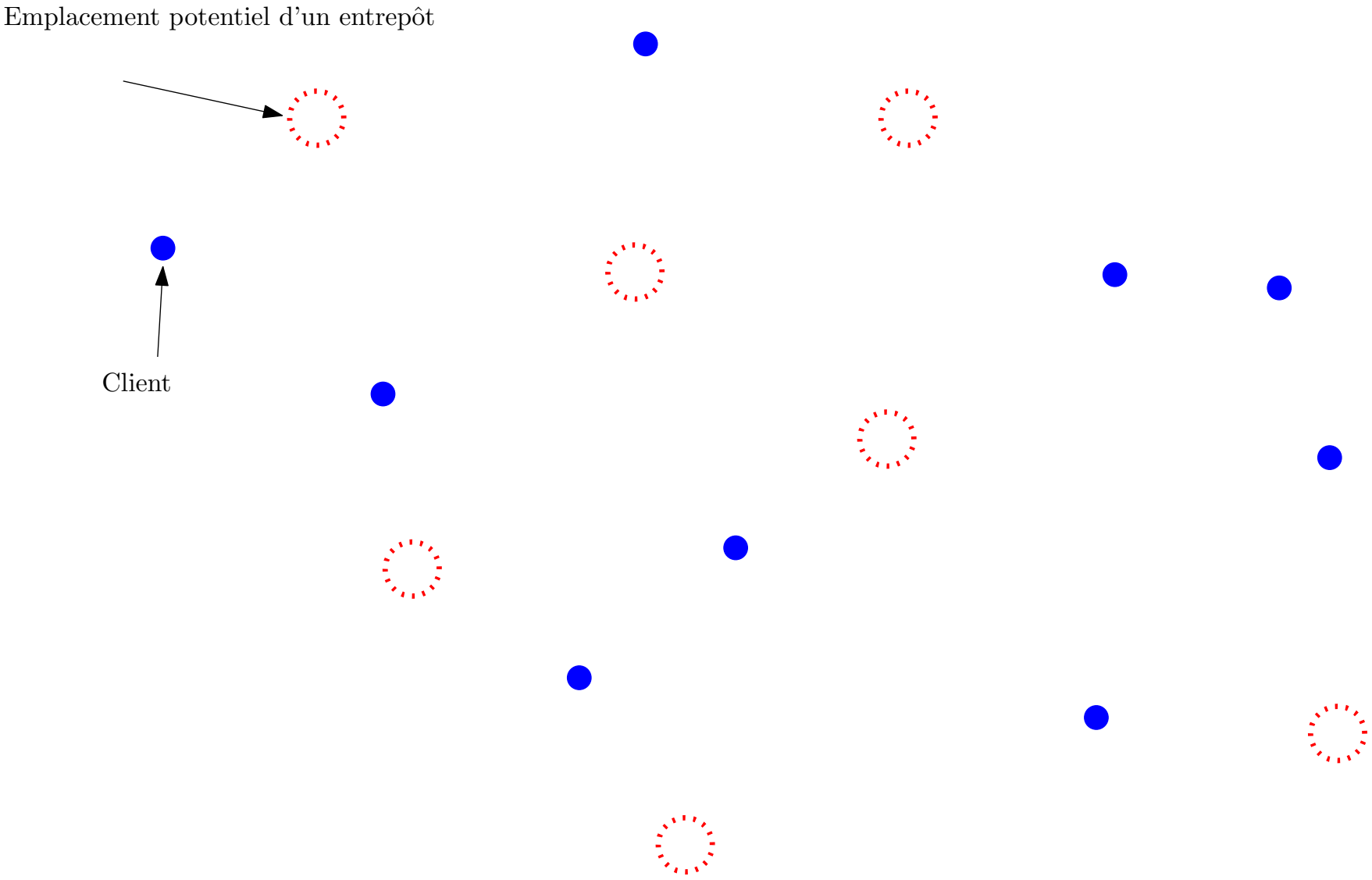
3. Positionnement d'entrepôts

4. Exercices

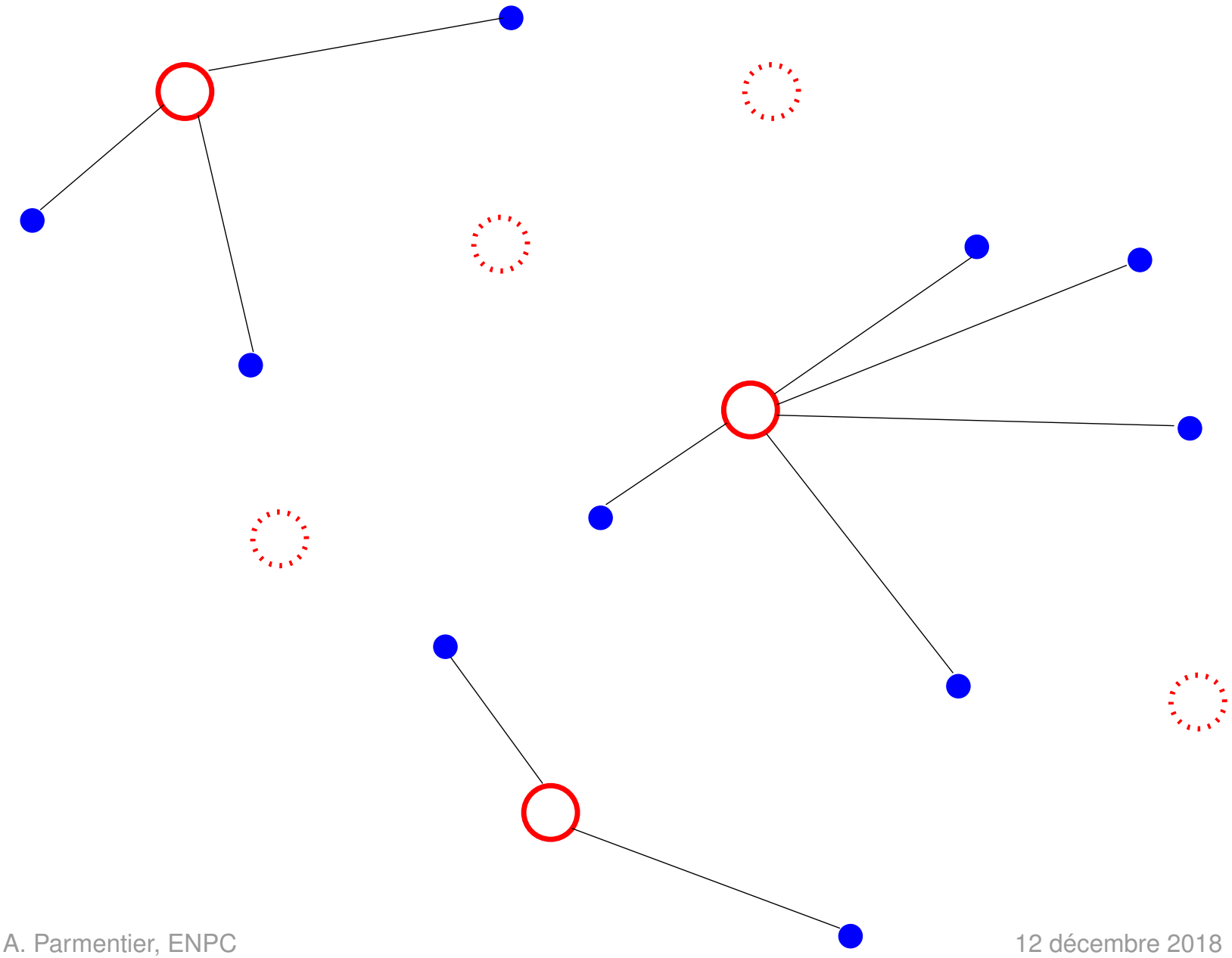
Présentation du problème

De nombreuses décisions économiques mettent en jeu la sélection ou le positionnement de dépôts, d'usines, de relais, d'hôpitaux, etc. afin de répondre de manière optimale à la demande.

Illustration



Illustration



Formalisation du problème

Problème du positionnement d'entrepôts.

Données. Clients $\mathcal{D}$, entrepôts potentiels $\mathcal{F}$, coût fixe $f_i \in \mathbb{R}_+$ d'ouverture pour entrepôt $i \in \mathcal{F}$, coût de service $c_{ij} \in \mathbb{R}_+$ pour $i \in \mathcal{F}$ et $j \in \mathcal{D}$.

Demande. Trouver $X \subseteq \mathcal{F}$ et affectation $\sigma : \mathcal{D} \to X$ tels que

$$\sum_{i \in X} f_i + \sum_{j \in \mathcal{D}} c_{\sigma(j)j}$$

soit minimale.

Complexité

Proposition

Le problème du positionnement d'entrepôts est **NP***-difficile.*

Formulation sous forme PLNE

Donner un PLNE qui modélise le problème de positionnement d'entrepôts.

Formulation sous forme PLNE

Donner un PLNE qui modélise le problème de positionnement d'entrepôts.

$$
\begin{array}{lll}
\text{Min} & \displaystyle\sum_{i\in\mathcal{F}} f_i y_i + \sum_{i\in\mathcal{F}}\sum_{j\in\mathcal{D}} c_{ij}x_{ij} & \\
\text{s.c.} & x_{ij} \leq y_i & i \in \mathcal{F},\ j \in \mathcal{D} \\
 & \displaystyle\sum_{i\in\mathcal{F}} x_{ij} = 1 & j \in \mathcal{D} \\
 & x_{ij} \in \{0,1\} & i \in \mathcal{F},\ j \in \mathcal{D} \\
 & y_i \in \{0,1\} & i \in \mathcal{F}.
\end{array}
$$

Recherche locale

1. Comment encoder une solution ?
2. Quel voisinage proposez vous ?

Recherche locale

1. Comment encoder une solution ?
2. Quel voisinage proposez vous ?

1. Le voisinage naturel est basé sur la remarque suivante : *L'espace des solutions peut être identifié à l'ensemble des parties* $\boldsymbol{X}$ *de* $\mathcal{F}$.

Recherche locale

1. Comment encoder une solution ?
2. Quel voisinage proposez vous ?

1. Le voisinage naturel est basé sur la remarque suivante : *L'espace des solutions peut être identifié à l'ensemble des parties $\boldsymbol{X}$ de $\mathcal{F}$.*
2. Modifications locales efficaces : passer de $\boldsymbol{X}$ à $\boldsymbol{X}'$ avec
   - 2.1 $\boldsymbol{X}' := \boldsymbol{X} \setminus \{\boldsymbol{x}\}$ pour un $\boldsymbol{x} \in \boldsymbol{X}$, (*drop*)
   - 2.2 $\boldsymbol{X}' := \boldsymbol{X} \cup \{\boldsymbol{x}'\}$ pour un $\boldsymbol{x}' \in \mathcal{F} \setminus \boldsymbol{X}$ (*add*) ou
   - 2.3 $\boldsymbol{X}' := \boldsymbol{X} \setminus \{\boldsymbol{x}\} \cup \{\boldsymbol{x}'\}$ (*swap*).

Exercice

On reprend le problème de positionnement d'entrepôts vu en cours, mais avec cette fois une contrainte supplémentaire : chaque entrepôt ne peut desservir qu'un nombre limité de clients. Modéliser le problème suivant comme un problème linéaire en nombres entiers.

Données : Un ensemble fini de clients $\mathcal{D}$, un ensemble fini d'entrepôts potentiels $\mathcal{F}$, un coût fixe $f_i \in \mathbb{R}_+$ d'ouverture et une capacité K_i pour chaque entrepôt $i \in \mathcal{F}$, et un coût de service $c_{ij} \in \mathbb{R}_+$ pour chaque $i \in \mathcal{F}$ et $j \in \mathcal{D}$.

Demande : Trouver un sous-ensemble $X \subseteq \mathcal{F}$ (dits *entrepôts ouverts*) et une affectation $\sigma : \mathcal{D} \to X$ des clients aux entrepôts ouverts, tel que pour tout i, l'entrepôt i ne desserve pas plus de K_i clients, de façon à ce que la quantité

$$\sum_{i \in X} f_i + \sum_{j \in \mathcal{D}} c_{\sigma(j)j}$$

soit minimale.

Correction

$$
\begin{array}{lll}
\text{Min} & \displaystyle\sum_{i\in\mathcal{F}} f_i y_i + \sum_{i\in\mathcal{F}}\sum_{j\in\mathcal{D}} c_{ij}x_{ij} & \\
\text{s.c.} & \displaystyle\sum_{j\in\mathcal{D}} x_{ij} \leq K_i y_i & i \in \mathcal{F} \\
& \displaystyle\sum_{i\in\mathcal{F}} x_{ij} = 1 & j \in \mathcal{D} \\
& x_{ij} \in \{0,1\} & i \in \mathcal{F},\ j \in \mathcal{D} \\
& y_i \in \{0,1\} & i \in \mathcal{F}.
\end{array}
$$

Sommaire de la partie

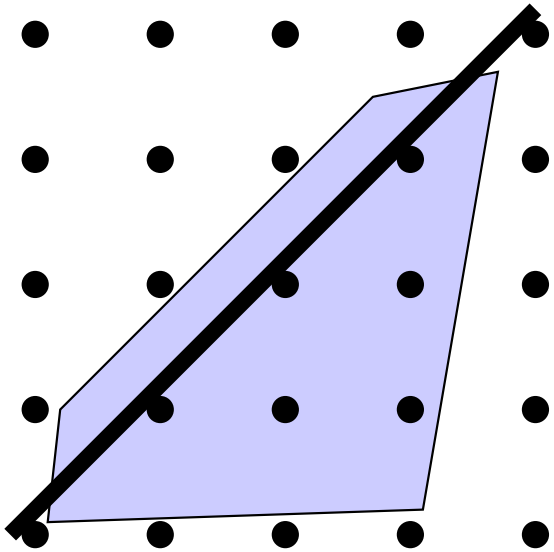
1. Sac à dos

2. Bin Packing

3. Positionnement d'entrepôts

4. Exercices
4.1 Inégalités valides
4.2 Relaxation Lagrangienne

Inégalité valide : rappel



Une inégalité valide n'a d'intérêt que si elle réduit la taille du polytope !

Inégalité valide pour le sac à dos

On considère un problème de sac-à-dos avec $a_1, \ldots, a_n, b$ des réels positifs. On note S l'ensemble des solutions réalisables

$$S := \left\{ x \in \{0,1\}^n : \sum_{i=1}^{n} a_i x_i \leq b \right\}.$$

On appelle *ensemble dépendant* tout ensemble $C \subseteq \{1, \ldots, n\}$ tel que $\sum_{i \in C} a_i > b$.

1. Montrer que tout $x \in S$ satisfait l'inégalité

$$\sum_{i \in E(C)} x_i \leq |C| - 1,$$

   où $E(C) = C \cup \{j : a_j \geq \max_{i \in C} a_i\}$.
2. Est-ce le cas pour des solutions du relâché continu.
3. Supposons que le résolve le relâché continu : comment identifier l'ensemble C pour lequel la contrainte est la plus violée ? A quoi cela sert-il ?

Correction

1. $\boldsymbol{C}$ n'est pas faisable, donc au plus $|\boldsymbol{C} - \mathbf{1}|$ éléments sont sélectionnés à l'intérieur.
2. Par exemple, si l'on deux items de poids $\boldsymbol{a_i} = \mathbf{2}$, une capacité $\boldsymbol{W} = \mathbf{3}$, et $\boldsymbol{C} = \{\mathbf{1}, \mathbf{2}\}$, alors la solution du relâché linéaire $\boldsymbol{x_1} = \mathbf{1}$ et $\boldsymbol{x_2} = \mathbf{1}/\mathbf{2}$ viole l'inégalité valide.
   De manière générale, il suffit par exemple de prendre la solution optimale du relâché continu pour avoir un contre exemple. (presque sûrement par rapport à la mesure uniforme sur les instances).
3. Par exemple par programmation dynamique (ou par un problème de plus court chemin sous-contrainte.)

Inégalité valide pour le bin-packing

On rappelle la modélisation PLNE du problème du bin-packing. On suppose que l'on dispose d'un stock de K boîtes de taille 1.

$$\begin{array}{lll} \min & \sum_{j=1}^{K} z_j & \\ \text{s.c.} & \sum_{j=1}^{K} y_{ji} = 1 & \text{pour } i = 1, \dots, n \\ & \sum_{i=1}^{n} a_i y_{ji} \leq z_j & \text{pour } j = 1, \dots, K \\ & y_{ji}, z_i \in \{0, 1\} & \text{pour } i = 1, \dots, n; \\ & & j = 1, \dots, K \end{array}$$

où $z_j = 1$ si la boîte j est utilisée et $y_{ji} = 1$ si l'objet i est mis dans la boîte j.

1. Montrer qu'en ajoutant les contraintes $z_j \geq z_{j+1}$, pour $j = 1, \dots, K-1$, on continue à modéliser le problème du bin-packing.
2. Même question avec la contrainte $\sum_{j=1}^{K} z_j \geq \lceil \sum_{i=1}^{n} a_i \rceil$.
3. Expliquer l'intérêt de ce type d'inégalité dans une résolution d'un problème de bin-packing dans une approche branch-and-bound.

Correction

1. Cela oblige juste à utiliser les premiers conteneurs, mais n'impacte pas la valeur de la solution optimale.
2. Borne inférieure que l'on a donné.
3. Briser la symétrie pour la première (très important) : divise par $\binom{K}{OPT}$ la taille de l'ensemble des solutions optimales. Restreindre la taille du polytope pour la seconde.

Relaxation Lagrangienne pour le bin packing

On rappelle le PLNE pour le bin-packing

$$
\begin{array}{lll}
\min & \displaystyle\sum_{j=1}^{k} z_j & \\
\text{s.c.} & \displaystyle\sum_{j=1}^{k} y_{ji} = 1 & i = 1, \ldots, n \\
& \displaystyle\sum_{i=1}^{n} a_i y_{ji} \leq W z_j & j = 1, \ldots, k \\
& y_{ji}, z_j \in \{0, 1\} & i = 1, \ldots, n;\ j = 1, \ldots, k
\end{array}
$$

1. Procéder à la relaxation Lagrangienne de $\sum_{j=1}^{k} y_{ji} = 1$
2. Montrer que la borne $z_{RL}(\lambda)$ se calcule facilement (quel problème doit on résoudre à chaque fois)
3. Montrer que la borne donnée par la relaxation lagrangienne est meilleure que la borne donnée par la relaxation linéaire

Relaxation Lagrangienne pour le bin packing

La correction est dans le nouveau poly, Section 11.3.1