

Sélection en temps linéaire - IGI3006 - Michel Couprie

A. Motivation

Le prince-émir du Kotar (pays abritant de gigantesques gisements de gaz naturel et de pétrole) offre une somme fabuleuse à l'ingénieur qui répondra à la question suivante. On va construire dans ce pays un réseau de gazoducs reliant les différents sites d'extraction de gaz, de manière à ce que chaque site soit relié par un gazoduc nord-sud à un unique gazoduc est-ouest (voir la figure 1).

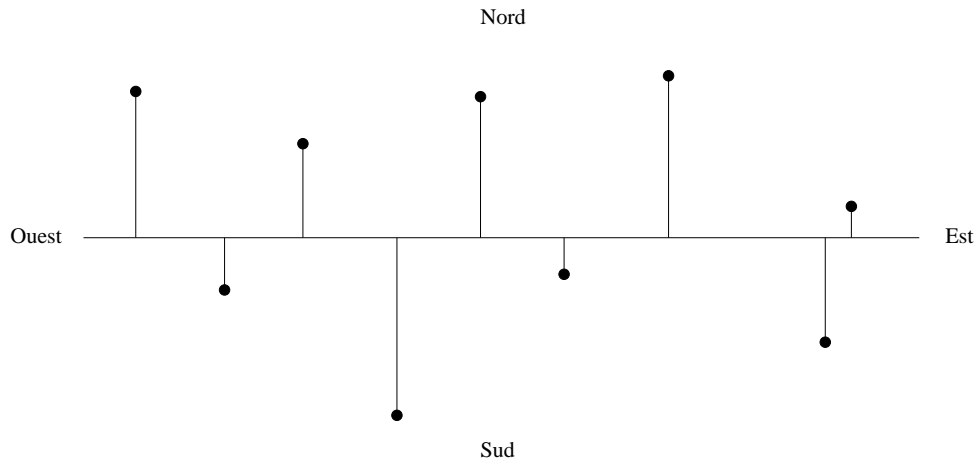


FIG. 1 – Les gazoducs du Kotar

La question est de savoir à quelle latitude placer le gazoduc est-ouest pour que le coût de construction soit minimal. Autrement dit, la longueur totale des canalisations à construire doit être minimale. On suppose pour simplifier que les latitudes des sites sont toutes différentes.

Question

Montrez par un contre-exemple simple que, dans le cas général, c'est-à-dire quelle que soit la configuration des sites, la latitude du site le plus au nord (ou le plus au sud) n'est pas toujours optimale.

Question

Montrez par un contre-exemple que la latitude moyenne des sites n'est pas toujours la solution optimale dans le cas général.

Question

Quelle est la solution optimale dans le cas général ? Justifiez votre réponse.

B. Les problèmes de la médiane et de la sélection

On dispose de n valeurs numériques, entiers ou réels, notées V_i ($i \in \{1, \dots, n\}$). Ces valeurs sont dans un ordre quelconque (*i.e.*, elles ne sont pas triées suivant les indices croissants ou décroissants). On les suppose toutes différentes entre elles.

La *médiane* de V est la $k^{\text{ième}}$ plus petite valeur parmi les V_i , avec $k = \lfloor \frac{n}{2} \rfloor$. Autrement dit, il y a exactement $k - 1$ valeurs dans V qui sont strictement inférieures à la médiane.

Un premier problème est de trouver un algorithme efficace qui retourne la médiane d'un tableau V de valeurs.

Un second problème, dit de la sélection, est de trouver un algorithme efficace qui retourne la $k^{\text{ième}}$ plus petite valeur parmi les V_i , avec k donné à l'avance (paramètre).

Il est bien évident que le problème de la médiane est un cas particulier du problème de la sélection.

C. Un algorithme naïf

Proposer deux algorithmes, l'un naïf, l'autre un peu moins naïf et plus efficace que le premier, pour résoudre le problème de la sélection. Quelle est la complexité de calcul de chacun d'eux ?

D. Médiane et sélection en temps linéaire

Montrez que, si l'on dispose d'un algorithme permettant de calculer la médiane en temps linéaire, alors on peut résoudre le problème de la sélection (pour un rang k quelconque) en temps linéaire.

Pour cela, proposez un algorithme pour la sélection basé sur l'approche diviser pour régner, qui utilise l'algorithme linéaire de calcul de la médiane, et analysez sa complexité.

E. Sélection en temps linéaire en moyenne

Proposez un algorithme de complexité linéaire en moyenne, basé sur l'approche diviser pour régner, pour résoudre le problème de la sélection. Analysez sa complexité.

F. Sélection en $O(n)$ dans le pire des cas

Proposer un algorithme pour résoudre le problème de la sélection, basé sur le principe "diviser pour régner", et dont la complexité de calcul est en $O(n)$.