

## 1 Chmura rozproszonych danych

### 1.1 Model

Dane są następujące parametry:

- $T_j$  - wektor zawierający czasy potrzebne na przeszukanie  $j$ -tego serwera.
- $q_{ij}$  - macierz zawierająca informację o tym, które cechy  $i$  zawiera  $j$ -ty serwer (1 - obecność informacji, 0 - brak informacji).
- $k$  - ilość serwerów (wnioskowana na podstawie wektora  $T_j$ )
- $n$  - ilość cech (wnioskowana na podstawie wektora  $q_{ij}$ )

#### 1.1.1 Zmienne decyzyjne

Zmienną decyzyjną jest wektor  $x$  o długości  $k$  odpowiadającej liczności serwerów. Wektor decyduje czy w ostatecznym przeszukiwaniu uwzględniany jest  $j$ -ty serwer ( $x_j = 1$ ) czy nie ( $x_j = 0$ ).

#### 1.1.2 Ograniczenia

Przynajmniej jeden wybrany serwer  $j$  zawiera dostęp do cechy  $i$ -tej

$$\forall_{i \in [n]} \left( \sum_{j=1}^k x_j \cdot q_{ij} \geq 1 \right)$$

#### 1.1.3 Funkcja kosztu

Dążymy do minimalizowania czasów dostępu do wszystkich serwerów tak aby odczytać wszystkie cechy. Koszt opisuje następująca funkcja

$$\min \sum_{j=1}^k T_j \cdot x_j$$

## 1.2 Przykładowe dane

Dla następujących danych:

- $T = [1, 2, 5, 5]$

$$\bullet q = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix}$$

solver GLPK znalazł rozwiązanie

$$x = [1, 0, 0, 1]$$