

Lösungen zum Index-Aufgabenblatt 7, WS 2013/2014

Aufgabe 7.1: Exact Match Query bei MOLHPE

Vorgehensweise während 1. Expansion:

- Berechne den Pseudoschlüssel $\Psi_j(x_j) = (b_1, \dots, b_w)$ mit $w \geq L_j + 1$ für jede Dimension j .
- Für jedes j , $1 \leq j \leq k$, berechne die vorläufige Komponentenadresse i_j :

$$i_j = \sum_{r=1}^{L_j} b_r \cdot 2^{r-1}$$

- Mit Hilfe der Expansionszeiger ep_1, \dots, ep_k kann nun entschieden werden, ob die Seite mit Adresse $G(i_1, \dots, i_k)$ schon expandiert wurde.

Fall 1: die Seite mit Adresse $G(i_1, \dots, i_k)$ wurde noch nicht expandiert

- greife auf diese Seite zu.

Fall 2: die Seite mit Adresse $G(i_1, \dots, i_k)$ wurde bereits expandiert

- i_{ex} muß neu bestimmt werden:

- Betrachte die binäre Darstellung von $i_{ex} = \sum_{r=1}^{L_{ex}-1} b_r \cdot 2^{r-1}$.

- Berechne den Anfangspunkt a des Seitenpaares, in dem i_{ex} liegt:

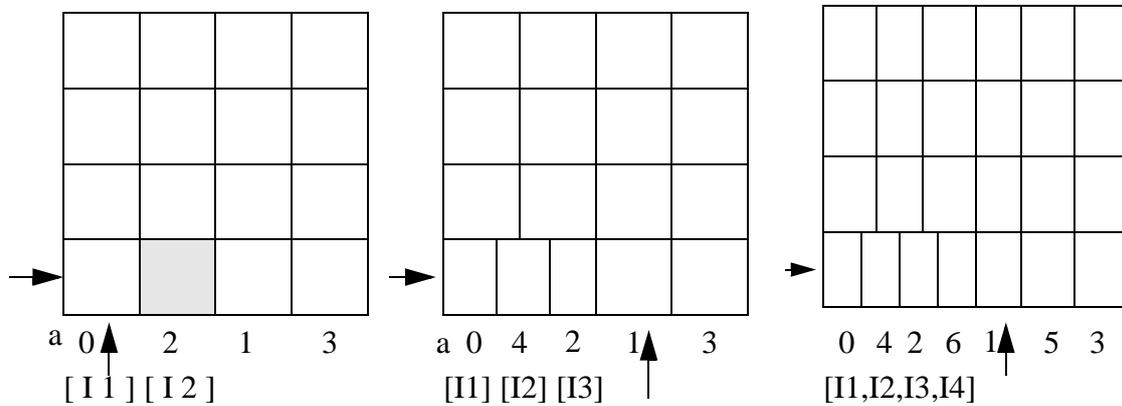
$$a = \frac{\sum_{i=1}^{L_{ex}-1} b_{L_{ex}-i} \cdot 2^{i-1}}{2^{L_{ex}-1}}$$

- Für dieses a berechne nun

$$i_{ex} = \begin{cases} i_{ex}, & \text{falls } x_{ex} \in I_a^1 \\ i_{ex} + 2^{L_{ex}}, & \text{falls } x_{ex} \in I_a^2 \\ i_{ex} + 2^{L_{ex}-1}, & \text{falls } x_{ex} \in I_a^3 \end{cases}$$

mit
$$I_a^j = \left[a + \frac{j-1}{3 \cdot 2^{L_{ex}-1}}, a + \frac{j}{3 \cdot 2^{L_{ex}-1}} \right)$$

- Greife nun zur Seite mit Adresse $G(i_1, \dots, i_{ex}, \dots, i_k)$ zu.



Vorgehensweise während 2. Expansion wie oben, außer:

Neuberechnung von i_{ex} :

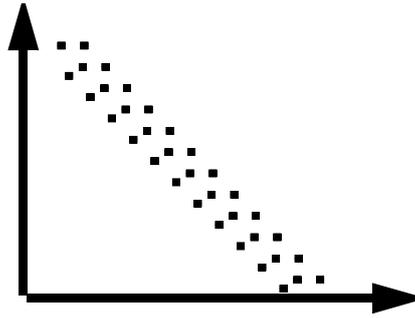
$$\begin{aligned}
 & i_{ex} && , \text{ falls } x_{ex} \in I_a^1 \\
 & i_{ex} + 2^{L_{ex}} && , \text{ falls } x_{ex} \in I_a^2 \\
 & i_{ex} + 2^{L_{ex}-1} && , \text{ falls } x_{ex} \in I_a^3 \\
 & i_{ex} + 2^{L_{ex}} + 2^{L_{ex}-1} && , \text{ falls } x_{ex} \in I_a^4
 \end{aligned}$$

mit
$$I_a^j = \left[a + \frac{j-1}{4 \cdot 2^{L_{ex}-1}}, a + \frac{j}{4 \cdot 2^{L_{ex}-1}} \right)$$

Aufgabe 8.2: (MOLHPE mit Quantilverfahren)

(+) Quantilverfahren ermöglicht gute Anpassung an Ungleichverteilungen in "einzelnen Dimensionen" (z.B. in den einzelnen Dimensionen normalverteilt)

(-) Abhängigkeiten der Attribute in den einzelnen Dimensionen (d.h. Korrelation) können durch das Quantilverfahren nicht nivelliert werden.



die Projektion der Daten auf einzelne Dimensionen schon gleichverteilt. Trotzdem insgesamt keine Gleichverteilung möglich

-> keine Möglichkeit, die Regionen der Datenseiten an die Verteilung der korrelierten Daten anzupassen (leerer Raum wird partitioniert)