

Prof. Dr. Hans-Peter Kriegel
Dr. Matthias Renz

Übungen zur Vorlesung
Index- und Speicherungsstrukturen für Datenbanksysteme
Wintersemester 2007/08

Blatt 12

Besprechung der Aufgaben: Mittwoch, 23.01.2008

Aufgabe 11.1: Eine wichtige Familie von Metriken wird von den sogenannten L_p -Distanzen gebildet. Für $x, y \in \mathcal{R}^n$ ist eine L_p -Distanz definiert wie folgt:

$$L_p(x, y) = \left(\sum_{i=1}^n |x_i - y_i|^p \right)^{\frac{1}{p}}$$

In der Vorlesung wurden Ihnen die Euklidische Distanz (L_2), die Manhattan-Distanz (L_1) und die Maximums-Distanz (L_∞) vorgestellt. In welchem Verhältnis stehen diese drei Distanzmaße zueinander?

Was bedeutet das im Kontext der Anfragebearbeitung?

Aufgabe 11.2: Man gebe die Berechnung der MINDIST-Funktion zwischen einem Anfragepunkt und einer rechteckigen Seitenregion (achsenparallel) in einem d -dimensionalen Datenraum für folgende Metriken an:

- (a) Manhattan-Distanz (L_1)
- (b) Maximums-Distanz (L_∞)
- (c) gewichtete Euklidische Distanz.

Ist dieselbe Vorgehensweise einer dimensionsweisen Aufsummierung auch für positiv definite quadratische Formen möglich?

Aufgabe 11.3: In der Grafik (siehe unten) sind für die Objekte einer Datenbank DB die jeweiligen lower/upper-bounding Filter-Distanzen (*lower*, *upper*) und exakte Distanzen (*exact*) bzgl. eines Anfrageobjektes Q angegeben. Die Objekte sind entsprechend Ihrer lower-bounding Distanz zu Q aufsteigend sortiert. Die jeweiligen Objekt-Ids sind auf der X-Achse und die zugehörigen Distanzwerte auf der Y-Achse angegeben.

Man ermittle grafisch für die folgenden Fälle, welche Objekte einem Verfeinerungsschritt unterworfen werden müssen und welche Objekte zum Resultat gehören bei der mehrstufigen Bereichsanfrage (RQ-MultiStep(DB, Q, ϵ)) für $\epsilon = 0.3$:

- Fall 1: Es ist nur die lower bounding Filterdistanz *lower* und die exakte Distanz *exact* verfügbar.
- Fall 2: Es sind alle Filterdistanzen *lower* und *upper* und die exakte Distanz *exact* verfügbar.

Übungen zur Vorlesung
Index- und Speicherungsstrukturen für Datenbanksysteme
Wintersemester 2007/08

