

Spatial, Temporal, and Multimedia Databases
 SoSe 2008

Übungsblatt 4: Algorithmen zur Ähnlichkeitssuche, optimale kNN-Anfrage, Ranking-Anfrage

Besprechung: 28.05.2008, 29.05.2008

Aufgabe 4-1 *Mehrstufige Anfragebearbeitung II*

Gegeben sei eine Datenbank mit 20 Objekten und ein Anfrageobjekt q . Auf diese Datenbank soll eine mehrstufige k -NN-Anfrage mit $k=10$ ausgeführt werden. Man ermittle, welche Objekte einem Verfeinerungsschritt unterworfen werden wenn man davon ausgeht, dass im Filterschritt sowohl die lower-bound Filterdistanz als auch die upper-bound Filterdistanz zur Verfügung stehen. Hierfür ist das Verfahren *k-NN-MultiStep-Optimal* (Skript Folie 87) anzuwenden.

Die jeweiligen exakten Distanzwerte und Filter-Distanzwerte zum Anfrageobjekt q sind aus der Grafik (siehe Abbildung 1) zu lesen.

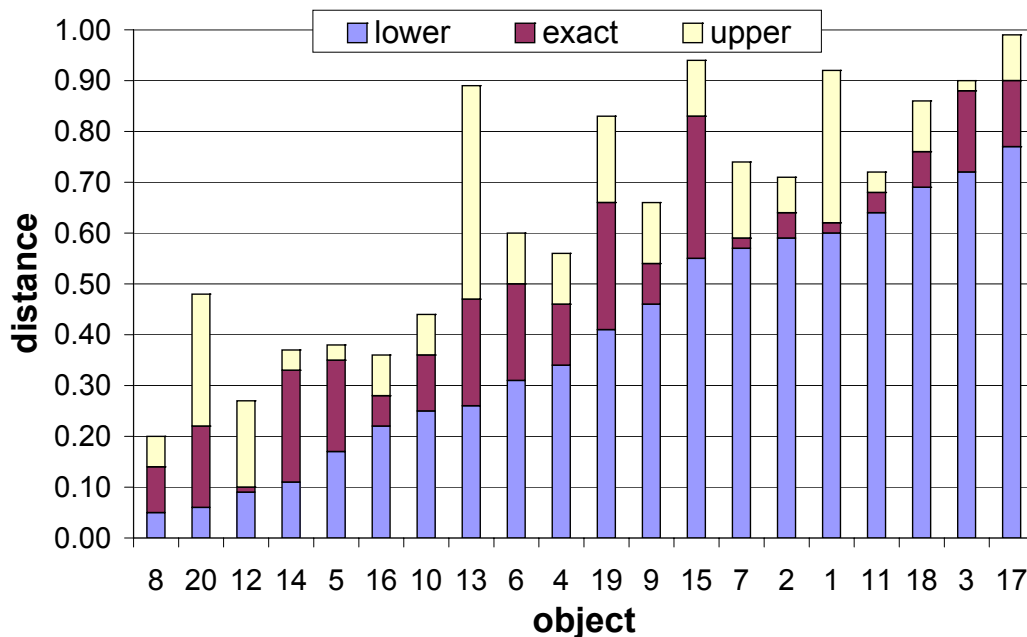


Abbildung 1: Distanzliste als Grafik (zu Aufgabe 5-1).

Aufgabe 4-2 *Basisalgorithmen zu Ranking Queries*

Überlegen Sie sich drei Basis-Algorithmen für die Ranking-Query ohne Indexunterstützung. Berücksichtigen Sie hierbei, daß nicht die *gesamte* Datenbank in den Arbeitsspeicher paßt. Die Anzahl der `get-next` Aufrufe umfaßt jedoch potentiell die gesamte Datenbankgröße.

Diskutieren Sie die alternativen Lösungen unter Berücksichtigung folgender Benutzerszenarien:

- (a) Ein Benutzer ruft typischerweise zwischen 10 und 100 Anfrageergebnisse ab.
- (b) Ein Programm, das die Anfrageergebnisse weiterverarbeitet, ruft zwischen 10% und 100% aller Datenbankobjekte ab.