

Prof. Dr. Hans-Peter Kriegel
Dr. Martin Ester
Marco Pötke

Übungen zur Vorlesung
Geo-Informationssysteme
Sommersemester 1999

Blatt 10

Aufgabe 10.1: Partitioning Around Medoids

Beim Clustering durch PAM (Partitioning Around Medoids, siehe Skript S. 249) tritt das Problem auf, die potentielle Kostenänderung einer Partitionierung zu berechnen, die das Ersetzen eines Medoids durch einen Nichtmedoid bewirken würde.

- a) Für welche Objekte ändert sich ihr Beitrag zu den Kosten einer Partitionierung, wenn ein Medoid durch einen Nichtmedoid ersetzt wird? Welche Fälle können hierbei unterschieden werden?
- b) Um wieviel ändert sich der Kostenbeitrag jeweils? Wieviele Objekte müssen für die Berechnung der Kostenänderung betrachtet werden?

Aufgabe 10.2: DBSCAN (*schriftlich bearbeiten*)

- a) Beweisen Sie folgende Aussagen (vgl. Skript S. 259):
 - 1.) Sei p ein Kernpunkt einer Datenbank D . Dann ist die Menge aller Punkte aus D , die von p dichte-erreichbar sind, ein Cluster.
 - 2.) Sei C ein Cluster in einer Datenbank D und p ein beliebiger Kernpunkt dieses Clusters. Dann ist C gleich der Menge aller Punkte aus D , die von p dichte-erreichbar sind.
- b) Für welche Punkte einer Datenbank ist die von DBSCAN ermittelte Clusterzugehörigkeit eindeutig, für welche eventuell nicht?
- c) Wie könnte man das Resultat von DBSCAN ganz eindeutig machen?

Besprechung der Aufgaben: Mittwoch, 21.07.1999