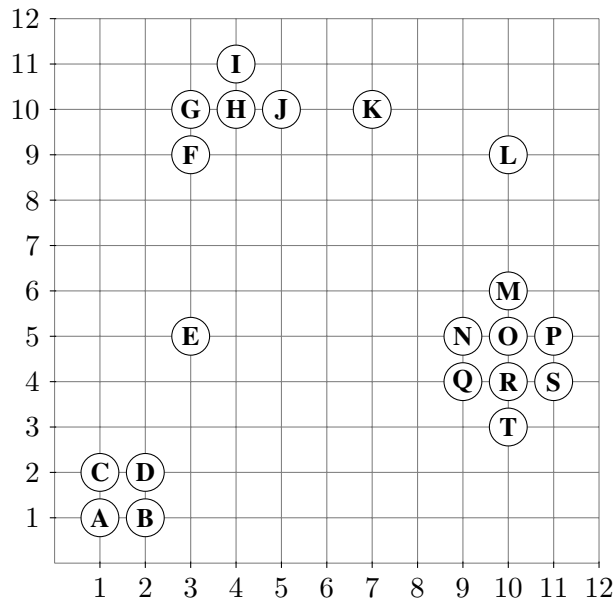


Knowledge Discovery in Databases
 SS 2013

Übungsblatt 4: Clusteranalyse – Single-Link und OPTICS

Aufgabe 4-1 Single-Link Hierarchical Clustering

Gegeben sei der folgende Datensatz:



Als Distanzfunktion zwischen den Punkten dient Ihnen jeweils wieder die Manhattan-Distanz (L_1 -Norm):

$$L_1(x, y) = |x_1 - y_1| + |x_2 - y_2|$$

Berechnen Sie zwei Dendrogramme für diesen Datensatz. Als Distanzfunktion zwischen Mengen von Objekten verwenden Sie

- (a) den Single-Link Ansatz,
- (b) den Average-Link Ansatz.

Tipp: Innere Knoten müssen nicht binär sein, d.h. sie können mehr als zwei Söhne haben.