

Zusammenfassung

In der Praxis gibt es viele Probleme, bei denen verschiedene Objekte, die jeweils in einer gewissen Relation (z.B. Ähnlichkeit oder Entfernung) zueinander stehen, in unterschiedliche Gruppen möglichst sinnvoll aufgeteilt werden sollen. Zum Beispiel können die Objekte Messdaten sein und das Ziel ist es die Daten so in Gruppen aufzuteilen, dass die Daten möglichst nah beieinander liegen bzw. ähnlich sind. Dieses kann man mathematisch modellieren, indem man einen, den Objekten und den Relationen entsprechenden Graphen betrachtet und dessen Knotenmenge in disjunkte Mengen zerlegt, die wir Cluster nennen. Eine Aufteilung der Knoten in Cluster ist dann ein Clustering. Das Clusteringproblem gibt es in viele verschiedenen Varianten. Je nach Problem haben wir eine bestimmte Zielfunktion und die Anzahl der Cluster ist vorgegeben oder nicht. Beim Complete-linkage Problem, das ich betrachte, geht es darum einen vollständigen Graphen in eine vorgegebene Anzahl von Clustern zu unterteilen, so dass die größte Entfernung einer Kante, die innerhalb der Cluster verläuft, kleinstmöglich ist.

Da die exakten Entfernungen in der Praxis nicht immer exakt ermittelt werden können, gehe ich davon aus, dass die Kantenentfernungen innerhalb einer Unsicherheitsmenge \mathcal{U} liegen. In meinem Vortrag werde ich vorstellen, wie sich die unsicheren Entfernungen allgemein auf das Complete-linkage Problem auswirken. Dannach werde ich einige bekannte Robustheitskonzepte untersuchen. Dabei gehe ich erstmal davon aus, dass wir eine allgemeine Unsicherheitsmenge \mathcal{U} haben. Zum Schluss betrachte ich dann noch intervallbasierte Unsicherheitsmengen.