

---

Irmela Herzog und Frank Siegmund

**Clusteranalyse räumlicher Daten  
mit Hilfe der Gemeinsamen Nachbarschafts-Gruppierung**

Eine häufige Aufgabe der Archäologen ist es, Objekte in Gruppen einzuteilen, z.B. um eine Typologie zu erstellen oder um geschlossene Funde verschiedenen Zeitstufen zuzuordnen. Um bei der Einteilung einer größeren Menge von Objekten Zeit zu sparen und mehr Objektivität zu erlangen, kann man mathematische Verfahren zur Gruppierung anwenden.

Zunächst muß man definieren, welchen Ähnlichkeits- bzw. Distanzwert man zwei zu klassifizierenden Objekten zuweisen möchte. Für unterschiedliche Datenarten sind jeweils andere Ähnlichkeitsmaße sinnvoll. So empfehlen wir für Meßdaten den euklidischen oder den mittleren Abstand (nach Standardisierung der Werte), die gleichen Distanzmaße und der Chi<sup>2</sup>-Abstand sind für gezählte Häufigkeiten geeignet. Für Anwesenheits-/Abwesenheitsdaten eignen sich besonders die Ähnlichkeitsmaße von Jaccard und Dice, bei kategorischen Daten ist auch das Maß von Russel und Rao verwendbar.

Die bisher erwähnten Distanzmaße berücksichtigen die geographische Lage bei räumlichen Daten nicht. Man erwartet jedoch oft, daß die errechneten Gruppen räumlich zusammenhängen. Deshalb haben wir ein Distanzmaß konstruiert, das die geographische Lage in einem vorher festzulegenden Prozentwert berücksichtigt. Wenn man z.B. 10 % der geographischen Information mit einbeziehen möchte, erhält man:

$$d = 0,1 \cdot d_{\text{Geographie}} + 0,9 \cdot d_{\text{Inhalt}}$$

wobei geographischer und Inhaltsabstand so normiert werden, daß sie zwischen 0 und 1 liegen.

Als mathematisches Gruppierungsverfahren halten wir die von R.A. Jarvis und Edward A. Patrick entwickelte Shared Near Neighbour Cluster Analysis (Gemeinsame-Nachbarschafts-Gruppierung, kurz GN-Gruppierung) für besonders geeignet. Der Benutzer wählt aus, wie viele Nachbarn eines Objektes berücksichtigt werden sollen und steuert dadurch, ob eher globale (in der Ebene kreisscheibenförmige) oder eher lokale (in der Ebene kettenförmige) Gruppierungen entstehen.

Anhand von zwei archäologischen Beispielen zeigen wir Schwächen und Stärken dieses Verfahrens. Das fränkische Gräberfeld von Köln-Müngersdorf weist 152 relativ wenig beraubte Bestattungen auf. Die GN-Gruppierung ohne Berücksichtigung der geographischen Lage reproduzierte das mit konventionellen Methoden erzielte Ergebnis am besten. Die Hoffnung, durch Einbeziehen der räumlichen Lage fundarme Gräber und solche unterschiedlichen Geschlechts besser einbinden zu können, erfüllte sich nicht, u.a. weil das Verfahren

das Wachstumsmodell des Gräberfeldes nicht berücksichtigt. Beim jungpaläolithischen Siedlungsplatz von Sprendlingen bestehen die Daten aus gezählten Häufigkeiten auf dem Niveau von Viertelquadraten. Durch Einbeziehen von 10% Lageinformation in das Abstandsmaß konnte das konventionelle Ergebnis am besten reproduziert werden, wobei für die Peripherie und das Zentrum des Siedlungsplatzes unterschiedliche Nachbarzahlen gewählt wurden, da das Zentrum mehr lokale Strukturen aufweist.

Die obigen Ergebnisse wurden mit Programmen errechnet, die demnächst im Rahmen des Bonner archäologischen Seriations- und Statistikpaketes erhältlich sein werden. Neben dem Clusterverfahren mit den oben beschriebenen Distanz- und Ähnlichkeitsmaßen sind Dienstprogramme vorhanden, die ein Denrogramm erzeugen, die Ergebnisse kartieren oder in Diagramm- und Tabellenform darstellen.

### Literatur

- G. Bosinski, E. Cziesla u.a., 1985, Sprendlingen. Ein Fundplatz des mittleren Jungpaläolithikums in Rheinhessen. Jahrb. RGZM 32, 1985, 5-91.
- J.E. Doran u. F.R. Hodson, 1975, Mathematics and Computers in Archaeology. Edinburgh University Press 1975, 23-24, 136-143, 173-186.
- R.A. Jarvis, 1978, Shared Near Neighbor Maximal Spanning Trees for Cluster Analysis. Proceedings of the Fourth International Joint Conference on Pattern Recognition, Kyoto 1978, 308-313.
- F. Siegmund, 1989, Fränkische Funde vom deutschen Niederrhein und der nördlichen Kölner Bucht. Phil. Diss. Köln, 1989.
- R.R. Sokal u. P.H.A. Sneath, 1973, Numerical Taxonomy. W.H. Freeman and Company, San Francisco, London 1973, 121-135.