

Praehistorische Zeitschrift

Herausgegeben von
Niels Bantelmann · Bernhard Hänsel
Michael Müller-Wille

66. Band · 1991 · Heft 2

Sonderdruck

Walter de Gruyter · Berlin · New York

Clusteranalyse räumlicher Daten mit Hilfe der Gemeinsamen-Nachbarschafts-Gruppierung (GN-Gruppierung)*

von Irmela Herzog, Bonn, und Frank Siegmund, Göttingen

In der Archäologie sind häufig Funde oder Befunde zu Gruppen zusammenzufassen; statistische Verfahren zur Clusteranalyse dienen dieser Zielsetzung. Voraussetzung der Clusterverfahren ist die Festlegung von Maßen für die Ähnlichkeit zwischen zwei zu klassifizierenden Objekten. Für unterschiedliche Datenarten sind jeweils andere Ähnlichkeitsmaße sinnvoll. Ein neues Ähnlichkeitsmaß berücksichtigt neben den Inhalten archäologischer Einheiten auch ihre topographische Lage. Das Gruppierungsverfahren von Jarvis und Patrick, bei dem der Bearbeiter beeinflussen kann, ob eher lokale oder eher globale Gruppen entstehen, wenden wir mit und ohne das neue Ähnlichkeitsmaß auf zwei archäologische Beispiele an. Die errechneten Gruppierungen für das Gräberfeld von Köln-Müngersdorf (470–650 n. Chr.) stehen im Ergebnis hinter konventionellen Methoden zurück; das „Wachstum“ eines Gräberfeldes läßt sich wider Erwarten nur schwer mit einem topographisch beeinflussten Ähnlichkeitsmaß beschreiben. Eine Lage-berücksichtigende Gruppierung des Fundplatzes von Spredlingen (30 000–20 000 v. Chr.) führt zu einer sinnvollen Zonengliederung. Computer-Programme zur Berechnung und Analyse von Gruppierungen stehen zur Verfügung.

En archéologie il s'avère souvent nécessaire de classer les trouvailles et les complexes, c'est ce que se proposent de faire les méthodes statistiques pour les analyses par clusters. La définition d'indices de ressemblance entre deux objets à classer est à la base de la classification par clusters. Il est conseillé pour des données de genres différents d'utiliser à chaque fois des indices de ressemblance différents. Un nouvel indice de ressemblance prendra en considération, en dehors des contenus des unités archéologiques, aussi leur situation topographique. La méthode de classification de Jarvis et Patrick, dans laquelle le chercheur peut influencer dans une certaine mesure, la formation de groupes locaux ou globaux, est appliquée ici avec et sans indice de ressemblance sur deux exemples archéologiques. La classification après calculs pour la nécropole de Cologne-Müngersdorf (470–650 ap. J. C.) donne des résultats moins satisfaisants que les méthodes traditionnelles. La «croissance» d'une nécropole, à l'encontre de nos espoirs s'avéra difficile à décrire à l'aide d'un indice de ressemblance, reposant sur des données topographiques. Par contre la classification du site paléolithique de Spredlingen (30 000–20 000 av. J. C.) fournit une série de zones d'habitat tout à fait cohérentes. Les programmes pour ordinateur, utilisés pour les calculs et les analyses sont disponibles.

In archaeology finds and sites must often be grouped. This may be done through the use of statistical techniques of cluster analysis. The prerequisite of clustering is the choice of measures of similarity between two objects to be classified. Different types of data call for different measures of similarity. A new measure of similarity takes the content and the topographic position of archaeological units into account. Jarvis and Patrick's shared near neighbour clustering which can be controlled by the user in order to generate groups with a range from the global to the local was applied to two archaeological examples with and without new measures of similarity. The clusterings arrived at by the analysis of Cologne-Müngersdorf (470–650 A. D.) are inferior to those arrived

* Die Idee, die GN-Gruppierung auf archäologische Daten anzuwenden, verdanken die Autoren I. Scollar. Er trug auch maßgeblich zur Entwicklung des topographisch beeinflussten Distanzmaßes bei. Hierfür danken wir ihm an dieser Stelle herzlich. – Eine leicht veränderte Fassung des

vorliegenden Beitrags haben wir am 18. Mai 1989 im Rahmen der Sitzung der Arbeitsgemeinschaft „Quantitative Methoden in der Archäologie“ beim Deutschen Kongreß für Archäologie in Frankfurt vorgetragen.

at by traditional means. The expansion of this cemetery can not as was hoped be adequately described by a topographically influenced measure of similarity. Meaningful groupings result from clustering the site of Sprendlingen (30 000 – 20 000 B. C.) when topographic considerations are taken into account. Computer programs which calculate and analyze clusterings are available.

Eine häufige Aufgabe der Archäologen ist es, Objekte in Gruppen einzuteilen, z. B. um eine Typologie zu erstellen oder um verschiedene geschlossene Funde gemeinsamen chronologischen Einheiten zuzuordnen. Die Einteilung einer größeren Menge von Objekten in Klassen oder Typen ist zeitaufwendig und das Ergebnis häufig umstritten¹. Um Zeit zu sparen und größere Objektivität zu erlangen, kann man mathematische Verfahren zur Gruppierung anwenden. Es gibt eine Fülle solcher Methoden, die je nach Datenart unterschiedlich gut funktionieren. Wir stellen hier die Gemeinsame-Nachbarschafts-Gruppierung vor, die sich in vielen Fällen als geeignet erwiesen hat. Außerdem entwickeln wir ein neues Distanzmaß, das neben den Eigenschaften der zu gruppierenden Objekte auch ihre topographische Lage berücksichtigt. Zwei archäologische Beispiele zeigen Stärken und Grenzen des Verfahrens.

Distanz- und Ähnlichkeitsmaße

Allen mathematischen Gruppierungsverfahren ist gemeinsam, daß man zunächst ein angemessenes Ähnlichkeitsmaß definieren muß. Die Wahl des Ähnlichkeitsmaßes beeinflusst das Ergebnis mindestens so stark wie die Wahl des Gruppierungsverfahrens. Möchte man beispielsweise eine Typologie für Schnallen erstellen, ist nicht von vornherein klar, welchen Ähnlichkeitswert man zwei Schnallen zuweisen soll. Wenn man davon ausgeht, daß sich die Typologie im wesentlichen durch die äußere Form bestimmt, ist es sinnvoll, Meßdaten wie Länge des Dorns, Länge und Breite der Schnalle usw. zu erheben; soll die Typologie vor allem die Verarbeitung oder Verzierung der Schnalle berücksichtigen, ist es besser, das Vorhandensein bzw. die Abwesenheit entsprechender Merkmale zu registrieren. Diese unterschiedlichen Arten von Information erfordern verschiedene, jeweils angemessene Ähnlichkeitsmaße. Man unterscheidet Häufigkeits- und Anwesenheits-/Abwesenheitsdaten. Eine Übersicht über die Abstandsmaße bei Häufigkeitsdaten gibt Abb. 1².

Die Objekte, die gruppiert werden sollen, werden als Einheiten bezeichnet. Während für Häufigkeitsdaten im allgemeinen Distanzmaße berechnet werden, verwendet man für Anwesenheits-/Abwesenheitsdaten traditionell Ähnlichkeitsmaße. Es besteht ein einfacher Zusammen-

hang zwischen Ähnlichkeits- und Distanzmaßen: je größer die Ähnlichkeit, desto kleiner die Distanz; je größer die Distanz, desto kleiner die Ähnlichkeit. Da sich Distanz und Ähnlichkeit umgekehrt proportional zueinander verhalten, ist es einfach, ein Ähnlichkeitsmaß in ein Distanzmaß zu transformieren und umgekehrt.

Die Ähnlichkeitsmaße für Anwesenheits-/Abwesenheitsdaten sind in Abb. 2 aufgeführt³. Kategorische Daten sind z. B. Farben oder Materialien: Jeder untersuchte Gegenstand fällt sicher in eine Klasse des Kategoriensystems; folglich werden für jeden Gegenstand gleich viele Beobachtungen erhoben. Das Inventar geschlossener Funde dagegen bildet nicht-kategorische Daten; die Komplexe enthalten in der Regel unterschiedliche Mengen typisierter Objekte. Zudem gehören die einzelnen Objekte meist verschiedenen Klassen an, für die jeweils ein eigenes Kategoriensystem verwendet wird.

In der Archäologie liegen häufig auch räumliche Informationen vor, z. B. bei Gräberfeldern oder Siedlungsplätzen. Die bisher besprochenen Distanzmaße berücksichtigen die topographische Lage nicht. Man erwartet jedoch oft, daß die Gruppen räumlich zusammenhängen. In solchen Fällen kann es sinnvoll sein, die Lageinformation in gewissem Maße zu berücksichtigen, wobei Störeffekte geglättet werden.

Soll eine räumliche Gliederung ohne Berücksichtigung der Inhaltsinformation vorgenommen werden, so ist der normale topographische Abstand das geeignete Distanzmaß. Es liegt nahe, die Distanzmaße für den Inhalt und für die Topographie zu einem neuen Maß zu kombinieren. Dabei wünscht man sich einen Regler, mit dem der Einfluß der topographischen Komponente auf das Gesamtdistanzmaß bestimmt wird. Ein Distanzmaß d , das beispielsweise aus 10% topographischem und 90% Inhaltsabstand besteht, kann man nach dieser Grundidee in folgender Weise konstruieren:

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

Der Absolutwert des normalen topographischen Abstands hängt von der Wahl der Maßeinheit ab, weshalb das neue Maß d von dieser nicht unabhängig ist. Deshalb dividieren wir den topographischen Abstand

¹ Doran u. Hodson 1975, 158.

² Sokal u. Sneath 1973; Doran u. Hodson 1975, 23 f.; 136 f.

³ Ebd. 137 – 143; Sokal u. Sneath 1973. Auf die häufig anzutreffende Normierung des Russel/Rao- und des Simple Matching-Ähnlichkeitskoeffizienten wurde hier verzichtet, da einerseits durch die Normierung die Berechnung teurer wird, andererseits sich das Gruppierungsergebnis nicht ändert.

Unterart	Vorbereitung	Abstandsmaß
Meßdaten	Standardisieren	– euklidischer Abstand – Cityblock Abstand
gezählte Häufigkeiten	Möglichkeit des Wurzelziehens	– euklidischer Abstand – Cityblock Abstand – Chi-Quadrat Abstand

Euklidischer Abstand: $\sqrt{\sum (X_i - Y_i)^2}, i = 1 \dots n$
 Cityblock Abstand:
 (auch: mittlerer Abstand, Manhattan Abstand) $\sum |X_i - Y_i|, i = 1 \dots n$
 Chi-Quadrat Abstand: normiert über Einheiten, rechenaufwendig (siehe Dixon 1983)

Abb. 1. Abstandsmaße bei Häufigkeitsdaten

Unterart	Ähnlichkeitsmaße
kategorisch nicht kategorisch	Russel/Rao, Simple Matching(?), Yule(??) Jaccard, Dice, Russel/Rao(?)

Hilfsgrößen für die Definition von Ähnlichkeitsmaßen für Anwesenheits-/Abwesenheitsdaten:

- a = Anzahl gemeinsamer Merkmale von Einheit-1 (Einheit minus 1) und Einheit-2
- b = Anzahl Merkmale, die zu Einheit-1 aber nicht zu Einheit-2 gehören
- c = Anzahl Merkmale, die zu Einheit-2 aber nicht zu Einheit-1 gehören
- d = Anzahl Merkmale, die weder zu Einheit-1 noch zu Einheit-2 gehören

Ähnlichkeitsmaße:

1. Russel/Rao: a
günstig, wenn die Einheiten ungefähr gleich viele Merkmale enthalten.
2. Jaccard: $a / (a + b + c)$
schafft Ausgleich, wenn die Einheiten unterschiedlich merkmalsreich sind.
3. Dice: $2a / (2a + b + c)$
betont Gemeinsamkeiten gegenüber Unterschieden.
4. Simple Matching: $a + d$
Achtung: gemeinsam fehlende Merkmale erhalten gleiches Gewicht wie gemeinsam vorhandene Merkmale.
5. Yule: $(ad - bc) / (ad + bc)$
Zahl der übereinstimmenden Merkmale wird abgewogen gegenüber Zahl der Unterschiede (ungünstig und rechenaufwendig!)

Abb. 2. Ähnlichkeitsmaße für Anwesenheits-/Abwesenheitsdaten

durch den maximalen topographischen Abstand (Durchmesser) unseres Datensatzes. Dann liegen alle topographischen Abstände im Bereich zwischen 0 und 1. In gleicher Weise normieren wir den Inhaltsabstand und haben nun die Gewißheit, daß beide Abstände im vorher festgelegten Prozentsatz in dem neuen Distanzmaß berücksichtigt werden.

Clusterverfahren – eine kurze Übersicht

Nachdem wir das Problem, welches Distanzmaß wir zugrunde legen, geklärt haben, wenden wir uns dem Problem des Gruppierungsverfahrens zu. Zu den bekanntesten Methoden zählen Single-Linkage-, Average-Linkage-, Complete-Linkage- und K-means-Cluste-

ring⁴. Die ersten drei Methoden sind hierarchische Clusterverfahren, bei denen anfangs jede Einheit eine „Gruppe“ bildet und diese dann sukzessive zu immer größeren Gruppen zusammengefaßt werden, bis schließlich alle Einheiten zu einem Cluster vereint sind. Der Auswerter entscheidet, auf welchem Niveau die Gruppenbildung für ihn am sinnvollsten ist. Die typische Darstellung für das Ergebnis eines hierarchischen Clusterverfahrens ist ein Dendrogramm. K-means-Clustering ist nicht hierarchisch; hierbei geht man von einer Anfangseinteilung der Einheiten aus und versucht, diese Einteilung sukzessive so zu verbessern, daß das Clusterkriterium immer besser erfüllt ist. Der Nachteil dieser Methode ist, daß das Ergebnis von der Anfangs-

⁴ Doran u. Hodson 1975, 173 – 177; Jarvis 1978.

einteilung der Einheiten abhängt und so nicht in jedem Fall ein optimales Ergebnis erzielt wird.

Um die Eigenschaften eines Clusterverfahrens aufzuzeigen, betrachtet man Bilder von Punkten in der Ebene und wertet aus, wie die Punkte bei dem Verfahren in Gruppen eingeteilt werden. In diesem Fall kann man mit dem menschlichen Auge sehr gut das Ergebnis der Gruppierung überprüfen. Bei Anwendung von Single-Linkage-Clustering auf Punkte in der Ebene bilden sich häufig kettenähnliche Cluster. Da dann immer nur eine kleine Umgebung eines Punktes mit zum Cluster des Punktes gehört, spricht man auch von lokalen Clustern. Complete-Linkage-Clusteranalyse dagegen kann sehr gut kreisscheibenförmige Cluster von Punkten in der Ebene erkennen. Hier gehört eine große Umgebung eines Punktes mit zu seiner Gruppe, man spricht deshalb allgemein von globalen Clustern. Average-Linkage- und K-means-Clusteranalyse sind auf der Achse zwischen lokalen und globalen Clusterverfahren anzusiedeln, aber näher bei den globalen als bei den lokalen Methoden.

Worin unterscheiden sich die beiden hierarchischen Clusterverfahren Single-Linkage und Complete-Linkage, so daß mit dem einen Verfahren linien- und mit dem anderen eher kreisförmige Cluster erkannt werden können? Die Antwort liegt in dem Kriterium, wie Gruppen zu einem neuen Cluster verschmolzen werden. Bei Single-Linkage werden in jedem Schritt die beiden Gruppen zusammengefaßt, bei denen die Ähnlichkeit zwischen beliebigen Einheiten aus beiden Gruppen maximal ist; bei Complete-Linkage wird jede Einheit der einen Gruppe mit jeder Einheit der anderen Gruppe verglichen und der Minimalwert dieser Ähnlichkeiten berechnet; die beiden Gruppen mit dem höchsten minimalen Ähnlichkeitswert werden verschmolzen.

Single-Linkage-Clustering ist trotz seiner bekannten Nachteile weit verbreitet, da es einfach und mit wenig Speicheraufwand zu programmieren ist und einen weiteren mathematischen Vorteil auf seiner Seite hat: Es ist unabhängig von monotonen Transformationen der Ähnlichkeitswerte, d. h. man kann den Logarithmus oder das Quadrat dieser Werte nehmen, ohne das Cluster-Ergebnis zu beeinflussen. In diesem Fall ist es sinnvoll, bei der euklidischen Distanz auf das Wurzelziehen zu verzichten und so Rechenzeit zu sparen. Auch Complete-Linkage-Clustering hat diesen mathematischen Vorteil.

Das Clusterverfahren von Jarvis und Patrick

Die Idee von R. A. Jarvis und E. A. Patrick war es, eine Clustermethode zu konzipieren, bei der durch einen Parameter gesteuert werden kann, ob man eher lokale (Single-Linkage) oder eher globale (Complete-Linkage) Gruppierungen haben möchte, bzw. einen

steuerbaren Kompromiß zwischen beiden Verfahren⁵. Der Parameter ist die Zahl der nächsten Nachbarn, die man für jede Einheit betrachtet. Bei Berücksichtigung einer geringen Anzahl von Nachbarn hat das Ergebnis Ähnlichkeit mit dem der Single-Linkage-Analyse, erhöht man die Zahl der Nachbarn, spielen globale Verteilungsaspekte eine größere Rolle. Der nächste Nachbar ist hier im Sinne der oben betrachteten Distanz- bzw. Ähnlichkeitsmaße gemeint.

Wenn man k Nachbarn berücksichtigt, dann wird zunächst für jede Einheit eine Liste der nächsten k Einheiten aufgestellt. Auf der untersten Ebene werden alle Einheiten zu einem Cluster zusammengefaßt, die k gemeinsame Nachbarn haben. Auf der nächsthöheren Ebene werden die Cluster aus Einheiten gebildet, bei denen mindestens $k-1$ (k minus 1) Nachbarn übereinstimmen, und so weiter. Hat eine Einheit mit einer anderen m Nachbarn gemeinsam, so werden die beiden Einheiten nur dann auf Stufe m in einem Cluster zusammengefaßt, wenn sie gegenseitig in den jeweiligen Nachbarschaftslisten enthalten sind. So wird verhindert, daß „Außenseiter“ in ein Cluster integriert werden (siehe Abb. 3).

Dieses Verfahren wurde von Jarvis und Patrick als „Shared Near Neighbour Analysis“ bezeichnet, wir übersetzen dies mit Gemeinsame-Nachbarschafts-Gruppierung, kurz *GN-Gruppierung*. Wenn die räumliche Lage in das Distanzmaß miteinbezogen wird, sprechen wir von einer topographisch beeinflussten *GN-Gruppierung*, kurz *TGN-Gruppierung*. Die nächsten Nachbarn sind invariant unter monotonen Transformationen der Ähnlichkeitswerte und damit auch das Verfahren insgesamt.

Die Gruppierungen von Punkten in der Ebene (Abb. 3–6) zeigen, daß dieses Clusterverfahren auch in schwierigen Fällen eine plausible Lösung findet. Bei eher globalen Clustern wie in Abb. 3 und Abb. 5 ist es sinnvoll, eine relativ große Anzahl von Nachbarn zu betrachten, bei lokalen Clustern wie in Abb. 4 sollte nur eine kleine Nachbarschaft berücksichtigt werden. Ein besonderes Merkmal dieser Methode ist es, daß sie im Gegensatz zu den oben erwähnten Verfahren Cluster mit unterschiedlichen Punktdichten erkennen kann (Abb. 5). Solche Cluster ergeben sich zum Beispiel, wenn man Orte mit unterschiedlichen Siedlungsdichten trennen möchte.

Jarvis⁶ untersuchte die Stabilität des Verfahrens anhand von Punktgruppen in der Ebene. In einem Experiment verschob er die Punkte leicht, in einem anderen berücksichtigte er einen kleinen Teil der Punkte nicht. Das Ergebnis der Experimente war, daß das Clusterergebnis weitgehend unabhängig von Störungen blieb.

⁵ Jarvis u. Patrick 1973; Jarvis 1978.

⁶ Ebd.

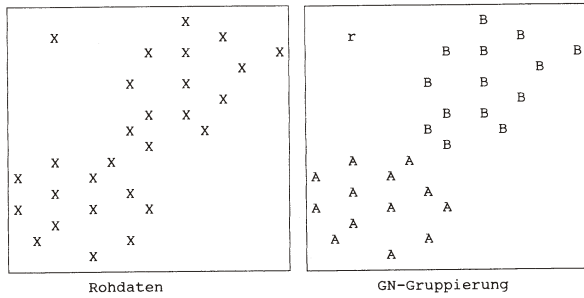


Abb. 3. Modell: Zwei sich berührende globale Gruppen. – Es wurden die neun nächsten Nachbarn betrachtet. Innerhalb einer Gruppe hat jede Einheit mit mindestens einer anderen Einheit der Gruppe sieben Nachbarn gemeinsam. Die einzeln liegende Einheit wurde keiner der beiden Gruppen zugeordnet. Insgesamt handelt es sich um eine eher globale Struktur

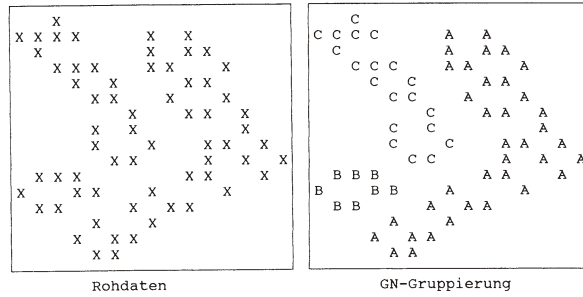


Abb. 4. Modell: Zwei nebeneinander liegende „Bananas“. – Hier handelt es sich eher um lokale Strukturen. Deshalb wurden nur die drei nächsten Nachbarn betrachtet. Innerhalb einer Gruppe hat jede Einheit mit einer anderen Einheit der Gruppe mindestens einen Nachbarn gemeinsam

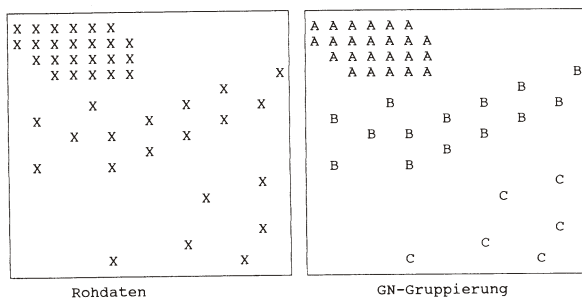


Abb. 5. Modell: Cluster mit unterschiedlichen Punktdichten. – Es handelt sich um eher globale Strukturen, die aber nicht kreisförmig sind. Für jede Einheit wurden die sieben nächsten Nachbarn bestimmt. Innerhalb eines Clusters hat jede Einheit mindestens einen Nachbarn, mit dem sie vier Nachbarn gemeinsam hat

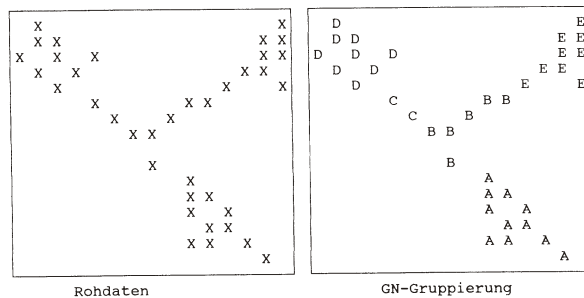


Abb. 6. Modell: Globale und lokale Strukturen gemischt. – Es wurden zehn Nachbarn betrachtet. Einheiten innerhalb eines Clusters haben mit einem ihrer Nachbarn mindestens acht Nachbarn gemeinsam

Bei Single-Linkage-Clustering ist das Ergebnis nicht notwendig eindeutig, wenn es keinen eindeutigen nächsten Nachbarn einer Einheit gibt, sondern m Nachbarn den gleichen Ähnlichkeitswert haben. Bei archäologischen Daten tritt dieses Problem besonders häufig bei Anwesenheits-/Abwesenheitsdaten auf, insbesondere beim Russel/Rao-Ähnlichkeitskoeffizienten. In abgeschwächter Form ist diese Schwierigkeit bei der GN-Gruppierung vorhanden, wenn nämlich die Zahl m der nächsten Nachbarn mit gleichem Ähnlichkeitswert die Zahl der betrachteten nächsten Nachbarn überschreitet. In den von uns verwendeten Programmen führt dieser Fall zu einer *Warnung* an den Benutzer; mögliche Abhilfe bietet die Erhöhung der Zahl der betrachteten Nachbarn oder die Verwendung eines anderen Ähnlichkeitskoeffizienten (Abb. 6).

Anwendungsbeispiele der GN- und TGN-Gruppierung an archäologischen Daten

Im folgenden werden Leistungsfähigkeit und Eigenschaften der GN- und TGN-Gruppierung im Hinblick

auf reale archäologische Probleme untersucht. Dazu werden zwei unterschiedliche und unseres Erachtens typische Testfälle herangezogen, für die jeweils bereits ein stabiles Ergebnis vorliegt, das mit konventionellen methodischen Ansätzen gewonnen wurde. Ziel dieser Versuche ist zunächst, die bereits konventionell erarbeiteten Ergebnisse mit Hilfe einer GN-Gruppierung möglichst gut zu reproduzieren.

Fallstudie fränkisches Gräberfeld

Als erster Testfall dient das fränkische Gräberfeld von Köln-Müngersdorf; es wurde in den 1920er Jahren recht sorgfältig und vollständig ergraben und 1955 von Fritz Fremersdorf publiziert⁷. Die 150 Bestattungen des Gräberfeldes⁸ gehören in die Zeit zwischen etwa 470 und 650 n. Chr. Von dem in der Merowingerzeit übli-

⁷ Fremersdorf 1955.

⁸ Bei den Befunden Nr. 3 und 7 ist es eher unwahrscheinlich, daß es sich um Gräber handelt; sie wurden nicht berücksichtigt.

Anz. Typen	Anz. Gräber
12:	1
11:	0
10:	0
9:	0
8:	2
7:	1
6:	6
5:	5
4:	5
3:	12
2:	27
1:	42
0:	49

Abb. 7. Köln-Müngersdorf. Zahl der ansprechbaren Typen pro Grab. – 49 Gräber enthalten keinen, 42 Gräber nur einen Typ usw. Mittlere Fundanzahl pro Grab, bezogen auf alle Gräber: 1,65; bezogen nur auf fundführende Gräber: 2,48

chen Grabraub wurde es eher unterdurchschnittlich beeinträchtigt und weist deshalb vergleichsweise „reiche“ Grabinventare auf (Abb. 7). Neben vielfältigen Detailuntersuchungen liegt eine neuere Gesamtanalyse vor, bei der mit Hilfe der chorologischen Methode sieben Belegungsphasen herausgestellt wurden (Abb. 8)⁹. Diese fügen sich zwanglos in ein umfassenderes Chronologiesystem ein, das auf dem Fundstoff des gesamten Niederrheins beruht¹⁰.

Modell

Ziel der Untersuchung ist es, die Bestattungen nach ihren Beigaben und ihrer Gräberfeldlage chronologisch zu gruppieren. Den konventionellen Untersuchungen zur chronologischen Gliederung fränkischer Gräberfelder liegen gemeinhin zwei unterschiedliche methodische Ansätze zugrunde: Die chorologische Methode nutzt das häufig recht regelmäßige Wachstum der Gräberfelder aus und argumentiert aufgrund des Vorkommens verschiedener Typen in als gleichzeitig erachteten Gräberfeldarealen¹¹. Die Untersuchung der Fundkombination beruht auf dem gemeinsamen Vorkommen von Typen in geschlossenen Grabinventaren; technische Hilfsmittel solcher Analysen sind Kombinationstabelle und Korrespondenzanalyse.

Mit der TGN-Gruppierung soll versucht werden, beide Ansätze miteinander zu verbinden; das oben entwickelte Distanzmaß ermöglicht es, neben dem Typenspektrum eines Grabes in wählbarer Gewichtung auch seine topographische Lage zu berücksichtigen.

⁹ Siegmund 1989, 61–71; die älteren Untersuchungen sind ebd. 61 Anm. 4 zusammengestellt.

¹⁰ Ebd. 87–114.

¹¹ Eine ausführliche Diskussion der Methoden ebd. 24–35.

Zwar wäre eine GN-Gruppierung auch für Inventare möglich, die nur einen Typ enthalten, doch führt die Berücksichtigung der Lage bei diesen Gräbern zu einer besseren Anbindung. Für eine Korrespondenzanalyse fallen solche fundarmen Inventare aus. Da in der Merowingerzeit Frauen- und Männergräber eine größere Zahl geschlechtsspezifischer Beigaben aufweisen, müssen sie bei der Korrespondenzanalyse getrennt behandelt werden; die GN-Gruppierung mit Berücksichtigung der topographischen Distanz würde neben der beiden Geschlechtern gemeinsamen Gefäßbeigabe die Lage als eine weitere nicht geschlechtsspezifische „Fundgattung“ hinzufügen. Da dann die geschlechtsspezifischen Beigaben das Bild nicht mehr so stark dominieren, könnten Frauen- und Männergräber einer gemeinsamen Analyse unterzogen werden. So würde die TGN-Gruppierung die Vorzüge von chorologischer Methode und Korrespondenzanalyse bei der Gräberfeldanalyse vereinen.

GN- und TGN-Gruppierung

Für das Gräberfeld von Müngersdorf wurde zunächst eine GN-Gruppierung ohne Berücksichtigung der topographischen Distanz durchgeführt¹². Es ergibt sich auf der Ebene 11 von 15 Nachbarn eine sinnvolle Gliederung des Materials in acht Gruppen; lediglich die erste Gruppe bedarf weiterer Untergliederung und wird hier auf der Ebene 14 von 15 Nachbarn betrachtet (Abb. 9–10)¹³.

Der Vergleich von konventionellem und Clusterergebnis zeigt das nur befriedigende Resultat der GN-Gruppierung, verweist aber auch auf die problematischen Bereiche des Gräberfeldes (Abb. 11). Die Belegungsphase A und die von fundreichen Frauengräbern bestimmte Phase B werden problemlos durch die GN-Gruppierung reproduziert. Auch die Phasen E bis G werden befriedigend wiedergefunden; dies begründet sich in einer Anzahl markanter und häufigerer Typen¹⁴, wobei zeitgleiche Frauen- und Männergräber offensichtlich hinreichend fest miteinander verbunden sind. Die Gräber die Belegungsphase D jedoch sind auf mehrere kleinere Cluster verteilt, die der Phase C fallen durchweg in den Restbereich¹⁵; in diesem mittleren

¹² Gräber ohne verwertbare Funde, für die nur ihre Lage als einzige Information zur Verfügung stand, wurden aus allen Berechnungen ausgeschlossen.

¹³ Wie oben dargestellt, ergeben sich bei hoher Nachbarzahl eher globale, bei kleiner Nachbarzahl eher lokale Strukturen. Für erste Berechnungen hat sich unseres Erachtens eine Nachbarzahl von 10% der Befundeinheiten als sinnvoll erwiesen.

¹⁴ Z. B. Breitsaxe, dreiteilige Gürtelgarnituren, monochromgeometrische Tauschiermuster.

¹⁵ In den Restbereich (Residuum) fallen die Bestattungen, die auf dem Niveau von hier elf gemeinsamen Nachbarn keiner

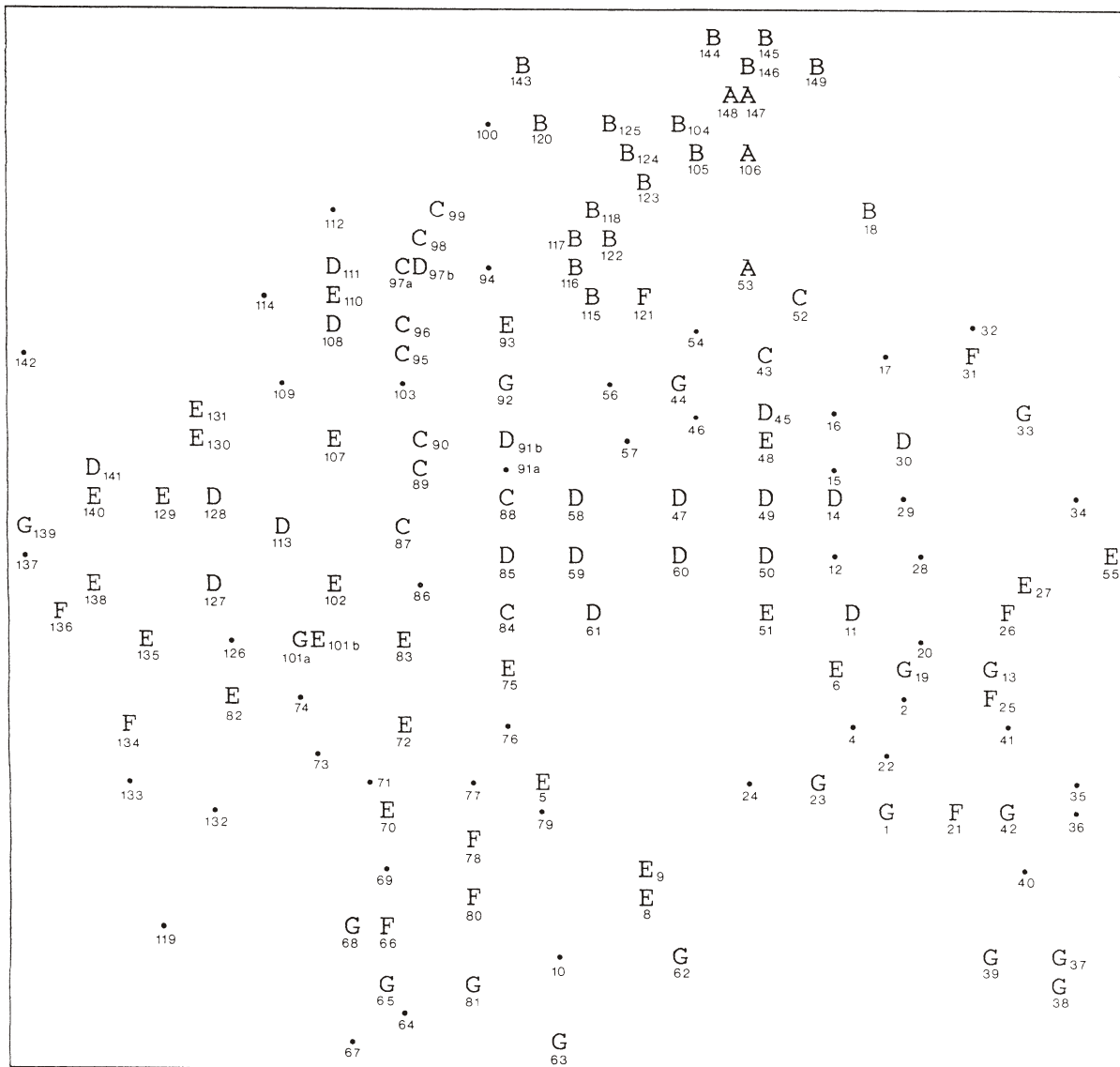


Abb. 8. Köln-Müngersdorf. Belegungsphasen nach einer konventionellen chorologischen Analyse.
Nach Siegmund 1989, 71 Abb. 10

Gräberfeldabschnitt fehlen häufigere Typen und die Verbindung von Frauen- und Männergräbern bereitet Probleme¹⁶.

Gruppe mit drei oder mehr Gräbern zugeordnet werden konnten.

¹⁶ Eine gegen die Regel gemeinsam für Frauen- und Männergräber berechnete Korrespondenzanalyse führt zu ähnlichen Ergebnissen: Die Belegungsphasen A und B sowie F und G werden gut reproduziert, während sich im Mittelabschnitt C bis E eine im Detail nur wenig befriedigende Ordnung ergibt. Für die, sicherlich korrektere, getrennte Behandlung von Frauen- und Männergräbern ist das Gräberfeld jedoch zu klein.

Entsprechend dem oben entwickelten Modell wurde eine TGN-Gruppierung unter Einbeziehung der topographischen Distanz berechnet. Das hier wiedergegebene Beispiel (Abb. 12) zeigt ein typisches Bild, das sich bei den unterschiedlichsten Versuchen immer wieder in ähnlicher Form einstellte; selbst bei nur 2%igem Einfluß der topographischen Lage auf die Gesamtdistanz kommt es zur Bildung lokaler Gruppen¹⁷. Das zumin-

¹⁷ Je höher der topographische Anteil gewichtet wird, desto stärker wird die Tendenz zu lokalen Clustern. Aber auch ein nur 1%iger Einfluß der Lage auf die Gesamtdistanz führt zu ähnlichen Bildern. Eine Erklärung neben anderen bildet die, statistisch gesehen, recht geringe Fundfrequenz pro Grab (vgl. Abb. 7).

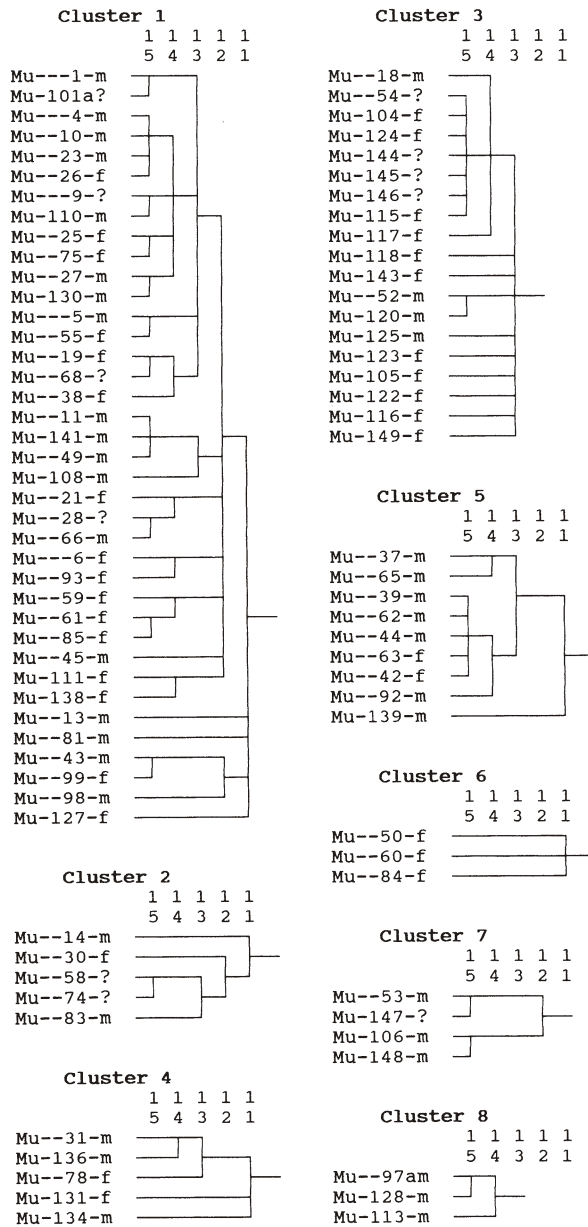


Abb. 9. GN-Gruppierung Köln-Müngersdorf: Dendrogramm

dest in der Peripherie kreisförmige Wachstum des Gräberfeldes wird nicht erfaßt. Das Ergebnis ist damit letztlich enttäuschend, denn mehrere der entstehenden Lokalgruppen beinhalten jeweils Inventare dreier Zeitphasen; auch die oben eigentlich gut getrennten Phasen E, F und G sind hier „vermischt“.

Bewertung

Der Versuch zeigt, daß der Ansatz für diese Art von Daten nur bedingt geeignet ist. Ein wichtiger Grund

Belegungs-Phasen konventionelle Analyse	Cluster GN-Gruppierung
A	7
B	3
C	Residuum
D	1-2, 1-5, 2, 6, 8
E	1-1
F	1-4, 4
G	1-3, 5

Abb. 11. GN-Gruppierung Köln-Müngersdorf: tabellarischer Vergleich mit dem konventionell gewonnenen Ergebnis (vgl. Abb. 10)

Cluster TGN-Gruppierung	Belegungs-Phasen konventionelle Analyse
1	E - F - G
2	E - F - G
3	B
4	G
5	D
6	D
7	D
8	G
9	E
10	C - D - E
11	A

Abb. 13. TGN-Gruppierung Köln-Müngersdorf: tabellarischer Vergleich mit dem konventionell gewonnenen Ergebnis (vgl. Abb. 12)

liegt sicherlich in dem Problem, daß bei Gräberfeldern die topographische Distanz nicht in alle Richtungen gleich bewertet werden kann. Verläuft das „Wachstum“ beispielsweise recht regelmäßig von Ost nach West, sind jeweils nördlich oder südlich benachbarte Gräber potentiell zeitähnlicher als gleich entfernte in östlicher oder westlicher Richtung. Viele andere Wachstumsmodelle sind denkbar und belegt, z. B. mono- oder polyzentrische und gerichtete, bei denen es zu plötzlichen Änderungen der generellen Wachstumsrichtung kommt. Es ist jedoch nicht vertretbar, für jedes Wachstumsmodell ein passendes Distanzmaß zu entwickeln.

Danach bleiben die konventionelle Gräberfeldanalyse per Kartierung einerseits und die nach Geschlechtern getrennte Korrespondenzanalyse auf breiterer Materialbasis die Verfahren der Wahl. Immerhin scheint es sinnvoll, versuchsweise auch eine einfache GN-Gruppierung durchzuführen; das Einbringen der räumlichen Distanzen scheint in diesem Fall ein Irrweg zu sein.

Fallstudie jungpaläolithischer Siedlungsplatz

Als zweiter Testfall dient der Fundplatz Sprendlingen in Rheinhessen, der anhand seiner Steinartefakte in das Mittlere Jungpaläolithikum (Gravettien, ca.

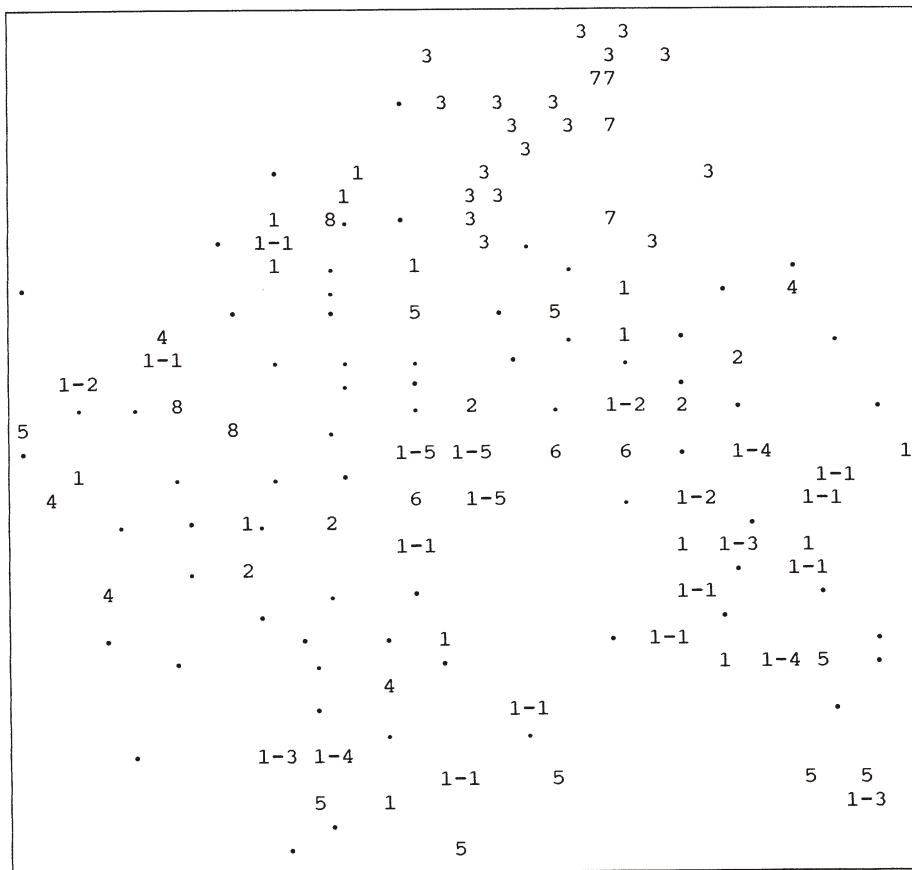


Abb. 10. GN-Gruppierung Köln-Müngersdorf: Kartierung (Jaccard-Koeffizient, 11 bzw. 14 von 15 Nachbarn) (vgl. Abb. 11)

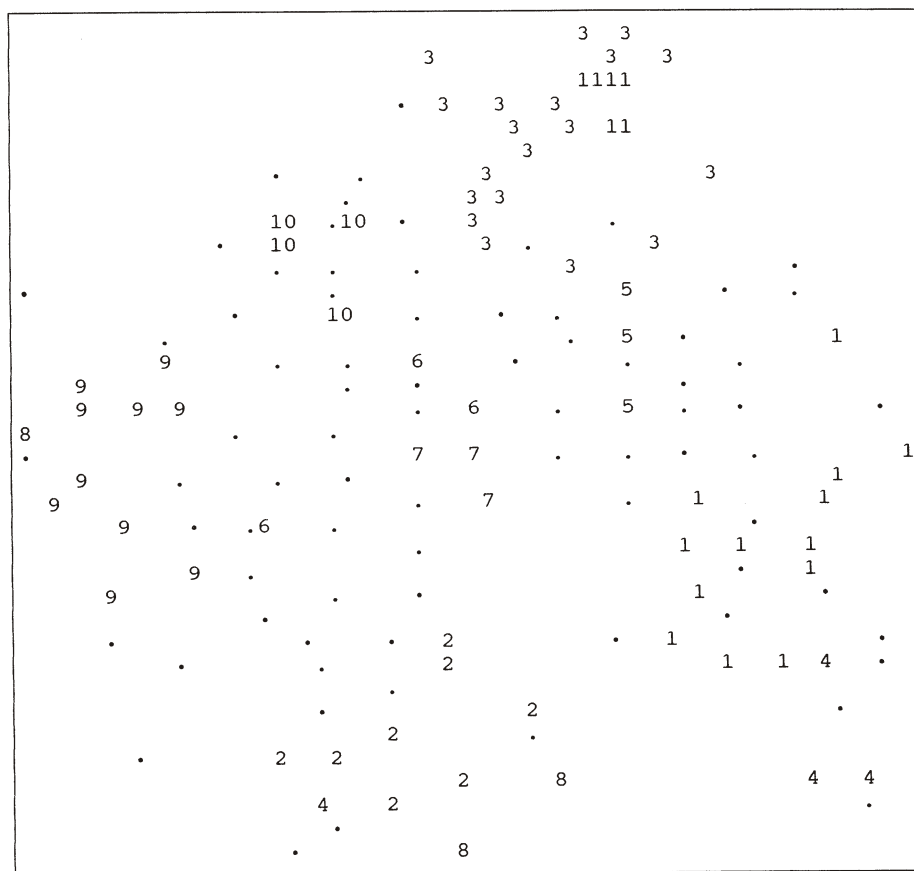


Abb. 12. TGN-Gruppierung Köln-Müngersdorf (2% topographische Komponente, 98% Jaccard-Koeffizient, 11 von 15 Nachbarn) (vgl. Abb. 13)

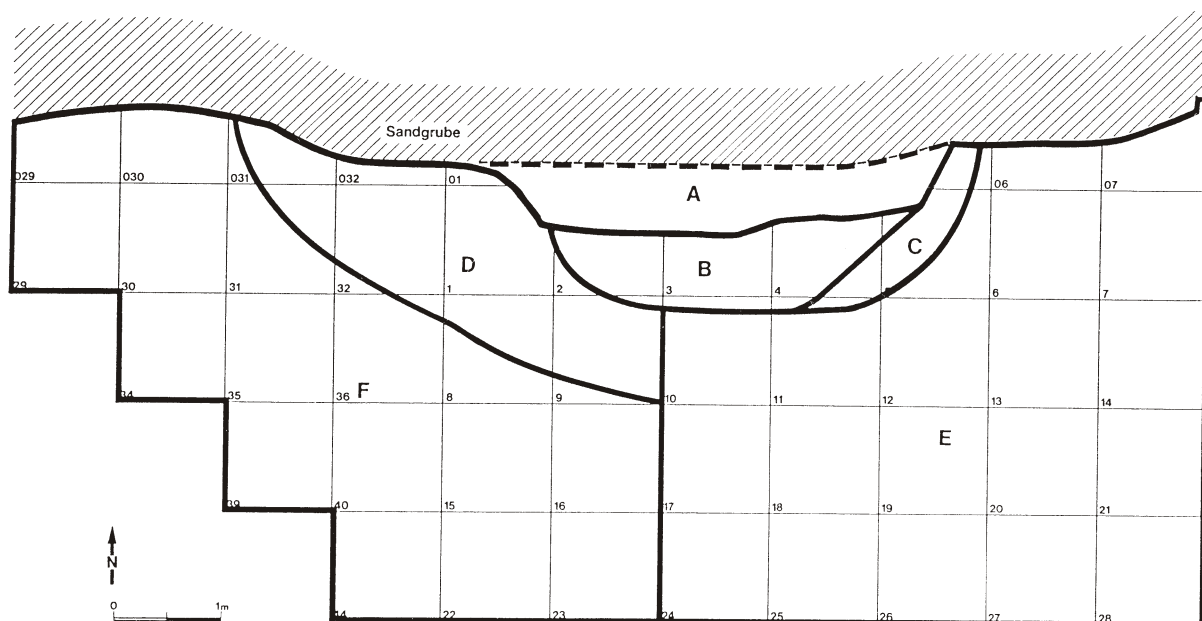


Abb. 14. Sprendlingen. Räumliche Gliederung des Siedlungsplatzes. Nach Bosinski u. a. 1985, 86 Abb. 64

Zone	A	B	C	D	E	F
Rotfärbung	!!	!	!	--	--	--
Steinartefakte	!!	!!	--	+	+	--
Absplisse	!!	!!	--	--	--	--
Steine	!	+	--	+	+	+
Knochen	!	!	--	--	--	--
Mollusken	!!	!!	--	--	+	--
Besonderheiten	Holzkohle, viele Kalk- steinplatten	Stichellamellen Rückenmesser, Mikrogravet- ten, Kratzer, Klingen		Stichel, Spitzklingen, zerbrochene Schlagsteine	zerbrannte Kalk- brocken	zwei zusammen- passende Kerne
Interaktion	Zusammenpassungen		keine ← Verbindung →			
Deutung	Zentrum + Feuerstelle	Randbezirk von A, Leder- und Fell- bearbeitung	Leerraum	Arbeitsplatz für grobe Tätigkeit	Vorplatz, Abraum der Feuerstelle	Peripherie
Symbole:	!! sehr viel	! viel	+ vorhanden	-- wenig	-- keine	

Abb. 15. Sprendlingen. Übersicht zur Zonengliederung des Fundplatzes. Nach Bosinski u. a. 1985 (vgl. hier Abb. 14)

30 000 – 20 000 v. Chr.) eingeordnet werden kann¹⁸. Von einer Sandgrube angeschnitten, konnte 1978 die noch unzerstörte Hälfte des Siedlungsplatzes ergraben und anschließend publiziert werden.

¹⁸ Bosinski u. a. 1985. – Allgemein: E. Czesla, Das Mittlere Jungpaläolithikum im westlichen Deutschland. Arch. Informationen 7, 1984, 28–39; G. Bosinski, Die jägerische Geschichte des Rheinlandes/Einsichten und Lücken. Jahrb. RGZM 30, 1983, 81–112; bes. 92 ff. mit Tab. 2.

Sprendlingen scheint als Testfall geeignet, da der Fundplatz in nahezu mustergültiger Weise mit vielen und genauen Kartierungen vorgelegt wurde, so daß alle interessierenden Informationen leicht aus der Publikation zu entnehmen sind. Als Ergebnis haben die Autoren eine klare Hypothese zur räumlichen Gliederung des Platzes entwickelt (Abb. 14)¹⁹; die inhaltliche Be-

¹⁹ Bosinski u. a. 1985, 85–89; insbesondere Abb. 64–65.

deutung dieser Zonen wurde hier nach den Thesen der Verfasser in Abb. 15 tabellarisch zusammengefaßt. Dabei umreißt die Zone A einen Bereich, in dem unmittelbar nach der Entdeckung der Fundstelle eine kurze Voruntersuchung stattfand; die Artefakte wurden hier noch nicht einzeln eingemessen und die Fläche nicht wie bei der späteren Ausgrabung nach Viertelquadraten aufgeteilt. Archäologisch entsprechen sich die Zonen A und B, wobei letztere lediglich den technisch besser ergrabenen südlichen Randbezirk darstellt²⁰.

Im Vergleich zu Köln-Müngersdorf weist der Fall Sprendlingen eine andere Art von Daten auf: Die „Befunde“ sind hier dicht aneinander grenzende Flächen, in denen auf dem Niveau von Viertelquadraten die Häufigkeit verschiedener Artefaktgattungen und -typen ausgezählt wurde²¹; als weiterer Befund kommen die Zusammenpassungen von Steinartefakten hinzu²².

Modell

Ziel der Analyse ist es, die einzelnen Befunde (hier: Viertelquadrate) zu Zonen ähnlicher Artefaktzusammensetzung zusammenzufassen und gegen andere Räume abzugrenzen, um so verschiedene Aktivitätsbereiche auf dem Siedlungsplatz nachzeichnen und in ihrem Wesen beschreiben zu können. Es ist anzunehmen, daß die menschlichen Siedlungsaktivitäten über ein einzelnes Viertelquadrat hinausgehen, jedoch nur einen begrenzten Bereich des Fundplatzes abdecken; demnach wären benachbarte Befunde potentiell ähnlicher als stärker entfernte. Diese Überlegung macht es auch hier wünschenswert, in die Berechnung der Ähnlichkeit neben dem Artefaktenspektrum die topographi-

sche Distanz zusätzlich einzubeziehen; eine TGN-Gruppierung müßte zu sinnvollen Ergebnissen führen.

GN- und TGN-Gruppierung

Zunächst wurde zum Vergleich eine GN-Gruppierung einzig aufgrund der Artefakte, ohne Berücksichtigung der topographischen Komponente, durchgeführt (Abb. 16) (vgl. Anm. 20). Im zentralen Bereich ergibt sich eine möglicherweise befriedigende Zonierung; in seiner Peripherie jedoch werden Flächen zusammengefaßt, die weit auseinander liegen, was nicht dem oben formulierten Modell entspricht. Gerade die fundarmen Quadrate fallen weitgehend in den ungegliederten Restbereich („Residuum“)²³.

TGN-Gruppierungen unter Einberechnung einer 10%igen topographischen Komponente führen dagegen zu dem angestrebten Ergebnis (Abb. 17–18)²⁴. Betrachtet man auf der Suche nach globalen Strukturen viele Nachbarn, ergibt sich eine Trennung in Zentrum und zwei unterschiedliche periphere Zonen; der dazwischen liegende Übergangsbereich fällt in das Residuum (Abb. 17)²⁵. Diese Trennung der fundärmeren Peripherie in eine östliche und eine westliche Zone war auch bei der konventionellen Bearbeitung herausgestellt worden²⁶. Bei einer Betrachtung relativ weniger Nachbarn zur Auffindung eher lokaler Strukturen fällt die gesamte Peripherie in lediglich ein Cluster (Abb. 18); hier wird dafür der zentrale Bereich sinnvoll gegliedert. Das Zentrum wird nach Osten hin von einem Cluster umgeben, während sich im Westen zwei kleinräumigere Zonen anschließen; dieses Bild entspricht weitgehend dem der konventionellen Bearbeitung.

Für eine detailliertere Diskussion des Ergebnisses werden im folgenden die Cluster der ersten TGN-

²⁰ Um eine bessere Vergleichbarkeit der Daten zu erreichen, wurden für alle Analysen die Mengenangaben für die Zone A durch 8 dividiert, da sie etwa acht Viertelquadrate umfaßt. Weiterhin wurden bei nur randlich angeschnittenen Einheiten die Häufigkeiten mit 2 multipliziert (Quadrate: 032; 01; 1–1; 2–1; 2–4; 3–1; 3–4; 4–1; 4–4; 5–1); – Die Ziffern nach dem Bindestrich benennen jeweils das Viertelquadrat, die Zahlen wurden (rechts oben beginnend) im Uhrzeigersinn vergeben.

²¹ Größere Artefakte wurden auf der Grabung einzeln eingemessen und in der Publikation entsprechend kartiert, nicht eingemessene Artefakte pro Viertelquadrat ausgezählt. Für die GN-Gruppierung wurden die Einzeleinmessungen auf das Viertelquadratsniveau generalisiert. – Die eingegebenen Daten sind, statistisch gesehen, „häßlich“; die absoluten Häufigkeiten nehmen zum Teil sehr hohe Werte an (z. B. Steinartefakte gesamt, Absplisse), zum Teil sehr niedrige Werte (z. B. Lamellen, Stichlamellen etc.). Daneben liegen Anwesenheits-/Abwesenheitsbeobachtungen vor, z. B. die Zusammenpassungen. Alle uns sinnvoll erscheinenden Transformationen (z. B. relative Häufigkeiten, Wurzeltransformation) wurden versuchsweise durchgeführt, ergaben aber durchweg unbefriedigende Bilder.

²² Zur Methode zuletzt (mit vertiefter Interpretation des Sprendlinger Befunds): Cziesla 1987.

²³ Die leeren Quadrate wurden in die TGN-Gruppierung einbezogen. Ein versuchsweise vorgenommener Ausschluß der leeren Befunde führte zu unbefriedigenden Bildern, was nach dem oben geschilderten Modell den Erwartungen entspricht.

²⁴ Der Ansatz einer 10%igen Gewichtung des Faktors topographische Distanz ist eine willkürliche Entscheidung des Bearbeiters. Versuche haben jedoch gezeigt, daß eine Veränderung dieses Wertes in der Spanne zwischen 5% und 15% kaum gravierende Konsequenzen für das Ergebnis hat. In der Tendenz führt eine niedrigere Gewichtung zu räumlich nicht zusammenhängenden Clustern, während eine hohe Gewichtung kleinräumige Gruppen fördert und der Residualbereich wächst.

²⁵ Der Wert von acht bzw. 32 Nachbarn geht auf 5% bzw. 20% der hier untersuchten 167 Viertelquadrate zurück. – Der breite „Residualgürtel“ zwischen Zentrum und Peripherie ist das typische Ergebnis eines gruppierenden Verfahrens: Die Extrema werden gut reproduziert, während die Bereiche allmählichen Übergangs zwischen den Gruppen jeweils unähnlich zu beiden sind und deshalb in das Residuum fallen.

²⁶ Bosinski u. a. 1985, 86 Abb. 64: Zonen „E“ und „F“.

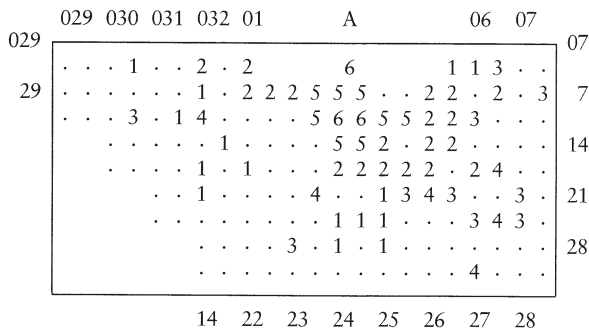


Abb. 16. Sprendlingen. GN-Gruppierung ohne topographische Komponente (4 von 8 Nachbarn)

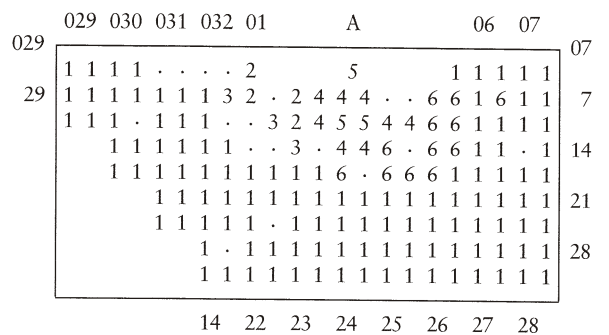


Abb. 18. Sprendlingen. TGN-Gruppierung mit 10% topographischer Komponente (4 von 8 Nachbarn)

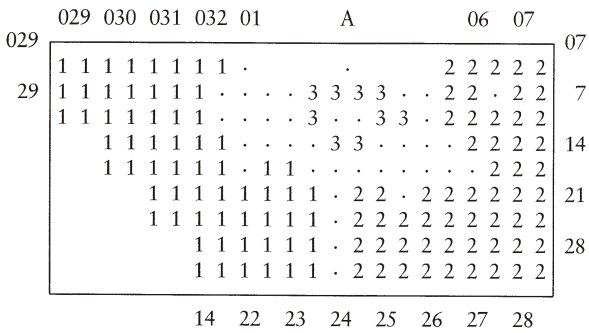
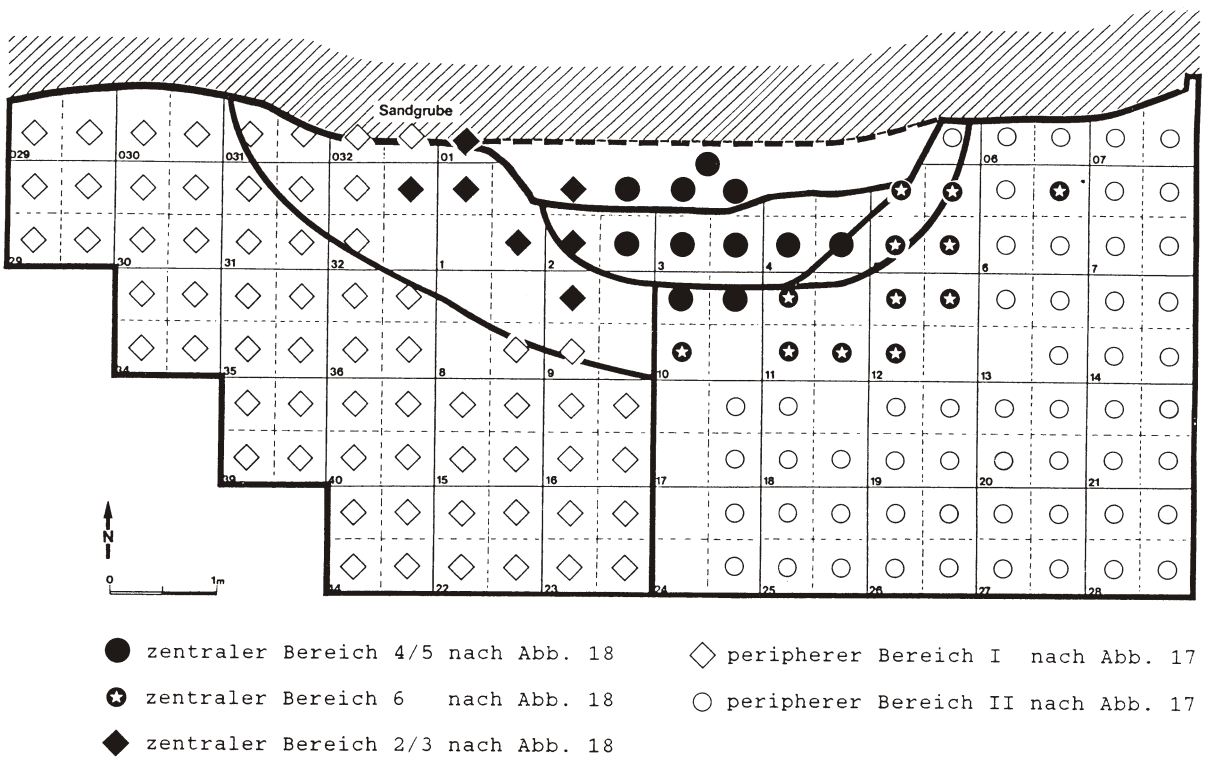


Abb. 17. Sprendlingen. TGN-Gruppierung mit 10% topographischer Komponente (28 von 32 Nachbarn)

Gruppierung mit römischen, die der zweiten mit arabischen Ziffern bezeichnet. Der Plan Abb. 19 zeigt die von G. Bosinski und E. Cziesla vorgeschlagene Gliederung und die hier errechneten Zonen: die Peripherie nach der TGN-Gruppierung mit 28 von 32 Nachbarn (Cluster I–II), den zentralen Bereich nach der TGN-Gruppierung mit vier gemeinsamen von acht betrachteten Nachbarn (Cluster 2–6).

Die Cluster 4 und 5 umfassen den Bereich der dichtesten Fundstreuung. Cluster 5 entspricht in seinem Inhalt Cluster 4 und beschreibt lediglich den extrem dicht besetzten Kernbereich; beide Inhalte wurden deshalb zur besseren Übersicht in der Tabelle Abb. 20



- zentraler Bereich 4/5 nach Abb. 18
- ★ zentraler Bereich 6 nach Abb. 18
- ◆ zentraler Bereich 2/3 nach Abb. 18
- ◇ peripherer Bereich I nach Abb. 17
- peripherer Bereich II nach Abb. 17

Abb. 19. Sprendlingen. Vergleich der Zonengliederung des Fundplatzes (nach Bosinski u. a. 1985) mit dem Ergebnis der TGN-Gruppierung (vgl. Abb. 20–21)

Merkmale	Cluster: 4+5	6	2+3	Rest	1
Rotfärbung (Abb. 12)	11	4	3	2	1
Steinartefakte Summe (Abb. 26)	314	57	48	104	32
Abspliß (Abb. 33)	264	47	36	57	19
Klinge, retuschiert (Abb. 48)	9		4	9	2
zerbrannter Kalkbrocken (Abb. 15)	47	19	18	33	22
aufeinander-gesinterte Stücke (Abb. 29)	3		5	9	
Knochen (Abb. 13)	160	5	9	14	9
Pferde-Knochen (Abb. 23)	6				
Rentier-Knochen (Abb. 23)	14	1			
Abschlag (Abb. 32)	24	5	2	8	4
Klinge, unretuschiert (Abb. 48)	6	1	1	1	
Quarz (Abb. 17)	7	3		4	
Lamelle (Abb. 42)	7	1		8	
Stichellamelle (Abb. 42)	9		2	5	
Rückenmesser (Abb. 48)	7			4	
Gravette-Spitze (Abb. 48)	1			1	
Kratzer (Abb. 31)	4	1			
Spitzklinge (Abb. 31)	1		3	4	3
Stichel (Abb. 31)	3	1	9	5	
Schlagstein 2 (Abb. 16)			2	1	
Schlagstein 3 (Abb. 16)			1	1	2
Schlagstein 1 (Abb. 16)				1	1
Kern (Abb. 31)				3	1
Aneinanderpassung 20 (Abb. 53)	2				
Aneinanderpassung 21 (Abb. 53)	2				
Aneinanderpassung 22 (Abb. 53)	2				
Aneinanderpassung 24 (Abb. 53)	1			1	
Aneinanderpassung 25 (Abb. 53)	1			1	
Aufeinanderpassung 19 (Abb. 53)	1			2	
Aufeinanderpassung 14 (Abb. 53)	1			2	
Anpassung 5 (Abb. 53)	1			1	
Aneinanderpassung 18 (Abb. 53)	1				1
Aneinanderpassung 23 (Abb. 53)	3				1
Aneinanderpassung 16 (Abb. 53)	1	2			
Aneinanderpassung 17 (Abb. 53)	1	1			
Aufeinanderpassung 6 (Abb. 53)		1		1	
Aufeinanderpassung 13 (Abb. 53)		1		1	
Aneinanderpassung 10 (Abb. 53)			2		
Aneinanderpassung 3 (Abb. 53)			1	1	
Anpassung 4 (Abb. 53)			1	1	
Aneinanderpassung 8 (Abb. 53)			2	1	
Aneinanderpassung 9 (Abb. 53)			2	1	
Aneinanderpassung 11 (Abb. 53)			1		1
Aneinanderpassung 1 (Abb. 53)				2	
Aneinanderpassung 2 (Abb. 53)				3	
Aneinanderpassung 12 (Abb. 53)				2	
Aneinanderpassung 7 (Abb. 53)				1	1
Aneinanderpassung 15 (Abb. 53)				1	1

Abb. 20. Sprendlingen. Typ-Statistik für 4 von 8 Nachbarn (Abbildungsnummern wie Bosinski u. a. 1985)

zusammengefaßt. Die Cluster 4 und 5 decken den größten Teil der intensiv rot gefärbten Fläche ab, hier liegt das Häufigkeitsmaximum der Steinartefakte und Absplisse sowie der Knochen. Von den Werkzeugen finden sich die meisten Lamellen und Stichellamellen, Rückenmesser und Kratzer in diesem Bereich. Die beiden Cluster entsprechen in Fläche und Inhalt der von Bo-

sinski und Cziesla herausgestellten Zone A/B (Abb. 15; 19).

Östlich davon grenzt das Cluster 6 an; es weist einen ähnlichen Inhalt wie die Cluster 4/5 auf, bei allerdings deutlich geringeren absoluten Häufigkeiten. Es entspricht inhaltlich der Zone C, die einen Ausdünnungsbereich der zentralen Zone A/B darstellt; das Ergebnis

Merkmale	Cluster: I	II	III	Rest
Steinartefakte Summe (Abb. 26)	18	21	188	423
Abschlag (Abb. 32)	2	1	19	21
Abspliß (Abb. 33)	4	13	145	261
Quarz (Abb. 17)	2	1	3	8
Klinge, retuschiert (Abb. 48)	1	1	6	16
zerbrannter Kalkbrocken (Abb. 15)	9	16	26	88
Klinge, unretuschiert (Abb. 48)			2	7
Kratzer (Abb. 31)			2	3
Lamelle (Abb. 42)			8	8
Stichellamelle (Abb. 42)			10	6
Rückenmesser (Abb. 48)			8	3
Gravette-Spitze (Abb. 48)			2	
Pferde-Knochen (Abb. 23)			1	5
Rentier-Knochen (Abb. 23)			9	6
Kern (Abb. 31)	2	2		
aufeinander-gesinterte Stücke (Abb. 29)	1		3	13
Stichel (Abb. 31)	3	1	3	11
Spitzklinge (Abb. 31)	2		1	8
Schlagstein 1 (Abb. 16)	2			
Schlagstein 2 (Abb. 16)	1			2
Schlagstein 3 (Abb. 16)	1			3
Rotfärbung (Abb. 12)		3	9	9
Knochen (Abb. 13)		6	120	71
Aneinanderpassung 22 (Abb. 53)			2	
Aneinanderpassung 17 (Abb. 53)			1	1
Aneinanderpassung 20 (Abb. 53)			1	1
Aneinanderpassung 21 (Abb. 53)			1	1
Aneinanderpassung 8 (Abb. 53)				3
Aneinanderpassung 9 (Abb. 53)				3
Aneinanderpassung 10 (Abb. 53)				2
Aneinanderpassung 16 (Abb. 53)				3
Aneinanderpassung 24 (Abb. 53)				2
Aneinanderpassung 25 (Abb. 53)				2
Aneinanderpassung 12 (Abb. 53)	2			
Anpassung 5 (Abb. 53)	1		1	
Aneinanderpassung 15 (Abb. 53)	1		1	
Aneinanderpassung 1 (Abb. 53)	1			1
Aneinanderpassung 2 (Abb. 53)	1			2
Aneinanderpassung 3 (Abb. 53)	1			1
Anpassung 4 (Abb. 53)	1			1
Aufeinanderpassung 6 (Abb. 53)	1			1
Aneinanderpassung 7 (Abb. 53)	1			1
Aneinanderpassung 11 (Abb. 53)	1			1
Aufeinanderpassung 13 (Abb. 53)	1			1
Aufeinanderpassung 14 (Abb. 53)	1			2
Aneinanderpassung 18 (Abb. 53)		1	1	
Aneinanderpassung 23 (Abb. 53)		1	3	
Aufeinanderpassung 19 (Abb. 53)		1	1	1

Abb. 21. Spredlingen. Typ-Statistik für 28 von 32 Nachbarn (Abbildungsnummern wie Bosinski u. a. 1985)

der TGN-Gruppierung legt allerdings nahe, diesen Bereich etwas weiträumiger zu fassen.

Auch die Cluster 2 und 3 unterscheiden sich untereinander lediglich etwas in den absoluten Häufigkeiten ihres Inhalts und wurden deshalb zur besseren Übersicht in der Tabelle Abb. 20 zusammengefaßt. Sie wei-

sen kaum Rotfärbung und wesentlich weniger Steinartefakte auf, die für Cluster 4/5 typischen Werkzeuge sind selten. Dagegen treten hier Stichel und Spitzklingen auf sowie zwei der drei Schlagsteine. Damit entsprechen sie in Inhalt und Ausdehnung weitgehend der Fundplatzzone D (Abb. 15; 19).

Dieses Bild wird durch die Zusammenpassungen unterstrichen (Abb. 20)²⁷. Cluster 4 und 5 sind vielfach miteinander verbunden, einige Linien dieser Zonen erreichen auch Cluster 6. Dieser östliche Fundplatzbereich der Cluster 4–6 ist nie durch Zusammenpassungen mit der westlichen Zone der Cluster 2/3 verbunden.

Ein ähnliches Bild ergibt sich bei den Clustern I und II für die Gliederung der Peripherie (Abb. 21). Die Cluster I und II sind fundarm, die Masse der Funde liegt im zentralen Bereich (Cluster III und Residuum). Die vorhandenen Funde verbinden im Westen Cluster I mit den Clustern 2 und 3 und im Osten Cluster II mit dem Bereich der Cluster 4–6. Die sich für I und II ausschließenden Zusammenpassungen unterstreichen diese Trennung. Demnach entsprechen Cluster I und II den Zonen E und F der konventionellen Analyse und spiegeln jeweils fundarme Wegwerfzonen zweier unterschiedlicher Aktivitätszentren wider²⁸.

Bewertung

Wie der oben dargelegte Vergleich mit dem konventionell begründeten Ergebnis zeigt, ist die TGN-Gruppierung für Daten dieser Art ein gut geeignetes Verfahren. Ihre Anwendung ist sinnvoll und kann empfohlen werden.

Die Cluster- und Dienstprogramme des Bonner archäologischen Statistikpakets, Version 4.0

Die Programme, mit denen die hier vorgestellten Ergebnisse erzeugt wurden, sind in dem Bonner archäologischen Seriations- und Statistikpaket enthalten (Herzog u. Scollar 1987; 1990). Die Cluster- und Dienstprogramme bieten eine Fülle von Möglichkeiten an: Für Anwesenheits-/Abwesenheitsdaten werden die fünf Ähnlichkeitsmaße, für Häufigkeitsdaten die drei Abstandsmaße angeboten, die in Abb. 1 und 2 aufgeführt wurden. Außerdem kann man die Topographie, wie oben erwähnt, in die Gruppierung miteinbeziehen.

Bei diesem Verfahren muß nicht wie bei Complete-Linkage oder Average-Linkage eine Tabelle mit den Ähnlichkeitswerten für jedes Paar von Einheiten (Ähnlichkeitsmatrix) gespeichert werden; diese Tabelle wird

zwar berechnet, aber nur die k nächsten Nachbarn werden für jede Einheit gespeichert, wobei k die vom Benutzer gewählte Zahl der betrachteten Nachbarn ist. Speicherplatzbedarf und Rechenzeit wachsen mit der Zahl k . Für einen recht großen Testdatensatz mit 610 Gräbern, 231 Typen und 1848 Objekten betrug der Speicherbedarf bei 15 Nachbarn 77826 bytes, die Rechenzeit bei Verwendung des Jaccard-Ähnlichkeitskoeffizienten etwas mehr als 18 Minuten (für einen XT-kompatiblen Rechner mit 10 MHz). Ist ein Koprozessor vorhanden, verkürzt sich die Rechenzeit um mehr als die Hälfte. Die Gruppierungsprogramme sind so konzipiert, daß immer nur so viel Speicherplatz belegt wird, wie auch gebraucht wird. Die GN-Gruppierung ist zwar speicher- und rechenaufwendiger als eine effektiv programmierte Single-Linkage-Clusteranalyse, aber deutlich weniger aufwendig als Complete- oder Average-Linkage-Clustering.

Zur Darstellung der Clusterergebnisse stehen drei Programme zur Verfügung. Das erste erzeugt ein Dendrogramm, hat aber auch die Möglichkeit, das Dendrogramm auf einer vom Benutzer vorgegebenen Ebene zu schneiden, so daß die Dendrogramme der Cluster auf dieser Ebene erstellt werden. Ein weiteres Programm erstellt eine Tabelle, die angibt, welcher Typ in welchen Clustern vorhanden ist (ähnlich wie in Abb. 20–21), sowie für jedes Cluster ein Diagramm, das veranschaulicht, mit welcher Häufigkeit die Typen in diesem Cluster auftreten. Zum Vergleich ist jeweils auch die Gesamthäufigkeit des Typs in Klammern angegeben. Schließlich wird ein Programm angeboten, das die Clusterergebnisse topographischer Daten in Kartenform darstellt.

Alle drei Dienstprogramme erzeugen in irgendeiner Form eine graphische Ausgabe, die sich allein aus den Standardzeichen eines IBM-kompatiblen Druckers zusammensetzt. Dies reduziert den Grad der Auflösung, so daß ein größerer, stark strukturierter Plan nicht mehr detailgetreu auf einem DIN A4-Blatt darstellbar ist; doch ist das Programm in der Lage, große Pläne auf Papierbahnen zu erzeugen, die später aneinandergeliebt werden können. Bei den Symbolen, die die Cluster markieren, muß man sich auf den IBM-Zeichensatz beschränken. Aber die Beschränkung hat auch Vorteile: Jeder handelsübliche Drucker, der an einen IBM-PC oder kompatiblen angeschlossen ist, kann diese Graphiken ausgeben. Außerdem besteht die Möglichkeit, die Ausgabe nach Belieben zu editieren, z. B. indem man eine zusätzliche Überschrift einfügt oder die Graphik in einen Text integriert.

Literaturverzeichnis

Bosinski u. a. 1985: G. Bosinski, E. Cziesla u. a., Sprendlingen/Ein Fundplatz des mittleren Jungpaläolithikums in Rheinhausen. *Jahrb. RGZM* 32, 1985, 5–91.

²⁷ Die ebd. 66 Abb. 53 abgebildeten Zusammenpassungen mußten für unsere Analysen durchnumeriert werden; da hier jedoch Details kaum eine Rolle spielen, wird auf eine Korrelation von Plan und Numerierung verzichtet.

²⁸ Östlich des Zentrums zeichnet sich mit einer breiten Ausräumzone (Cluster 6) und einem peripheren Wegwerfereich (Cluster II) ein recht kontinuierliches Ausdünnen der zentralen Artefaktstreuung (Cluster 4/5) ab. Wir bezweifeln, daß diese Beobachtung mit der Rekonstruktion einer Behausung vereinbar ist, wie sie G. Bosinski vorschlägt (ebd. 88 Abb. 65).

- Cziesla 1987: E. Cziesla, L'analyse des raccords ou le concept du dynamisme en préhistoire. Bull. Soc. Préhist. Luxembourgeoise 9, 1987 (1988), 77–111.
- Dixon 1983: W. J. Dixon (Hrsg.), BMDP Statistical Software (University of California Press, Berkeley 1983) 459.
- Doran u. Hodson 1975: J. E. Doran u. F. R. Hodson, Mathematics and Computers in Archaeology (Edinburgh University Press 1975) 23–24; 136–143; 158; 173–177.
- Fremersdorf 1955: F. Fremersdorf, Das fränkische Reihengräberfeld von Köln-Müngersdorf. Germ. Denkmäler Völkerwanderungszeit 6 (Berlin 1955).
- Herzog u. Scollar 1987: I. Herzog u. I. Scollar, Ein ‚Werkzeugkasten‘ für Seriation und Clusteranalyse. Arch. Korrb. 17, 1987, 273–279.
- 1990: –, The Bonn Seriation and Archaeological Statistics Package Version 4.1 (Remagen 1990).
- Jarvis u. Patrick 1973: R. A. Jarvis u. E. A. Patrick, Clustering Using a Similarity Measure Based on Shared Near Neighbors. IEEE Transactions on Computers, Vol. C–22, No. 11, 1973, 1025–1034.
- Jarvis 1978: R. A. Jarvis, Shared Near Neighbor Maximal Spanning Trees for Cluster Analysis. Proceedings of the Fourth International Joint Conference on Pattern Recognition (Kyoto 1978) 308–313.
- Siegmund 1989: F. Siegmund, Fränkische Funde vom deutschen Niederrhein und der nördlichen Kölner Bucht (Phil. Diss. Köln 1989).
- Sokal u. Sneath 1973: R. R. Sokal u. P. H. A. Sneath, Numerical Taxonomy (W. H. Freeman and Company, San Francisco–London 1973) 121–135.

Irmela Herzog, Rheinisches Amt für Bodendenkmalpflege, Colmantstr. 14–16, D (W) 5300 Bonn 1

Frank Siegmund, Seminar für Ur- und Frühgeschichte, Universität, Nikolausberger Weg 15, D (W) 3400 Göttingen