

Sonderdruck aus

METHODS OF OPERATIONS RESEARCH

57

XI. SYMPOSIUM ON OPERATIONS RESEARCH

Organized by
Gesellschaft für Mathematik, Ökonomie und Operations
Research e.V.

University of Darmstadt, September 1-3, 1986

Proceedings

Edited by: Wolfgang Domschke, Darmstadt
 Werner Krabs, Darmstadt
 Jürgen Lehn, Darmstadt
 Peter Spellucci, Darmstadt

athenäum

EIN VERGLEICH ZWEIMODALER CLUSTERANALYSEVERFAHREN

M. Both/W. Gaul

Institut für Entscheidungstheorie und Unternehmensforschung
Universität Karlsruhe (TH)

Zusammenfassung:

Während einmodale Clusteranalyseverfahren wohlbekannt sind, beginnen zweimodale Techniken der Clusteranalyse, die sich als Operations Research - Optimierungsmodelle formulieren lassen, gerade erst, in den Bereich der Datenanalyse einzudringen. Beziehungen zwischen Elementen von Mengen verschiedener Modalität (Objekte, Eigenschaften, Subjekte, usw.), die durch Assoziationstabellen, Verwechslungstabellen, Kontingenztabellen, Fluktuationstabellen, usw., beschrieben werden können, sind Beispiele solcher zweimodaler Datensätze, die durch diese neuen zweimodalen Clusteranalyseverfahren ausgewertet werden können. Zwei eigne Techniken (eine hierarchische und eine nicht-hierarchische Version) werden vorgestellt und im Rahmen eines Marketing-Beispiels verglichen.

Summary:

Whereas one-mode clustering is well-known, two-mode clustering techniques which can be formulated as operations research optimization models are just entering the data analysis area. Relations between elements of sets of different modes (objects, properties, subjects, etc.) as described by association tables, confusion tables, contingency tables, fluctuation tables, etc. are examples of such two-mode data which can be evaluated by these new two-mode clustering procedures. Two own techniques (a hierarchical and a non-hierarchical version) are presented and compared on the basis of a marketing example.

1. Einleitung

Im Bereich der Datenanalyse gewinnen Verfahren zur Analyse mehrmodaler Daten an Bedeutung. Dabei gibt der Modus eines interessierenden Datensatzes die Anzahl der verschiedenen Mengen von Elementen an, die durch die Daten beschrieben werden. Beispiele aus dem Marketing für zweimodale Datensätze sind Assoziationsmatrizen, die etwa Beziehungen zwischen Produkten und relevanten Produkteigenschaften charakterisieren, Fluktuationmatrizen, die Veränderungen im Markenwahlverhalten von Konsumenten anzeigen, und Verwechslungsmatrizen, die etwa Fehlzuordnungen zwischen Werbebotschaften und Markenimages beschreiben. Zur Behandlung solcher Daten sind in der jüngsten Vergangenheit sowohl hierarchische (De Sarbo, De Soete (1984), De Soete et al. (1984), Espejo, Gaul (1986)) als auch nicht-hierarchische (De Sarbo (1982), Both, Gaul (1985)) Verfahren entwickelt worden.

Mit der Anzahl alternativ einsetzbarer Methoden wachsen aber auch die Schwierigkeiten, die sich bei der Auswahl der für eine bestimmte Problemstellung am besten geeigneten Methoden ergeben. Deshalb sind vergleichende Studien von Interesse, die unterschiedliche Verfahren im Hinblick auf die damit erzielten Ergebnisse (Güte der Anpassung, Interpretierbarkeit der Lösungen) gegenüberstellen. Für den einmodalen Fall vergleicht Eckes (1982) das nicht-hierarchische überlappende auf dem additiven Modell (siehe Shepard, Arabie (1979)) basierende Clusteranalyseverfahren MAPCLUS (Arabie, Carroll (1980)) mit einem gebräuchlichen hierarchischen Verfahren, für den zweimodalen Fall stellen Both, Gaul (1985) dem nicht-hierarchischen GENCLUS (De Sarbo (1982)) Modell, das auf Techniken der kombinatorischen Optimierung zurückgreift, ihr eigenes auf Penalty-Ansätzen basierendes nicht-hierarchisches Verfahren PENCLUS gegenüber, wohingegen Espejo, Gaul (1986) einen Vergleich von zweimodalen hierarchischen Clusteranalyseverfahren durchführen.

In der vorliegenden Arbeit werden zweimodale hierarchische und nicht-hierarchische, überlappende Methoden der Clusteranalyse anhand von Eigenentwicklungen (Both, Gaul (1985), Espejo, Gaul (1986)) miteinander verglichen. Dazu werden im zweiten Abschnitt die zugrundeliegenden Modelle vorgestellt und die verwendeten Lösungsansätze skizziert. Im Anschluß daran wird im dritten Abschnitt anhand von Problemstellungen aus dem Marketing aufgezeigt, inwiefern diese Verfahren zu einem besseren Verständnis der zu untersuchenden Sachverhalte beitragen können. Abschließend werden die festgestellten Gemeinsamkeiten und Unterschiede der beiden Ansätze zusammengefaßt und Hinweise darauf gegeben, wie sich die Vorzüge beider Verfahren verbinden lassen.

2. Zweimodale Clusteranalyse

Im folgenden werden neue Ansätze der zweimodalen hierarchischen und nicht-hierarchischen Clusteranalyse diskutiert und miteinander verglichen. Ein offensichtliches Vergleichskriterium ist dabei die durch die Verfahren erzeugte Klassenbildung.

2.1. Hierarchische Verfahren

Hierarchische Clusteranalyseverfahren, für die im einmodalen Fall bereits eine Vielzahl unterschiedlicher Varianten entwickelt worden sind, zählen zu den am weitesten verbreiteten Ansätzen in der Datenanalyse. Ihr Einsatz bietet sich an, wenn in den zu untersuchenden Daten hierarchische

Strukturen vermutet werden, da ein derartiger Sachverhalt mit den Bedingungen an die zu erzeugende Hierarchie von Klassen - je zwei Klassen sind entweder disjunkt oder die eine Klasse enthält die andere Klasse - korrespondiert. Derartige Beziehungen lassen sich auch im zweimodalen Fall anschaulich in Form eines Dendrogramms darstellen, das Auskunft darüber gibt, in welcher Reihenfolge einzelne Objekte bzw. Klassen zu größeren Einheiten zusammengefaßt werden.

Sei $S = (s_{ij})$ die gegebene zweimodale $(n \times m)$ -Matrix, die die Beziehungen zwischen den Elementen $i \in O_1 = \{1, \dots, n\}$ und $j \in O_2 = \{1, \dots, m\}$ beschreibt, wobei O_1, O_2 die beiden Mengen unterschiedlicher Modalität bezeichnen. Um die Verfahren anwenden zu können, muß S zunächst in eine $(n \times m)$ -Distanzmatrix $D = (d_{ij})$ (falls S nicht schon Distanzmatrix ist) transformiert werden. Danach stellt sich das Problem, aus D eine sogenannte $((n+m) \times (n+m))$ -Ultrametrik-Matrix $U = (u_{i_1 i_2})$ zu erzeugen, die die "ultrametrische" Ungleichung

$$u_{i_1 i_2} \leq \max \{ u_{i_1 k}, u_{i_2 k} \}, \quad i_1, i_2, k \in \{1, \dots, n+m\} \quad (1)$$

erfüllt und die in den zweimodalen Matrizen S bzw. D zum Ausdruck gebrachten Beziehungen zwischen den Elementen der Mengen O_1, O_2 bestmöglich wiedergibt. Die Bedingung (1) ist notwendig und hinreichend für die Repräsentation der in U enthaltenen Information in Form eines eindeutig bestimmten Dendrogramms. Fig. 1 zeigt verschiedene Möglichkeiten, wie man eine Ultrametrik-Matrix U erzeugen kann.

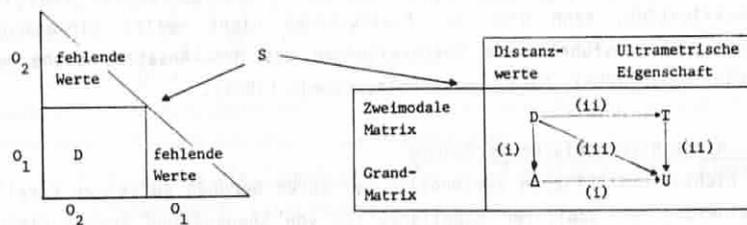


Fig. 1: Möglichkeiten zur Erzeugung einer Ultrametrik-Matrix im zweimodalen Fall

(i) Ausgehend von S kann D als bekannter Teil einer symmetrischen, sogenannten $((n+m) \times (n+m))$ -Grand-Matrix $\Delta = (\delta_{i_1 i_2})$ mit fehlenden Werten, die

zuerst geeignet zu bestimmen sind, aufgefaßt werden. Zu der so erzeugten $((n+m) \times (n+m))$ -Matrix Δ wird dann eine bestmögliche Ultrametrik-Matrix U mittels eines einmodalen Verfahrens gesucht.

(ii) Ausgehend von S kann D zur Konstruktion einer $(n \times m)$ -Matrix $T = (t_{ij})$ benutzt werden, die die "zwei Klassen ultrametrische" Ungleichung

$$t_{ij} \leq \max \{t_{ik}, t_{lk}, t_{lj}\}, \quad i, l \in O_1, j, k \in O_2 \quad (2)$$

erfüllt. Die Bedingung (2) ist notwendig und hinreichend für die Repräsentation der in T enthaltenen Information (unter Berücksichtigung einer speziellen Erweiterung von T zu U) in Form eines eindeutig (bis auf die interne Struktur von einmodalen Sub-Dendrogrammen) bestimmten Dendrogramms. Diese Methode benutzen De Soete et al. (1984) in ihrem hier 2MLS (2 Mode Least Squares) genannten Penalty-Ansatz.

(iii) Man kann auch direkt von D ausgehend versuchen, eine bestmögliche Ultrametrik-Matrix U zu bestimmen. De Soete (1984) braucht noch eine $((n+m) \times (n+m))$ -Start-Matrix zur Initialisierung seines hier 1MLS (1 Mode Least Squares) genannten Penalty-Ansatzes während Espejo, Gaul (1986) die gesuchte Ultrametrik-Matrix über eine fehlende Werte berücksichtigende Version MVAL (Missing Values Average Linkage) des Average-Linkage Verfahrens berechnen.

Auf die 1MLS, 2MLS Verfahren, die eine Kleinste-Quadrate-Zielfunktion mit einem Penalty-Ansatz verbinden, der die Nebenbedingungen (1) bzw. (2) berücksichtigt, kann hier aus Platzgründen nicht weiter eingegangen werden; für ausführlichere Beschreibungen des MVAL-Ansatzes siehe man Espejo, Gaul (1986), Gaul, Böckenholt, Espejo (1986).

2.2. Nicht-hierarchische Verfahren

Die nicht-hierarchischen zweimodalen Verfahren beruhen auf einer Verallgemeinerung des additiven Modellansatzes von Shepard und Arabie (1979) und lassen überlappende Klassenbildung zu. Dieser Möglichkeit nicht-disjunkter Klassifikationen kommt z.B. bei Positionierungs- und Segmentierungsfragestellungen im Marketing besondere Bedeutung zu (Arabie et al. (1981)), denn es kann durchaus erwünscht und sinnvoll sein, daß Objekte bzw. Subjekte mehreren Klassen gleichzeitig angehören, ohne daß diese Klassen - wie in hierarchischen Ansätzen vorausgesetzt - ineinander enthalten sind.

Obgleich einige Verfahren zur Generierung überlappender Clusterkonfigurationen schon seit längerer Zeit bekannt sind (siehe z. B. die in Eckes (1982) zitierte Literatur), wurde diese Art der Clusteranalyse erst mit dem von Shepard, Arabie (1979) entwickelten additiven Modell, das als Spezialfall des Kontrastmodells von Tversky (1977) aufgefaßt werden kann, in den Blickpunkt des Interesses gerückt. Das wesentliche Kennzeichen dieser Vorgehensweise besteht darin, daß die Ähnlichkeit zwischen zwei Objekten als ein additives Maß ihrer gemeinsamen Charakteristika angesehen wird. Demnach wird die durch das Modell rekonstruierte Ähnlichkeit zwischen zwei Objekten als Summe der Maße (Gewichte) all jener Cluster dargestellt, die beide Objekte als Elemente enthalten. Für den einmodalen Fall wurde von Arabie, Carroll (1980) das auf dem ALS (Alternating-Least-Squares)-Prinzip basierende Lösungsverfahren MAPCLUS vorgeschlagen, das auf Penalty-Ansätze zurückgreift. Eine Erweiterung des additiven Modells zur Behandlung zweimodaler Fragestellungen wurde von De Sarbo (1982) vorgenommen, wobei das von ihm eingesetzte Lösungsverfahren GENCLUS Ansätze aus der kombinatorischen Optimierung zur Bestimmung der die Klassenzugehörigkeit bezeichnenden Parameter verwendet. Dagegen werden im von Both, Gaul (1985) beschriebenen zweimodalen Clusteranalyseverfahren PENCLUS Penalty-Ansätze zur Berechnung der Klassenzugehörigkeitsparameter und zur Einhaltung von Nebenbedingungen an die resultierenden Konfigurationen herangezogen. Die dort eingeschlagene Vorgehensweise wird im folgenden kurz beschrieben.

Für eine gegebene zweimodale $(n \times m)$ -Ähnlichkeitsmatrix $S = (s_{ij})$ wird eine $(n \times m)$ -Matrix $\tilde{S} = (\tilde{s}_{ij})$ der Form

$$\tilde{S} = P W Q' + C \quad (3)$$

gesucht, die die aufsummierten quadratischen Abweichungen zwischen den Elementen von S und \tilde{S} minimiert. Dabei bezeichnet

$P = (p_{i1})$ eine die Zugehörigkeit der n Zeilenobjekte zu k Klassen wiedergebende $(n \times k)$ -Matrix,

$W = (w_{11}^*)$ eine $(k \times k)$ -Gewichtungsmatrix,

$Q = (q_{j1}^*)$ eine die Zugehörigkeit der m Spaltenobjekte zu k Klassen wiedergebende $(m \times k)$ -Matrix,

$C = (c)$ eine $(n \times m)$ -Matrix, deren Elemente alle gleich sind und die als additive Konstante interpretiert werden können.

Dabei müssen die Elemente der Matrizen P und Q , die die Klassenzugehörigkeit spezifizieren, binär sein:

$$\begin{array}{lll}
 p_{il} \in \{0,1\}, & i = 1, \dots, n, & l = 1, \dots, k, \\
 q_{jl}^* \in \{0,1\}, & j = 1, \dots, m, & l^* = 1, \dots, k.
 \end{array} \quad (4)$$

Weiterhin können zusätzliche Nebenbedingungen spezifiziert werden, die a-priori vorhandene Kenntnisse der zu untersuchenden Sachverhalte oder angestrebte Eigenschaften der Clusterkonfiguration berücksichtigen; der gewünschte Klassifikationstyp kann beispielsweise wie folgt angegeben werden:

$$\sum_{l=1}^k p_{il} \left\{ \begin{array}{l} = \\ \geq \end{array} \right\} 1, \quad i = 1, \dots, n, \quad (5(a))$$

$$\sum_{l^*=1}^k q_{jl}^* \left\{ \begin{array}{l} = \\ \geq \end{array} \right\} 1, \quad j = 1, \dots, m. \quad (5(b))$$

Die Bedingung (5(a)) stellt sicher, daß nur nichtüberlappende Cluster gebildet werden, während (5(b)) den allgemeineren Fall überlappender Cluster zuläßt. Das zu behandelnde restringierte Optimierungsproblem läßt sich in seiner einfachsten Form wie folgt darstellen:

$$\begin{aligned}
 \text{Min } Z &= \sum_{i=1}^n \sum_{j=1}^m (s_{ij} - \bar{s}_{ij})^2 \\
 &= \sum_{i=1}^n \sum_{j=1}^m (s_{ij} - \sum_{l=1}^k \sum_{l^*=1}^k p_{il} w_{ll}^* q_{l^*j} - c)^2
 \end{aligned} \quad (6)$$

bzgl. P , W , Q und C unter Einhaltung von Nebenbedingungen des Typs (4) und (5). Zur Lösung von (6) wird ein ALS-Ansatz mit einer iterativen Schätzung von Parameteranteilen eingesetzt. Dabei können die Matrizen W und C mittels Regression geschätzt werden, während für die Berechnung der Matrizen P und Q das restringierte Optimierungsproblem (6) unter Verwendung von Penaltyfunktionen (siehe Gaul, Hartung (1979), Rao (1978)) in eine dazu äquivalente Folge unbeschränkter Optimierungsprobleme, die mit Hilfe eines konjugierten Gradientenverfahrens gelöst werden, transformiert wird. Dazu werden die Nebenbedingungen (4) und (5) wie folgt modifiziert: Aus Bedingung (4) für die einzelnen Elemente der Matrix P erhält man die Binaritäts-Bedingung

$$BB(P) = \frac{1}{n \cdot k} \sum_{i=1}^k \sum_{l=1}^k (p_{il}(1-p_{il}))^2 = 0 \quad (7)$$

während aus (5(a)) die Überlappungs-Bedingung

$$OB(P) = \frac{1}{n} \sum_{i=1}^n (\max \{ 0, 1 - \sum_{l=1}^k p_{il} \})^2 \quad (8)$$

resultiert. Für den nichtüberlappenden Fall wird $\frac{1}{n} \sum_{i=1}^n (\sum_{l=1}^k p_{il} - 1)^2 = 0$ verwandt. Entsprechendes gilt bei Zugrundelegung der Matrix Q.

Das bedingte Optimierungsproblem bzgl. P- die anderen Parameterteilmengen werden festgehalten - hat somit für den überlappenden Fall folgende Gestalt (die Notation für die Matrix Q erfolgt analog):

$$\text{Min } Z_{\text{pen}}(P) = Z(P) + \rho^{IT} (BB(P) + OB(P))$$

wobei Z_{pen} die Zielfunktion des Penalty-Ansatzes ist und ρ^{IT} den Penalty-Parameter im jeweiligen Iterationsschritt bezeichnet. Zur Lösung dieses nichtrestringierten Optimierungsproblems werden folgende partiellen Ableitungen benötigt:

$$\frac{\partial Z}{\partial p_{ab}} = -2 \sum_{j=1}^m [(s_{aj} - \sum_{l=1}^k \sum_{l=1}^k p_{al} w_{1l} * q_{1*j} - c) (\sum_{l=1}^k w_{b1} * q_{1*j})],$$

$$\frac{\partial BB}{\partial p_{ab}} = \frac{2}{n \cdot k} p_{ab} (1 - 3 p_{ab} + 2 p_{ab}^2),$$

$$\frac{\partial OB}{\partial p_{ab}} = -\frac{2}{n} \max \{ 0, 1 - \sum_{l=1}^k p_{al} \}.$$

f weitere Verfahrensaspkte kann im Rahmen der vorliegenden Arbeit cht eingegangen werden; eine detaillierte Beschreibung dieser Gesichtspunkte wird in Both, Gaul (1985) gegeben.

Marketing-Beispiel und Verfahrensvergleich

die im vorangehenden Kapitel beschriebenen unterschiedlichen Vorgehensweisen der Clusteranalyse zu verdeutlichen und die beiden vorgestellten Ansätze miteinander zu vergleichen, werden im folgenden aus einem Experiment über die Anzeigenwerbung für Zigarettenmarken gewonnene zwei Datensätze mit Hilfe jeweils eines hierarchischen und eines nicht-hierarchischen Verfahrens ausgewertet.

Bei der Vermarktung von Imagery-Produkten wie z. B. Zigaretten und Kosmetika, die sich bezüglich ihrer Produkteigenschaften häufig nur geringfügig unterscheiden, wird bekanntlich der Werbung eine herausragende Bedeutung beigemessen. Denn ihr kommt die Aufgabe zu, ein Markenimage aufzubauen bzw. zu pflegen, das als Bündel von Begriffen und Assoziationen aufgefaßt werden kann, welches der Marke im Wahrnehmungsraum der Konsumenten eine eigenständige, nach Möglichkeit von der Konkurrenz abgegrenzte Position verschaffen soll. In diesem Zusammenhang sind u.a. folgende Fragestellungen von Interesse: "Welche Wahrnehmungsdimensionen werden mit den Marken und mit den zugehörigen Anzeigen in Verbindung gebracht?", "Entsprechen die vermittelten Werbebotschaften dem vorhandenen Markenimage?", "Welche Bedeutung haben Werbeslogans bei der Beurteilung von Anzeigen bzw. für die Wiedererkennung von beworbenen Marken?". Zur Beantwortung derartiger Fragestellungen wurde ein Experiment durchgeführt, dessen Aufbau nachfolgend skizziert wird (Eine ausführlichere Beschreibung dieses Experiments wird in Both, Gaul (1985) gegeben.).

Für acht Zigarettenmarken (Camel, HB, Krone, Lord, Marlboro, R6, Stuyvesant, West), die einem Anteil von ca. 61% des deutschen Marktes abdecken, wurden jeweils zwei Werbeanzeigen zur Präsentation ausgewählt. Vier weitere Marken (Dunhill, Ernte, Kim, Philip Morris) mit einem Marktanteil von insgesamt 7% wurden ohne Werbeanzeigen hinzugenommen. Es wurden zwei auf unterschiedliche Weise maskierte Anzeigenserien verwendet: Bei einer Serie wurden lediglich die Markennamen und Zigarettenpackungen verdeckt, während bei der anderen Serie zusätzlich die Slogans entfernt wurden.

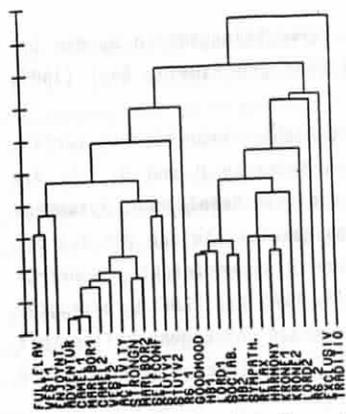
Während des Experiments mußten zunächst alle zwölf Marken unter Verwendung von Ratingskalen im Hinblick auf die Eigenschaften "voller Geschmack", "Abenteuer", "gute Laune", "Aktivität", "Sympathie", "Entspannung", "Geselligkeit", "Genuß", "Harmonie", "Härte", "Exklusivität", "Selbstsicherheit" und "Tradition" beurteilt werden. Anschließend wurden den in zwei Gruppen aufgeteilten Teilnehmern - 126 Studenten aus Marketing- und BWL-Vorlesungen - kurz die markierten Werbeanzeigen vorgeführt, wobei eine Zuordnung der Anzeigen zu den vorgegebenen Marken zu treffen war. Zum Schluß wurden die Anzeigen ein zweites Mal präsentiert; dabei mußten von den Probanden Beurteilungen der Anzeigen bezüglich der oben angegebenen Eigenschaften abgegeben werden. Die auf diese Weise erhobenen Daten wurden als Assoziationsmatrizen (Marken/Anzeigen und Eigenschaften) sowie als Verwechslungsmatrizen (Anzeigen und Marken) angeordnet und unterschiedlich (mit Slogan, ohne Slogan, Männer, Frauen, Raucher, Nicht-raucher) aggregiert. Im weiteren werden einige Ergebnisse aus der Analyse

der Assoziationsmatrizen vorgestellt (Die Verwechslungsdaten wurden bereits von Böckenholt (1986), Both, Gaul (1985) und Espejo, Gaul (1986) ausgewertet.).

Zum Vergleich der mit den hierarchischen und nicht-hierarchischen Verfahren erzielten Ergebnisse werden die beiden Matrizen P und Q, die die Klassenzugehörigkeit der Zeilen- bzw. Spaltenobjekte bezeichnen, zusammengefaßt und unter die jeweiligen Dendrogramme gesetzt, so daß die aus dem nicht-hierarchischen Verfahren resultierenden Cluster leicht rekonstruiert werden können. Fig. 2 stellt die mittels MVAL und PENCLUS erzeugten Konfigurationen für die Assoziationen von Anzeigen und Eigenschaften (mit Slogans bzw. ohne Slogans) sowie für die Assoziationen von Marken und Eigenschaften (Männer bzw. Frauen) gegenüber.

Das iterative PENCLUS-Verfahren wurde hierfür mit Klassenzugehörigkeitsmatrizen P und Q initialisiert, die auf mit hierarchischen Verfahren erzielten Resultaten basieren. Bei den nachfolgend vorgestellten Ergebnissen, die mit einer symmetrischen Gewichtungsmatrix W gewonnen wurden, ergaben sich keine überlappenden Cluster, da die zwischen den Klassen bestehenden Beziehungen durch die Nebendiagonalelemente von W zum Ausdruck gebracht werden können. Mit einer diagonalen Gewichtungsmatrix wurden dagegen überlappende Cluster erzeugt, die jedoch eine geringere Anpassungsgüte aufweisen. Dieser Sachverhalt liegt darin begründet, daß die hier untersuchten Assoziationsdaten im Gegensatz zu den Verwechslungsdaten (siehe Both, Gaul (1985)) keine Diagonalstruktur aufweisen, weshalb eine Modellspezifikation mittels diagonalen Gewichtungsmatrix hier weniger gut geeignet ist.

Der Vergleich zwischen den mit MVAL und PENCLUS erzielten Ergebnissen erfolgt zum einen an Hand der Klassenzugehörigkeit der einzelnen Elemente und zum anderen unter Verwendung von Maßen zur Beurteilung der erreichten Anpassungsgüte. In Tab. 1 sind der hierfür herangezogene Theil'sche Inequality Coefficient (TIC) und die Variance Accounted For (VAF) gemeinsam mit den zugehörigen Gewichtungsmatrizen W zusammengestellt. Man erkennt Verbesserungen beider Gütemaße - wie es sein muß - bei einer Erhöhung der Klassenanzahl von drei auf vier Klassen für das PENCLUS-Verfahren. Die mit der hierarchischen MVAL-Prozedur erzeugten Konfigurationen weisen einen höheren Anteil erklärter Varianz auf, während die TIC-Werte für die PENCLUS-Lösungen besser sind, so daß bezüglich der Anpassungsgüte keine eindeutige Oberlegenheit eines der beiden Verfahren erkennbar ist.

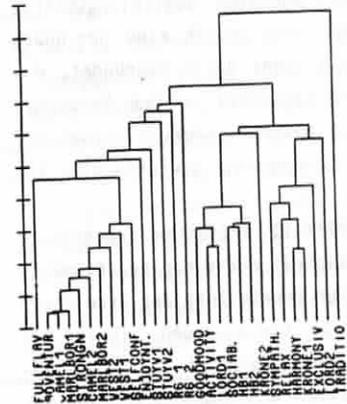


Anzeigen-Eigenschaften (mit Slogans)

1	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
2	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
3	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X

1	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
2	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
3	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
4	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X

Fig. 2(a)

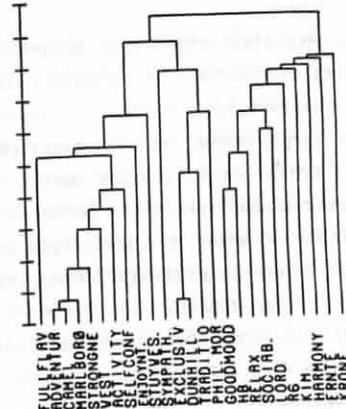


Anzeigen-Eigenschaften (ohne Slogans)

1	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
2	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
3	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X

1	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
2	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
3	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
4	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X

Fig. 2(b)

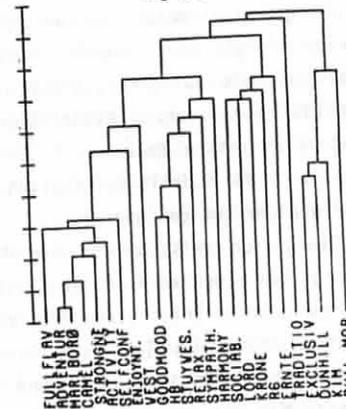


Marken-Eigenschaften (Männer)

1	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
2	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
3	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X

1	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
2	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
3	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
4	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X

Fig 2(c)



Marken-Eigenschaften (Frauen)

1	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
2	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
3	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X

1	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
2	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
3	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
4	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X

Fig. 2(d)

Fig. 2: Zweimodale Clusteranalyse der Assoziationsdaten mittels MVAL und PENCLUS

Anzeigen-Eigenschaften (mit Slogans)					Marken-Eigenschaften (Männer)				
Verfahren	Kl.	TIC	VAF	W	Verfahren	Kl.	TIC	VAF	W
PENCLUS	3	0.098	0.494	0.77 0.60 0.42 0.60 0.56 0.43 0.42 0.43 0.58	PENCLUS	3	0.089	0.518	0.64 0.40 0.38 0.40 0.59 0.40 0.38 0.40 0.47
	4	0.093	0.544	0.77 0.60 0.46 0.40 0.60 0.56 0.43 0.43 0.46 0.43 0.74 0.54 0.40 0.43 0.54 0.57		4	0.083	0.580	0.64 0.42 0.40 0.35 0.42 0.56 0.44 0.39 0.40 0.44 0.55 0.46 0.35 0.39 0.46 0.42
MVAL		0.112	0.691		MVAL		0.095	0.639	
Anzeigen-Eigenschaften (ohne Slogans)					Marken-Eigenschaften (Frauen)				
Verfahren	Kl.	TIC	VAF	W	Verfahren	Kl.	TIC	VAF	W
PENCLUS	3	0.108	0.503	0.69 0.48 0.39 0.48 0.38 0.41 0.39 0.42 0.56	PENCLUS	3	0.116	0.387	0.66 0.36 0.40 0.36 0.57 0.40 0.40 0.40 0.46
	4	0.101	0.555	0.68 0.47 0.44 0.39 0.47 0.34 0.31 0.38 0.44 0.31 0.69 0.52 0.39 0.38 0.52 0.57		4	0.110	0.460	0.64 0.37 0.45 0.36 0.37 0.57 0.41 0.35 0.45 0.41 0.55 0.44 0.36 0.35 0.44 0.42
MVAL		0.120	0.583		MVAL		0.114	0.516	

Tab. 1 Gütemaße und Gewichtungsmatrizen für PENCLUS und MVAL

Vergleicht man den MVAL- und den PENCLUS-Ansatz hinsichtlich der gebildeten Cluster, so lassen sich folgende Feststellungen treffen, die am Beispiel der Analyse der Assoziationen von Anzeigen und Eigenschaften für die Versuchsgruppe mit Slogans (Fig. 2(a)) illustriert werden. Bei einer vorgegebenen Anzahl von drei Klassen enthält das Cluster mit dem höchsten Gewicht beide Anzeigen für Camel und Marlboro sowie die Anzeige West 2, die am stärksten mit den Eigenschaften "Abenteuer", "Aktivität" und "Härte" assoziiert werden. Das zweite Cluster bilden beide Anzeigen für Stuyvesant und R6 sowie die Anzeige West 1, die am ehesten mit den Eigenschaften "voller Geschmack", "Selbstvertrauen" und "Genuß" in Verbindung gebracht werden. Das dritte Cluster mit den restlichen Anzeigen und Eigenschaften wird bei einer Erhöhung der Klassenanzahl aufgespalten. Dabei erhält das Cluster mit beiden Anzeigen für HB sowie der Anzeige Lord 1 und den Eigenschaften "gute Laune" und "Geselligkeit" ein hohes Gewicht, was auf eine vergleichsweise große Homogenität schließen läßt. Somit können die wesentlichen, in den Assoziationsdaten enthaltenen Strukturen bereits mit Hilfe weniger mittels PENCLUS erzeugter Cluster aufgedeckt werden. Eine vergleichbare Gruppierung von Anzeigen und

Eigenschaften kann dagegen mit einem Cut im mit MVAL generierten Dendrogramm nicht erlangt werden. Denn es zeigt sich, daß die Eigenschaften "Tradition" und "Exklusivität", die für eine Charakterisierung der Anzeigen am wenigsten geeignet erscheinen, bei einer hierarchischen Analyse eigenständige Cluster bilden und erst auf einem hohen Niveau mit den zuvor gebildeten Klassen fusioniert werden. Während bei PENCLUS die zwischen den Klassen bestehenden Beziehungen durch die Nebendialogelemente der Gewichtungsmatrix W zum Ausdruck gebracht werden, geben die Dendrogramme Auskunft über die Strukturen innerhalb der Klassen. Hiermit wird im einzelnen erkennbar, welche Anzeigen besonders stark mit den jeweiligen Eigenschaften assoziiert werden und in welcher Reihenfolge die einzelnen Elemente zu größeren Einheiten zusammengefaßt werden.

Betrachtet man die für die Beurteilung der Marken erlangten Konfigurationen (Fig. 2(c), 2(d)), so soll hier nur auf folgenden Unterschied in der Struktur zwischen den Assoziationsdaten für Anzeigen und Eigenschaften sowie Marken und Eigenschaften hingewiesen werden. Es wird von allen Verfahren jeweils ein Cluster gebildet, das die Marken Dunhill und Philip Morris, für die keine Werbeanzeigen präsentiert wurden, zusammen mit der Eigenschaft "Exklusivität" enthält. Bei den weiblichen Versuchspersonen wird im Gegensatz zu den Männern in dieses Cluster noch die Frauenzigarette Kim aufgenommen (Fig. 2(d)), was auf unterschiedliche geschlechtsspezifische Assoziationen hinweist. Während bei den Frauen diese Clusterbildung von beiden Verfahren einheitlich durchgeführt wird, erkennt man bei den Männern (Fig. 2(c)) einen Unterschied bzgl. der (Nicht-)Hinznahme der Marke Stuyvesant in das entsprechende Cluster, dem hier auch noch die Eigenschaft "Tradition" zugeordnet wurde.

Zusammenfassend läßt sich festhalten, daß die mittels MVAL und PENCLUS erzeugten diskreten Repräsentationen der Assoziationsdaten die wichtigen Strukturen gut erkennen lassen, und daß auf dieser Grundlage Hinweise für die Ausgestaltung von Werbebotschaften gegeben werden können. Der Verfahrensvergleich hat gezeigt, daß diese Verfahren Informationen z.B. über die relevanten Cluster, die Strukturen innerhalb von Klassen und die Beziehungen zwischen Klassen liefern, die einander sinnvoll ergänzen. Für eine Clusteranalyse zweimodaler Datensätze wird daher empfohlen, beide Verfahren gemeinsam in der beschriebenen Weise einzusetzen, um somit die Vorteile des hierarchischen und nichthierarchischen Ansatzes miteinander zu verbinden.

Literatur:

- Arabie, P., Carroll, J. D. (1980): MAPCLUS: A Mathematical Programming Approach to Fitting the ADCLUS Model, *Psychometrika*, 45, 211-235.
- Arabie, P., Carroll, J. D., De Sarbo, W., Wind, J. (1981): Overlapping Clustering: A New Method for Product Positioning, *Journal of Marketing Research*, 18, 310 - 317.
- Böckenholt, I. (1986): Graphic Representation of Categorical Proximities: A Maximum Likelihood Method, in: Gaul, W., Schader, M. (Eds.): *Classification as a Tool of Research*, North Holland, Amsterdam, New York, Oxford, Tokyo.
- Both, M., Gaul, W. (1985): PENCLUS (Penalty Clustering) for Marketing Applications, Discussion Paper No. 82, Institute of Decision Theory and Operations Research, University of Karlsruhe.
- De Sarbo, W. S. (1982): GENCLUS: New Models for General Nonhierarchical Clustering Analysis, *Psychometrika*, 47, 449-475.
- De Sarbo, W. S., De Soete, G. (1984): On the Use of Hierarchical Clustering for the Analysis of Nonsymmetrical Proximities, *Journal of Consumer Research*, 11, 601-610.
- De Soete, G. (1984): Ultrametric Tree Representation of Incomplete Dissimilarity Data, *Journal of Classification*, 1, 235-242.
- De Soete, G., De Sarbo, W. S., Furnas, G. W., Carroll, J. D. (1984): The Estimation of Ultrametric and Path Length Trees from Rectangular Proximity Data, *Psychometrika*, 49, 289-311.
- Eckes, T. (1982): Zwei diskrete Modelle zu Ähnlichkeitsstrukturen im Vergleich: Das hierarchische und das additive Cluster-Modell, *Archiv für Psychologie*, 134, 281-302.
- Espejo, E., Gaul, W. (1986): Two-Mode Hierarchical Clustering as an Instrument of Marketing Research, in: Gaul, W., Schader, M. (Eds.): *Classification as a Tool of Research*, North Holland, Amsterdam, New York, Oxford, Tokyo.
- Gaul, W., Böckenholt, I., Espejo, E. (1986): Hierarchical Clustering with Missing Values in One- and Two-Mode Proximity Data, Discussion Paper No. 94, Institute of Decision Theory and Operations Research, University of Karlsruhe.
- Gaul, W., Hartung, J. (1979): A Barrier Method with Arbitrary Starting Point, *Mathematische Operationsforschung und Statistik, Ser. Optimierung*, 10, 317-323.
- Rao, S. S. (1978): *Optimization*, Wiley Eastern Ltd., New Dehli, Bangalore, Bombay.
- Shepard, R. N., Arabie, P. (1979): Additive Clustering: Representation of Similarities as Combination of Discrete Overlapping Properties, *Psychological Review*, 86, 87-123.
- Tversky, A. (1977): Features of Similarity, *Psychological Review*, 84, 327- 352.

Martin Both
 Institut für Entscheidungstheorie
 und Unternehmensforschung
 Universität Karlsruhe (TH)
 Postfach 63 80
 D-7500 Karlsruhe

Prof. Dr. Wolfgang Gaul
 Institut für Entscheidungstheorie
 und Unternehmensforschung
 Universität Karlsruhe (TH)
 Postfach 63 80
 D-7500 Karlsruhe