

Schwaiger, Manfred

Working Paper

Innovative Auswertungskonzepte für ungestützte Recalltests

Arbeitspapiere zur mathematischen Wirtschaftsforschung, No. 140

Provided in Cooperation with:

Universität Augsburg, Institut für Statistik und Mathematische Wirtschaftstheorie

Suggested Citation: Schwaiger, Manfred (1996) : Innovative Auswertungskonzepte für ungestützte Recalltests, Arbeitspapiere zur mathematischen Wirtschaftsforschung, No. 140, Universität Augsburg, Institut für Statistik und Mathematische Wirtschaftstheorie, Augsburg

This Version is available at:

<http://hdl.handle.net/10419/47645>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

**ARBEITSPAPIERE
ZUR
MATHEMATISCHEN WIRTSCHAFTSFORSCHUNG**

**Innovative Auswertungskonzepte
für ungestützte Recalltests**

M. Schwaiger

Heft 140/1996

**Institut für Statistik und Mathematische Wirtschaftstheorie
Universität Augsburg**

Dr. Manfred Schwaiger

Universitätsstr. 16, D-86135 Augsburg

Telefon: (0821) 598-4160

Telefax: (0821) 598-4226

e-mail: manfred.schwaiger@wiso.uni-augsburg.de

Innovative Auswertungskonzepte für ungestützte Recalltests

M. Schwaiger*

Kurzzusammenfassung:

Angesichts ständig zunehmender Werbeintensität ist die Aufmerksamkeitswirkung eines Werbemittels, die in der Praxis hauptsächlich mit Hilfe ungestützter Recalltests gemessen wird, ein wesentliches Qualitätsmerkmal. Die bisher gängigen Auswertungsmethoden beschränken sich auf die (univariate) Berechnung von Recallwerten, die über Referenzwerte mit konkurrierenden Werbemitteln verglichen werden. Der vorliegende Artikel zeigt, wie Recall-Daten durch Anwendung multivariater Verfahren, insbesondere zweimodaler Klassifikationsverfahren, informativer ausgewertet werden können. Der Informationsgewinn entsteht durch die Berücksichtigung von Ähnlichkeitsbeziehungen zwischen Werbemitteln und durch Visualisierung von Zusammenhängen, die bei univariater Auswertung verborgen bleiben.

1 Zur Kontrolle der Aufmerksamkeitswirkung von Werbemitteln

Die Werbeinvestitionen entsprechen im Jahr 1996 mit 56 Mrd. DM voraussichtlich etwas mehr als 1,55% des Bruttoinlandsproduktes in Deutschland (vgl. ZAW 1996, S. 10). Diese Zahl verdeutlicht, daß Unternehmen Werbung als einen wesentlichen Erfolgsfaktor der Wettbewerbswirtschaft auffassen. Unterstützt wird diese Haltung durch die aktuellen Marktbedingungen - insbesondere im Konsumgüterbereich. Gesättigte Märkte, Verdrängungswettbewerb, technisch weitgehend ausgereifte Produkte mit kaum erkennbaren Qualitätsunterschieden und ohne wirklich innovative Eigenschaften führen dazu, daß immer mehr Marktkommunikation betrieben werden muß, um auf Produkte aufmerksam zu machen und Konsumenten als Kunden zu gewinnen.

Empirischen Untersuchungen zufolge betrug die gesamtgesellschaftliche Informationsüberlastung, d.h. der Prozentsatz angebotener Information, der den Empfänger nicht erreicht, bereits 1987 über 98% (vgl. *Kroeber-Riel* 1987, S. 485ff.). Zwar sind in diesem Anteil auch die Streuverluste enthalten, also jener Teil an Informationen,

*PD Dr. Manfred Schwaiger, Universität Augsburg, Lehrstuhl für Mathematische Methoden der Wirtschaftswissenschaften, Universitätsstr. 16, D-86135 Augsburg, Tel.: 0821/598-4160, Fax 0821/598-4226

der an ein breites Publikum weitergegeben wird, aber nur für eine spezielle Zielgruppe bestimmt ist. Streuverluste sind aber nicht der ausschlaggebende Grund für den "information overload"; Tatsache ist vielmehr, daß selbst die Zielgruppe die für sie bestimmten Informationen nicht vollständig nutzt. Die Informationsüberlastung wird in den kommenden Jahren zusätzlich wachsen, weil neue Anbieter in den Markt eintreten und durch technischen Fortschritt neue Medien zur Informationsübertragung einsetzbar werden. Die Informationsaufnahmefähigkeit der Empfänger kann dagegen aufgrund biologischer Restriktionen nicht nennenswert gesteigert werden.

Werbung muß folglich Aufmerksamkeit erregen, damit sie sich durchsetzen und ins Bewußtsein oder Unterbewußtsein des Konsumenten vordringen kann. Der Kontrolle der Aufmerksamkeitswirkung wird daher in der einschlägigen Literatur zur Konsumentenforschung (vgl. hierzu bspw. *Kroeber-Riel* 1990 und 1992) ein erheblicher Stellenwert zugebilligt. Ungestützte Recalltests, die oft fälschlicherweise zur Ermittlung der Kommunikationsleistung eingesetzt und dann zu Recht als nicht valide kritisiert wurden, eignen sich zur Analyse der Aufmerksamkeitswirkung im Sinne der Durchsetzungsfähigkeit eines Werbemittels (1). Bei der Durchführung von Recalltests wird durch Befragung festgestellt, an welche Werbemittel bzw. an welche Teile eines Werbemittels sich die Testpersonen erinnern. Charakteristisch für den ungestützten Recall ist, daß auf Erinnerungshilfen (Abbildungen, Markenzeichen, Auswahllisten etc.) verzichtet wird, mithin also aktive Gedächtnisinhalte der Probanden abgerufen werden. Aus Platzgründen soll an dieser Stelle nicht auf einzelne Varianten von ungestützten Recalltests eingegangen werden, der Leser wird auf die einschlägige Literatur verwiesen (vgl. z.B. *Rehorn* 1988). In der Praxis werden überwiegend Folder-Tests für Anzeigen bzw. Studiotests für TV- und Hörfunkspots eingesetzt. Beim Folder-Test wird den Probanden eine Mappe mit Anzeigen, ggf. ergänzt um redaktionelle Beiträge, vorgelegt. Nach Durchsicht und Rückgabe der Mappe werden die Versuchspersonen befragt, an welche Anzeigen sie sich erinnern. Der Prozentsatz der Befragten, die sich an eine bestimmte Anzeige erinnern, wird als Kennzahl für die relative Aufmerksamkeitswirkung der betreffenden Anzeige angesehen (vgl. *Rehorn* 1988, S. 27). In ähnlicher Weise wird beim Studiotest verfahren, nur werden hier an Stelle der Präsentation eines Folders mit Printanzeigen Sequenzen von TV- oder Hörfunkspots vorgeführt.

Gängige Praxis ist es bis dato, lediglich den Anteil der Personen, die sich an den Testspot oder die Testanzeige erinnern konnten, auszuweisen und mit kategoriespezifischen Referenzwerten zu vergleichen.

Um aussagefähige Ergebnisse im Hinblick auf die Durchsetzungsfähigkeit zu erhalten, ist es jedoch realitätsnäher, zumindest die wichtigsten Konkurrenz-Werbemittel in den Test mit aufzunehmen, und damit einer relativen Beurteilung eines Werbemittels den Vorzug gegenüber einer absoluten einzuräumen. Ferner erscheint es sinnvoll, nicht einen einzelnen Recallwert für die gesamte Zielgruppe zu berechnen, sondern die segmentspezifische Durchsetzungsfähigkeit zu ermitteln.

Ziel dieser Arbeit ist es zu zeigen, wie auf ungestützten Recalltests basierende Ergebnisse durch Anwendung multivariater Verfahren, insbesondere zweimodaler Clusteranalysen, informativer ausgewertet werden können. Durch die Berücksichtigung von Ähnlichkeitsbeziehungen zwischen konkurrierenden Werbemitteln und durch Visualisierung von strukturellen Zusammenhängen, die bei univariater Auswertung verborgen bleiben, gelingt es, einen höheren Anteil der Ursprungsinformation zu konservieren.

2 Grundlagen der zweimodalen Klassifikation

Aufgabe der Klassifikationsverfahren allgemein ist die Zusammenfassung von Elementen zu Klassen, so daß innerhalb der Klassen möglichst große Ähnlichkeit und zwischen den Klassen möglichst große Unähnlichkeit besteht (vgl. *Opitz* 1980, S. 65ff.). Eine bestimmte Menge von Elementen wird dabei als *Modus* bezeichnet.

Elemente eines speziellen Modi können Objekte (z.B. Werbemittel), Merkmale (z.B. Items zur Beschreibung von Objekten), Meßbedingungen oder Zeitpunkte sein. Sollen Elemente verschiedener Modi, also beispielsweise Objekte *und* Merkmale, in einem Cluster zusammengefaßt werden, so wendet man Verfahren der multimodalen, andernfalls Verfahren der einmodalen Klassifikation an.

In der hier angewandten, zweimodalen Clusteranalyse werden Zeilen- und Spaltenelemente einer Datenmatrix simultan klassifiziert. Als Input wird eine Datenmatrix \mathbf{X} der Form

$$\mathbf{X} = (x_{ij})_{(n \times m)} = \begin{pmatrix} x_{11} & \cdots & x_{1m} \\ \vdots & \ddots & \vdots \\ x_{n1} & \cdots & x_{nm} \end{pmatrix}, \quad i = 1, \dots, n, \quad j = 1, \dots, m$$

benötigt, wobei der Index i die Zeilenelemente und der Index j die Spaltenelemente kennzeichnet. Im vorliegenden Fall wird x_{ij} den Anteil der Versuchspersonen aus Zielgruppensegment i kennzeichnen, die sich an Werbemittel j erinnerten.

Die allgemeinen Vorteile des Einsatzes zweimodaler Clusteranalysen werden durch den Vergleich mit einmodalen Verfahren deutlich: eine einmodale Klassifikation liefert idealerweise mehrere in sich homogene, untereinander jedoch heterogene Cluster von Elementen eines Modi, in diesem Fall von Zielgruppensegmenten. Man weiß, daß in einem Cluster zusammengefaßte Zielgruppensegmente im Hinblick auf die Struktur der erzielten Recallwerte ähnlich sind; wie ein bestimmter Cluster zu beschreiben ist, wie er sich von anderen Clustern unterscheidet, wird allerdings erst durch die erneute Analyse der Ausgangsdaten klar.

Hier stellt die zweimodale Klassifikation eine hervorragende Interpretationshilfe zur Verfügung: dadurch, daß Zielgruppensegmente und Werbemittel simultan klassifiziert werden, läßt sich ein Cluster ohne Rückgriff auf Einzelwerte in der Datenbasis und ohne Berechnung von Lage- und Streuungsparametern sofort treffsicher beschreiben, weil die Werbemittel, die von den Clusterobjekten vergleichsweise häufig erinnert wurden, ebenfalls im Cluster enthalten sind.

Weitere Vorteile, wie etwa der aus der Analyse des Fusionsprozesses resultierende Erkenntnisgewinn, werden später an den entsprechenden Stellen hervorgehoben.

Zweimodale Klassifikationsverfahren lassen sich – grob vereinfacht – in hierarchische und nicht-hierarchische Verfahren unterteilen. Während letztere lediglich eine fertige Klassifikation liefern, zeigen hierarchische Verfahren den Fusionsprozeß als Baumdiagramm (Dendrogramm) und liefern damit weitere, wertvolle Erkenntnisse.

2.1 Nicht-hierarchische zweimodale Klassifikationsverfahren

Der von McCormick/Schweitzer/White (vgl. *McCormick/Schweitzer/White* 1972, S. 993ff.) entwickelte Bond-Energy-Algorithmus (BEA) permutiert die Zeilen und

Spalten einer Datenmatrix \mathbf{X} so, daß eine Matrix mit dichten Blöcken numerisch hoher Werte entsteht. Erreicht wird dies durch Maximierung des Effektivitätsmaßes

$$\text{EM}(\mathbf{X}) = \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^m x_{ij} (x_{i-1,j} + x_{i+1,j} + x_{i,j-1} + x_{i,j+1}) \quad , \quad (1)$$

mit $x_{0,j} = x_{n+1,j} = x_{i,0} = x_{i,m+1} = 0$.

Die zu den Blöcken gehörenden Objekte und Merkmale bilden jeweils einen zweimodalen Cluster.

Als Hauptkritikpunkt am BEA ist das Fehlen einer eindeutigen Vorschrift zur Bildung von Clustern zu nennen. Selbst eine optimal permutierte Matrix läßt unter Umständen großen Spielraum bei der Interpretation, da die Abgrenzung von “Blöcken numerisch hoher Werte” subjektiven Einflüssen unterworfen ist. Darüber hinaus ist für die Güte der Klassifikation keine Kennzahl vorgesehen (vgl. *DeSarbo* 1982, S. 450).

Eine weitere Gruppe nicht-hierarchischer Verfahren bilden die GENNCLUS-Varianten, mit deren Hilfe zweimodale, disjunkte oder nicht-disjunkte Klassifikationen erzeugt werden können.

Unter Verwendung der Datenmatrix \mathbf{X} lautet das GENNCLUS-Modell

$$\hat{\mathbf{X}} = \mathbf{P}\mathbf{W}\mathbf{Q}^T + \mathbf{C} \quad \text{mit} \quad (2)$$

$$\begin{aligned} \hat{\mathbf{X}} = (\hat{x}_{ij})_{n \times m} &= \text{Matrix der geschätzten Ausprägungen von Objekt } O_i \text{ bzgl. Merkmal } M_j \\ & \quad i = 1, \dots, n, \quad j = 1, \dots, m \\ \mathbf{P} = (p_{ir})_{n \times k} &= \text{Matrix der binären Zugehörigkeitszahlen von Objekt } i \text{ zu Cluster } r \\ & \quad i = 1, \dots, n, \quad r = 1, \dots, k \\ \mathbf{Q} = (q_{jr})_{m \times k} &= \text{Matrix der binären Zugehörigkeitszahlen von Merkmal } j \text{ zu Cluster } r \\ & \quad j = 1, \dots, m, \quad r = 1, \dots, k \\ \mathbf{W} = (w_{rr'})_{k \times k} &= \text{Matrix der Beziehungen zwischen Cluster } r \text{ und } r' \\ & \quad \text{(Clustergewichtungen)} \\ & \quad r, r' = 1, \dots, k \\ \mathbf{C} = (c)_{n \times m} &= \text{Matrix mit konstanten Elementen.} \end{aligned}$$

Im Modell (??) sind die Matrizen \mathbf{P} , \mathbf{W} , \mathbf{Q} und \mathbf{C} zu schätzen. Hier wird unterstellt, daß die Ähnlichkeit zweier Elemente als additives Maß der Gewichte jener Cluster angesehen werden kann, denen beide Elemente angehören.

Zur Lösung von (??) wurden verschiedene Algorithmen vorgeschlagen (vgl. *DeSarbo* 1982, S. 453ff., *Both/Gaul* 1985, S. 2ff., *Gaul/Schader* 1996, S. 20f.). Bei allen angegebenen Lösungsverfahren muß die Klassenanzahl k a priori angegeben werden. Idealerweise erhöht man also – beginnend mit $k_0 = 2$ – die Anzahl der Cluster schrittweise um eins bis zu einer als sinnvoll erachteten Obergrenze k_{\max} . Anhand der Gütekriterien VAF oder TIC (2) ermittelt man diejenige Clusteranzahl, die die Ausgangsdaten in \mathbf{X} am besten reproduziert.

Zusammenfassend kann festgehalten werden, daß der entscheidende Vorteil der GENN-CLUS-Varianten darin liegt, daß der Anwender auch nicht-disjunkte Klassifikationen vornehmen, also Werbemittel oder Zielgruppensegmente ggf. mehreren Klassen zuordnen kann.

Dabei darf nicht übersehen werden, daß die Lösung von (??) mit Hilfe von Gradientenprojektionsverfahren erfolgt, so daß das Auffinden globaler Extremalstellen nicht gesichert ist. Je nach Beschaffenheit der Datenmatrix hängt die Güte der resultierenden Klassifikation in hohem Maße von der Startkonfiguration, d.h. von der Vorgabe der Matrizen \mathbf{P} und \mathbf{Q} ab.

2.2 Hierarchische zweimodale Klassifikationsverfahren

Ziel der von Eckes/Orlik (vgl. *Eckes/Orlik* 1991 und 1993) entwickelten Zentroid-Effekt-Methode ist es, die Objekte und Merkmale der Datenmatrix \mathbf{X} in zweimodale Cluster mit minimaler innerer Heterogenität zusammenzufassen. Als Maßstab für die Heterogenität einer Klasse \mathcal{K}_r dient das varianzähnliche Kriterium

$$\text{MQA}_r = \sigma_r + (\bar{x}_r - \mu)^2, \quad (3)$$

wobei σ_r die Varianz der r -ten Klasse, \bar{x}_r den Mittelwert der Ausprägungen in Klasse r und μ den maximalen Eintrag in der Datenmatrix bezeichnen.

Eine völlig andere Vorgehensweise verwenden die sog. indirekten Verfahren der zweimodalen Klassifikation, die durch folgende Vorgehensweise charakterisierbar sind.

1. Unter Verwendung der Ausgangsdaten der Matrix $\mathbf{X}_{(n \times m)}$ wird eine sog. Grand-Matrix $\mathbf{G}_{(n+m) \times (m+n)}$ konstruiert, die folgende Distanzen enthält:
 - a) Distanzen zwischen den Zeilenelementen O_i und $O_{i'}$, $i, i' = 1, \dots, n$, $i < i'$ (in Abb. ?? ist das der mit $d_{ii'}$ bezeichnete Block).
 - b) Distanzen zwischen den Spaltenelementen M_j und $M_{j'}$, $j, j' = 1, \dots, m$, $j < j'$ (Block $d_{jj'}$ in Abb. ??).
 - c) Distanzen zwischen den Zeilenelementen O_i und den Spaltenelementen M_j , $i = 1, \dots, n$, $j = 1, \dots, m$ (Block d_{ij} in Abb. ??).
2. Falls \mathbf{G} die ultrametrische Ungleichung (3) nicht erfüllt, sind geeignete Transformationen $f : \mathbf{G}_{(n+m) \times (m+n)} \rightarrow \mathbf{U}_{(n+m) \times (m+n)}$ vorzunehmen, die die Einhaltung der Ultrametrik garantieren. Nur dann ist die exakte hierarchische Repräsentation der Klassifikation in Form eines Dendrogrammes möglich.
3. \mathbf{G} wird dann als Distanzmatrix mit den bekannten einmodalen hierarchischen Verfahren der Clusteranalyse, die in der angegebenen Literatur gut dokumentiert sind (vgl. z.B. *Opitz* 1980, S. 65ff. und *Bausch/Opitz* 1993, S. 55ff.), bearbeitet.

Die vordringliche Aufgabe der indirekten Methoden zweimodaler Klassifikation ist also die Erzeugung einer Grand-Matrix, die die ultrametrische Ungleichung erfüllt. Üblicherweise wird die Datenmatrix \mathbf{X} mit Ähnlichkeitswerten zwischen Objekt i und Merkmal j als Input verwendet.

Zur Ermittlung der Grand-Matrix bieten sich drei theoretisch denkbare Wege an (vgl. *Both/Gaul* 1986, S. 595):

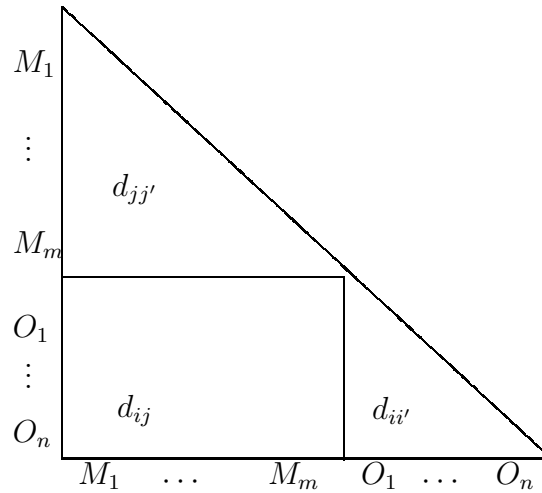


Abbildung 1: Grand-Matrix

- (1) Die Datenmatrix \mathbf{X} wird zu einer vorläufigen Grand-Matrix \mathbf{G} ergänzt, die anschließend in eine der ultrametrischen Ungleichung genügende Grand-Matrix \mathbf{U} umgewandelt wird. Bis dato wurde diese Vorgehensweise allerdings nicht aufgegriffen.
- (2) Die Datenmatrix \mathbf{X} wird zunächst in eine ultrametrische Matrix \mathbf{T} transformiert, die \mathbf{X} möglichst ähnlich ist. Erst dann erfolgt eine Ergänzung zur Grand-Matrix \mathbf{U} , wobei die Einhaltung der Ultrametrik erneut zu berücksichtigen ist. Diese Vorgehensweise wenden De Soete et al. im Two Mode Least Squares Algorithmus (2MLS) an (4).
- (3) Ein dritter Weg besteht darin, von \mathbf{X} ausgehend sofort eine bestmögliche Grand-Matrix \mathbf{U} zu bestimmen. Zwei Ansätze, die diese Möglichkeit aufgreifen werden nachfolgend kurz skizziert.

Das Grundprinzip der *Missing-Value-Algorithmen* zur zweimodalen Klassifikation (vgl. *Espejo/Gaul* 1986, S. 123) ist es, eine Datenmatrix $\mathbf{X}_{n \times m}$ gemäß der Vorschrift

$$d_{ij} = \max_{i,j} x_{ij} - x_{ij} \quad (4)$$

in eine Distanzmatrix $\mathbf{D} = (d_{ij})_{n \times m}$ zu transformieren, die als bekannter, in Abb. ?? als Block der d_{ij} bezeichneter Teil einer Grand-Matrix angesehen wird. Die restlichen Distanzwerte der Grand-Matrix werden als fehlend definiert und im Algorithmus nicht explizit berücksichtigt. Auf \mathbf{D} werden dann die aus der einmodalen Clusteranalyse bekannten Linkage-Verfahren sinngemäß übertragen.

Die bisher dargestellten hierarchischen Verfahren berücksichtigen Ähnlichkeiten in einer Weise, die bisweilen – vor allem im Rahmen der Werbewirkungskontrolle – als unzureichend angesehen werden muß. Der Grund dafür liegt in der Tatsache, daß sie

1. als Ähnlichkeiten de facto nur (geringe) Distanzen oder (hohe) Assoziationswerte zwischen den Zeilenelementen und den Spaltenelementen einer Datenmatrix \mathbf{X} verarbeiten und die Ähnlichkeitsbeziehungen der Zeilenelemente

untereinander sowie der Spaltenelemente untereinander im wesentlichen vernachlässigen, und

2. auf niedrigstem Fusionsniveau stets die Zusammenlegung eines Zeilenelementes und eines Spaltenelementes verlangen. Dadurch werden unter Umständen identische Zeilenelemente oder identische Spaltenelemente nicht sofort, d.h. zu Beginn des Fusionsprozesses, zu einem Cluster vereinigt. Das in der Clusteranalyse allgemeingültige Prinzip, Elemente anhand beobachteter oder geschätzter Ähnlichkeiten in homogene Gruppen zusammenzufassen, wird damit verletzt.

Exakt diesen Punkt greift der ESOCLUS-Algorithmus (vgl. *Schwaiger 1997*) mit der expliziten Berücksichtigung der Ähnlichkeiten der Zeilenelemente (Objekte) und der Spaltenelemente (Merkmale) einer Datenmatrix untereinander durch Berechnung der paarweisen City-Block-Distanzen auf.

Ausgehend von $\mathbf{X}_{(n \times m)}$ werden zunächst Distanzen zwischen Objekt i und Objekt i' gemäß

$$d_{ii'} = \frac{1}{m} \sum_{j=1}^m |x_{ij} - x_{i'j}|, \quad i, i' = 1, \dots, n, \quad (5)$$

und Distanzen zwischen Merkmal j und Merkmal j' gemäß

$$d_{jj'} = \frac{1}{n} \sum_{i=1}^n |x_{ij} - x_{ij'}|, \quad j, j' = 1, \dots, m, \quad (6)$$

berechnet. Die zur Vervollständigung der in Abb. ?? gezeigten Grand-Matrix benötigten Distanzen zwischen Objekt i und Merkmal j werden durch

$$d_{ij} = \max_{i,j} x_{ij} - x_{ij}, \quad i = 1, \dots, n, \quad j = 1, \dots, m. \quad (7)$$

ermittelt.

Zur sinnvollen Anwendung dieses Algorithmus ist es erforderlich, daß die Daten in \mathbf{X} identisches Skalenniveau und ähnliche Größenordnungen in den Merkmalsausprägungen aufweisen. Letzteres kann durch Standardisierung der Datenmatrix gewährleistet werden.

Mit (??)–(??) sind bis dato allerdings unterschiedliche Größenordnungen in den Distanzblöcken der Grand-Matrix und damit unterschiedliche Einflüsse der Distanzarten (Objekt-Objekt, Merkmal-Merkmal und Objekt-Merkmal) zu befürchten.

Sofern keine begründete Absicht zur Hervorhebung der Bedeutung einzelner Distanzblöcke besteht, ist es zweckmäßig, den Fusionsprozeß durch alle drei Distanzarten in identischer Weise determinieren zu lassen. Hierzu müssen die nach (??)–(??) gebildeten Distanzwerte normiert werden, indem eine Division der einzelnen Distanzen eines Blocks durch die jeweilige maximale Distanz innerhalb eines Distanzblocks erfolgt.

3 Ein Anwendungsbeispiel

Die folgenden Abschnitte zeigen die Relevanz zweimodaler Klassifikation anhand einer Fallstudie. Dabei wird insbesondere der geforderten Bewertung eines Werbemittels in Relation zu Konkurrenzwerbemitteln Rechnung getragen, indem einerseits Ähnlichkeitsbeziehungen zwischen Werbemitteln und zwischen einzelnen Segmenten innerhalb der Zielgruppe berücksichtigt und andererseits Kennzahlen der Aufmerksamkeitswirkung auf eben jene Zielgruppensegmente aufgeschlüsselt werden.

Kurz zusammengefaßt läßt sich die präsentierte Vorgehensweise am Beispiel des ungestützten Recalltests wie folgt beschreiben:

1. Die im Rahmen einer Stichprobenerhebung ausgewählten Versuchspersonen werden in homogene Cluster aufgespalten. Hierzu können entweder Klassenmerkmale vorgegeben (Geschlecht, regionale Herkunft, Alter etc.), oder einmodale Klassifikationsverfahren zu deren Bildung angewandt werden. Beispielsweise könnten vor Abfrage der Recallwerte den Testteilnehmern geeignete Items vorgelegt werden, die auf einer Rating-Skala abzufragen sind und die Bildung psychographischer Zielgruppen-Cluster erlauben.
2. Im Rahmen eines ungestützten Recalltests wird erhoben, an welche Werbemittel sich die Testpersonen erinnern.
3. Pro Werbemittel werden die clusterspezifischen Recallwerte gebildet.
4. Die Auswertung der Testergebnisse erfolgt mit Hilfe zweimodaler Klassifikationsverfahren, wobei das Hauptaugenmerk in erster Linie der Analyse des Fusionsprozesses und erst in zweiter Linie der resultierenden Klassenstruktur gilt.

Die wesentlichen Schritte dieser kurzen Übersicht werden nachfolgend detailliert erläutert.

3.1 Die notwendige Datenbasis

Im Rahmen eines Foldertests wurden Print-Werbemittel für die Parfums CASMIR, GAULTIER, YPNO, CERRUTI, DOLCE VITA, TALISMAN, NIGHTFLIGHT, ROMA und FEMME untersucht (5).

Im vorliegenden Fall wurden 50 BWL-Studenten beiderlei Geschlechts als Testpersonen ausgewählt und unter Verwendung einmodaler Clusteranalysen anhand einiger psychographischer Statements in vier Segmente unterteilt. Ohne näher auf die Vorgehensweise einzugehen beschreiben wir die Segmente als ‘Hedonisten’, ‘Karrieretypen’, ‘Frustrierte’ und ‘Unauffällige’. Zusätzlich zu dieser Unterteilung haben wir in der Folge das Geschlecht der Versuchspersonen als weiteres Gruppierungsmerkmal verwendet. Die in Tab. ?? ausgewiesenen Segmente müssen also nicht paarweise elementfremd sein; selbst einer weiteren Unterteilung (nach Alter etc.) stünden keine Einwände gegenüber.

Die werbemittelspezifischen Recallwerte der eben beschriebenen sechs Klassen zeigt Tab. ??, die zugleich als Datenmatrix $\mathbf{X} = (x_{ij})_{6 \times 9}$ die Ausgangsbasis für zweimodale Klassifikationsverfahren liefert. x_{ij} bezeichnet dabei den Anteil der Personen aus Cluster i ($i = 1, \dots, 6$), die sich an Werbemittel j ($j = 1, \dots, 9$) erinnerten.

	CAS- MIR	GAUL- TIER	YPNO	CER- RUTI	DOLCE VITA	TALIS- MAN	NIGHT- FLIGHT	ROMA	FEM- ME
Unauffällige	0,57	0,57	0,50	0,50	0,14	0,07	0,64	0,50	0,64
Frustrierte	0,42	0,58	0,17	0,08	0,08	0,17	0,25	0,17	0,17
Karrieretypen	0,08	0,17	0,25	0,42	0,25	0,25	0,67	0,42	0,33
Hedonisten	0,50	0,25	0,25	0,42	0,00	0,17	0,42	0,33	0,33
Männer	0,29	0,33	0,19	0,33	0,10	0,19	0,43	0,38	0,19
Frauen	0,48	0,45	0,38	0,38	0,14	0,14	0,55	0,34	0,52
Alle Segmente	0,40	0,40	0,30	0,36	0,12	0,16	0,50	0,36	0,38

Tabelle 1: Klassenspezifische Recallwerte

3.2 Die Auswertung der spezifischen Recallwerte

Eine genauere Betrachtung der Datenmatrix \mathbf{X} zeigt, daß die Klasse der ‘Unauffälligen’ insgesamt mit relativ hohen Recallwerten aufwarten kann, sieht man von den allgemein wenig erinnerten Anzeigen für DOLCE VITA und TALISMAN ab.

Ferner wird deutlich, daß die Anzeige für NIGHTFLIGHT sich stark gegen die Konkurrenten durchsetzen kann, also häufig erinnert wird.

Wie gut können solche (und weitere, nicht eben augenfällige) Sachverhalte nun visualisiert werden? Zur Klärung dieser Frage vergleichen wir die Ergebnisse der in Abschnitt ?? besprochenen Methoden. Die Klassenstrukturen der Zentroid-Effekt-Methode (ZEM), des Missing-Value-Average-Linkage Verfahrens (MV), des ESOCLUS-Algorithmus in der Average-Linkage-Variante (ESOCLUS) und des GENNCLUS sind aus Tab. ?? ersichtlich. Zur Festlegung der Klassenanzahl bei den hierarchischen Verfahren wurde jeweils das Ellenbogenkriterium herangezogen, das zeigt, bei welchem Übergang zur nächstniedrigeren Klassenanzahl die Klassifikationsbewertung sich am deutlichsten verschlechtert. Bei GENNCLUS wurden, ausgehend von verschiedenen Startkonfigurationen vier und fünf Klassen umfassende Lösungen erzeugt, wobei wir hier nur die 5-Klassen-Lösung mit den besten Gütemaßen ($VAF = 0,830$ und $TIC = 0,095$) präsentieren.

Einige Ergebnisse sind teilweise unbefriedigend: So weisen ZEM/MV z.B. Hedonisten der Klasse 1 mit NIGHTFLIGHT zu, obwohl deren Recallwert bzgl. CASMIR deutlich höher liegt. Frauen befinden sich ebenfalls in Klasse 1, obwohl sie relativ hohe Recallwerte auch bzgl. FEMME und CASMIR aufweisen.

Im Gegensatz zu ZEM/MV werden die Hedonisten von ESOCLUS ‘richtig’ klassifiziert. Aufgrund der Berücksichtigung von Ähnlichkeitsbeziehungen werden Karrieretypen und NIGHTFLIGHT separat ausgewiesen (Klasse 1), weil Karrieretypen bzgl. ihrer Recallwerte sehr deutlich von allen anderen Gruppen abweichen und NIGHTFLIGHT sich bzgl. der erzielten Recallwerte sehr deutlich von anderen Anzeigen absetzt. Diese Klassifikation ist stabil, da die Anwendung eines Austauschverfahrens die Klassen unverändert läßt.

Eine völlig andere, trotz hoher Gütemaße nur mit erheblichen Schwierigkeiten interpretierbare Klassifikation liefert GENNCLUS. So sind insbesondere die Klassen 3 (Frustrierte, DOLCE VITA, TALISMAN) und 4 (Männer, YPNO, FEMME, NIGHTFLIGHT) mit den Ausgangsdaten nicht in Einklang zu bringen. Gibt man anstelle einer zufälligen Startpartition z.B. die ESOCLUS-Lösung vor, so wird gegenüber letzterer lediglich ROMA in Klasse 1 verschoben, weitere Veränderungen

	ZEM und MV	ESOCCLUS	GENNCLUS
\mathcal{K}_1	Karrieretypen Frauen Männer Hedonisten NIGHTFLIGHT	Karrieretypen NIGHTFLIGHT	Karrieretypen CERRUTI ROMA
\mathcal{K}_2	Unauffällige FEMME CASMIR YPNO CERRUTI ROMA	Unauffällige FEMME YPNO CERRUTI ROMA	Unauffällige Frauen GAULTIER
\mathcal{K}_3	Frustrierte GAULTIER	Frustrierte GAULTIER	Frustrierte DOLCE VITA TALISMAN
\mathcal{K}_4	TALISMAN	TALISMAN DOLCE VITA	Männer YPNO FEMME NIGHTFLIGHT
\mathcal{K}_5	DOLCE VITA	CASMIR Hedonisten Frauen	CASMIR Hedonisten
\mathcal{K}_6		Männer	

Tabelle 2: Aus zweimodaler Klassifikation resultierende Klassen

ergeben sich nicht. Die VAF dieser Klassifikation ist allerdings negativ. Die Problematik der verwendeten Gütemaße, auf die wir in Abschnitt ?? detaillierter eingehen werden, ist offensichtlich.

Das noch fehlende Ergebnis des Bond-Energy-Algorithmus zeigt Tab. ?. Von der Umrandung eingefasst sind die Teile der Matrix, die nach Ansicht des Verfassers “dichte Blöcke numerischer hoher Werte” darstellen.

	TALIS- MAN	YPNO	GAUL- TIER	CAS- MIR	FEM- ME	NIGHT- FLIGHT	CER- RUTI	ROMA	DOLCE VITA
Männer	0,19	0,19	0,33	0,29	0,19	0,43	0,33	0,38	0,10
Karrieretypen	0,25	0,25	0,17	0,08	0,33	0,67	0,42	0,42	0,25
Frauen	0,14	0,38	0,45	0,48	0,52	0,55	0,38	0,34	0,14
Unauffällige	0,07	0,50	0,57	0,57	0,64	0,64	0,50	0,50	0,14
Hedonisten	0,17	0,25	0,25	0,50	0,33	0,42	0,42	0,33	0,00
Frustrierte	0,17	0,17	0,58	0,42	0,17	0,25	0,08	0,17	0,08

Tabelle 3: BEA-Endergebnis

Abgesehen davon, daß die Umrandung in Tab. ?? subjektiv gezogen wurde, fällt die Analyse des Ergebnisses nicht leicht. Auf jeden Fall sind die Klassen als überlappende Cluster zu sehen, weil klare Trennungen zwischen den Blöcken fehlen. Im Kern sind die Klasse 2 der ZEM/des MV und des ESOCCLUS (YPNO, CERRUTI, ROMA, FEMME und Unauffällige) erkennbar, die hier wegen der möglichen Überlappung

um CASMIR, GAULTIER und NIGHTFLIGHT erweitert werden. Gleiches gilt für die Klasse 1 der ZEM/des MV, die von ESOCLUS allerdings aufgespalten wurde.

Die Zuordnung der Frustrierten zu GAULTIER (ZEM, MV, ESOCLUS) bzw. CASMIR zu Hedonisten (ESOCLUS) ist ebenso nachvollziehbar, wie die Alleinstellung der Anzeigen für DOLCE VITA und TALISMAN.

Ohne Kenntnis der hierarchischen Klassifikation bietet der Bond-Energy-Algorithmus zumindest eine einfache Lesehilfe: Blöcke, die viele Zeilenelemente umfassen kennzeichnen Werbemittel, die hohe Aufmerksamkeit erregen (hier CASMIR und NIGHTFLIGHT), und Blöcke, die viele Spaltenelemente umfassen kennzeichnen Zielgruppensegmente, die insgesamt hohe Recallwerte zeigen (hier Frauen und Unauffällige).

Im übrigen wird deutlich, daß die Interpretation der BEA-Ergebnisse um so schwieriger wird, je geringer die numerischen Unterschiede in den Ausprägungen der Datenmatrix sind. Speziell zur Auswertung ungestützter Recalltests ist aus diesem Grund bei Anwendung des BEA eine gewisse Erfahrung im Umgang mit diesem Algorithmus nötig.

Betrachten wir abschließend eine überlappende 5-Klassen-Lösung von GENNCLUS, wobei wir als Startkonfiguration die disjunkte Klassifikation aus Tab. ?? vorgegeben haben. Für die in Tab. ?? gezeigte Lösung erhält man $VAF = 0,92$ und $TIC = 0,086$,

Klasse	Zugeordnete Elemente
\mathcal{K}_1	Unauffällige, Hedonisten CASMIR, GAULTIER, YPNO, CERRUTI, NIGHTFLIGHT, ROMA, FEMME
\mathcal{K}_2	Frustrierte, CASMIR, GAULTIER, YPNO, TALISMAN, NIGHTFLIGHT, ROMA, FEMME
\mathcal{K}_3	Karrieretypen, YPNO, CERRUTI, DOLCE VITA, TALISMAN, NIGHTFLIGHT, ROMA, FEMME
\mathcal{K}_4	Frauen, CASMIR, GAULTIER, YPNO, CERRUTI, NIGHTFLIGHT, ROMA, FEMME
\mathcal{K}_5	Männer, CASMIR, GAULTIER, YPNO, CERRUTI, TALISMAN, NIGHTFLIGHT, ROMA, FEMME

Tabelle 4: Überlappende GENNCLUS-Lösung mit 5 Klassen

also eine sehr ansprechende Güte.

Zur Interpretation im ökonomischen Sinne ermittelt man hier zweckmäßiger Weise, welche Werbemittel in den einzelnen Segmenten fehlen. Zunächst fällt auf, daß sich DOLCE VITA nur in einer Klasse (\mathcal{K}_3) befindet, während TALISMAN immerhin 3 Klassen (\mathcal{K}_2 , \mathcal{K}_3 und \mathcal{K}_5) zugewiesen wurde. Ferner wird deutlich, daß sich CASMIR und GAULTIER mit Ausnahme der Klasse 3 in allen Klassen wiederfinden.