

1. Einleitung

Im Rahmen der Forschungsarbeiten innerhalb des Graduiertenkollegs „Energiebereitstellung aus regenerativen Energiequellen“ und an der Professur für Systemtheorie wurden verschiedene Analyse- und Prognosemethoden für Zeitreihen auf Basis der Fuzzy Pattern Klassifikation erarbeitet. Insbesondere im Bereich des Energiemanagements konnte dabei die Leistungsfähigkeit der Verfahren nachgewiesen werden.

So wurden in Zusammenarbeit mit einem regionalen Energieanbieter Prognosemodelle für Energielastgänge auf Basis von 100 gemessenen Tagesverläufen erstellt, deren Vorhersageergebnisse für die ¼-Stunden-Prognose in 99,7% aller Fälle und für die 2-Stunden-Prognose in 95,0% aller Fälle einen Prognosefehler unter 4% aufweisen [4]. Für die Modelle werden dabei Wochen- und Feiertage, sowie Sommer- und Winterzeit unterschieden und für die vier entstehenden Kategorien separate dynamische Fuzzy Pattern Modelle erstellt. Die Basis dieser Modelle bildeten sogenannte Fuzzy Prototypen. Ein Fuzzy Prototyp bildet dabei das dynamische Verhalten einander ähnlicher Tagesverläufe sowohl als repräsentativen Verlauf, den sogenannten Repräsentanten, als auch in ihren Unwägbarkeiten (Meßfehlern, Änderungen des Verbraucherverhaltens, Wettereinflüsse usw.) als unscharfe (fuzzy) Beschreibung ab¹.

Die Vorhersage erfolgt über eine unscharfe Zuordnung des bekannten Zeitreihenabschnittes zu den Prototypen und einer Vorhersage für jeden einzelnen Prototyp. Diese Ergebnisse werden unscharf bewertet und auf Basis von Defuzzifikationsmethoden zu einer Gesamtprognose zusammengefaßt [5][4].

In einer weiteren Studie im Bereich der Windenergieprognose [3] wurde auf ein quasi-dynamisches Modell zurückgegriffen, bei dem die Dynamik der Zeitverläufe in sogenannten Merkmalen erfaßt wurde und ein (statischer) Fuzzy Pattern Klassifikator erstellt wurde, dessen Klassen typische Zustände des Windenergieverlaufes repräsentieren. Dabei wurden wiederum zueinander ähnliche zeitliche Messwertverläufe klassifiziert und zu Fuzzy Pattern Klassen zusammengefaßt.

Im konkreten Fall konnten dabei 7 Verlaufstypen (Klassen) anhand von zwei Merkmalen unterschieden werden. Für die Prognose sind den Verlaufstypen jeweils eine Prognosestrategie zugeordnet. Mittels den Methoden der Defuzzifikation werden aus den Einzelprognosen eine Gesamtprognose erzeugt. Für diesen Anwendungsfall lagen die Fehler für die ¼-Stunden-Vorhersage in über 90% aller Fälle bei 4%.

Die Leistungsfähigkeit der Zeitreihenanalyse und -prognose konnten darüber hinaus auch im Bereich Verkehr und Umwelt nachgewiesen werden [4].

In den bisherigen Anwendungen wurden für die Klassifikation der Meßreihen Methoden der Clusteranalyse sowie Expertenwissen verwendet. Dabei zeigte sich, daß die herkömmlichen Methoden der automatischen Klassifikation nur bedingt den Kriterien für Modelle auf Basis der Fuzzy Pattern Klassifikation genügen. Dadurch werden Gütebewertungen und darauf aufbauende Klassifikatoranpassungen notwendig, um einen geeigneten Klassifikator zu erhalten.

Auf Grund dieser Tatsache und basierend auf herkömmlichen Klassifikationsverfahren wurde eine automatische Klassifikation innerhalb der Theorie der Fuzzy Pattern Klassifikation entwickelt, deren Ergebnis ein vollständiger Fuzzy Pattern Klassifikator ist.

¹ siehe auch 2. Zeitreihenanalyse und -prognose mit Fuzzy Pattern Modellen

2. Zeitreihenanalyse und –prognose mit Fuzzy Pattern Klassifikation

Ein Modell der unscharfen Zeitreihenanalyse [3] stützt sich auf die Annahme, daß im Zeitverhalten wiederholt die gleichen Phänomene auftreten und diese ein bestimmtes zukünftiges Verhalten bedingen. Dabei wird von einem festen Zeitfenster ausgegangen. Auf Basis der Meßgrößen des Zeitfensters gilt es nun Merkmale zu finden, die die Prozeßzustände und die im Prozeß enthaltene Dynamik ausreichend beschreiben. Mittels dieser Merkmale werden für jedes Zeitfenster sogenannte Objekte definiert, die ihrerseits zu Klassen zusammengefaßt werden können (Abb. 1).

Ein derartiger Ansatz führt zu einem „statischen“ Fuzzy Pattern Klassifikator [1] (Abb. 2).

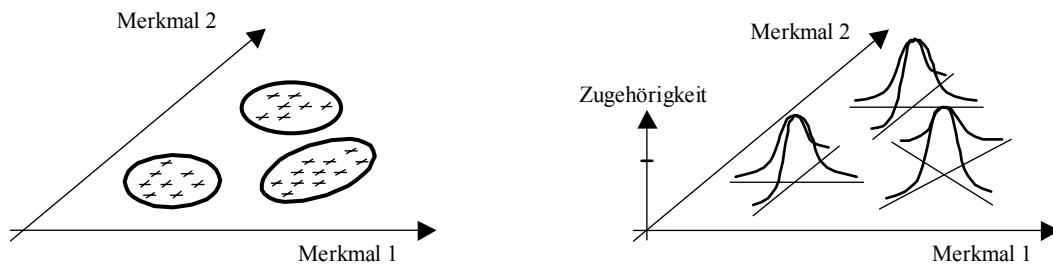


Abb. 1: Objektstrukturen eines statischen Klassifikators Abb. 2: Statischer Fuzzy Pattern Klassifikator

Für die Prognose wird zum Vorhersagezeitpunkt die bekannte Zeitreihe in einen Merkmalsvektor (Objekt) überführt, welche anhand des Klassifikators identifiziert wird, dadurch erhält es eine unscharfe Zuordnung zu jeder Klasse. Den verschiedenen Klassen ist wiederum jeweils eine Prognosestrategie zugeordnet, die im einfachsten Fall ein konkreter Prognosevektor ist. Diese Strategien führen für jede Klasse zu einer Prognose, die mittels der Methoden der Defuzzifikation, beispielsweise der unscharfen Wichtung, in einen Prognosevektor überführt werden. Die Klasse, zu der das Objekt die größte Zugehörigkeit aufweist, beeinflusst dabei die Prognose am stärksten.

Ein anderes Modell basiert auf der Zerlegung von Zeitverläufen anhand sogenannter Triggerpunkte in wohldefinierte Zeitreihenabschnitte (z. B. Tagesverläufe) [4] [5]. Bei einem solchen Ansatz bleiben i. Allg. die Meßwerte als Merkmale erhalten. Die Objekte sind in diesem Fall Zeitreihen (Abb. 3) und die Klassifikationsaufgabe erweitert sich dahingehend, daß eine Ähnlichkeit der Zeitreihen über dem gesamten Zeitabschnitt bewertet werden muß.

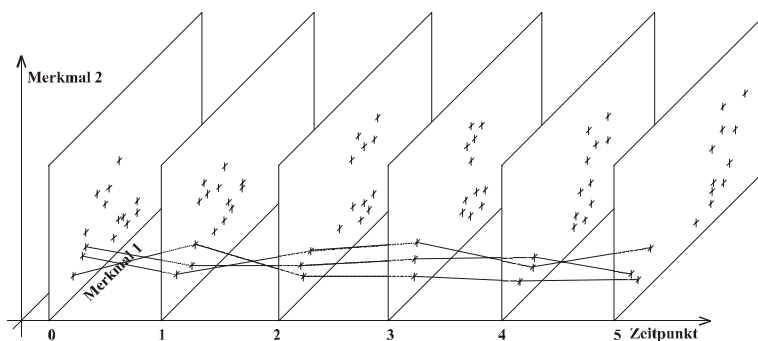


Abb. 3: Zeitreihen eines Zeitabschnittes

Die Ergebnisse der Klassifikation sind Klassen von Zeitreihen (Abb. 4), die zu jedem einzelnen Zeitpunkt einen vollständigen „statischen“ Klassifikator bilden [5].

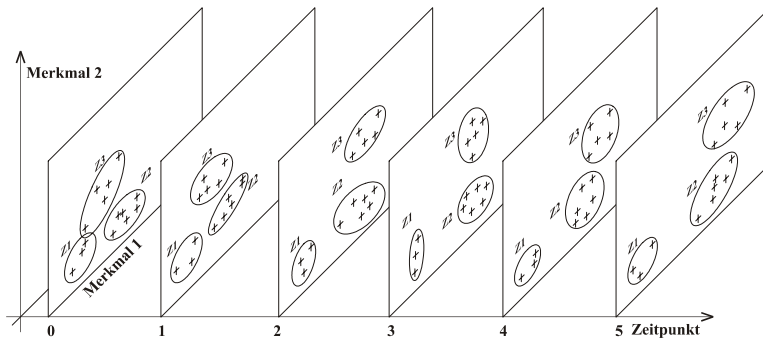


Abb. 4: Klassen von Zeitreihen

Diese Klassifikatoren werden in einem zweiten Schritt in Fuzzy Pattern Klassifikatoren überführt [5], deren Klassen in ihrem zeitlichen Verhalten sogenannte Fuzzy Prototypen darstellen (Abb. 5).

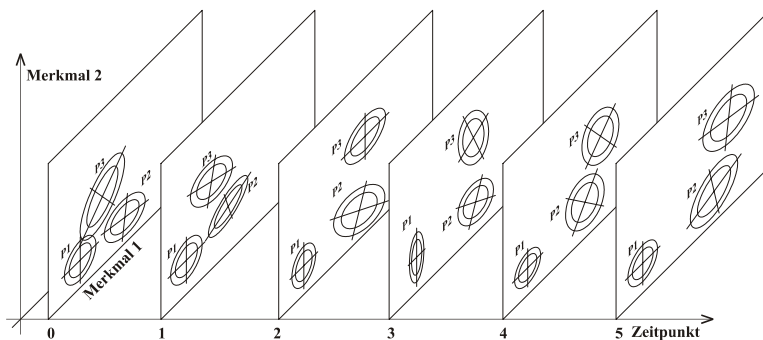


Abb. 5: Fuzzy Prototypen

Diese Fuzzy Prototypen repräsentieren im betrachteten (getriggerten) Zeitabschnitt Muster von typischen Zeitverläufen einschließlich auftretender Schwankungen.

Voraussetzung für die Prognose von Zeitreihen mittels eines derartigen Modells ist, daß die vorherzusagende Zeitreihe im Kontext mit diesem Modell steht, d.h. die Zeitreihe kann einem (oder auch mehreren) Prototypen zugeordnet werden, welche deren typischen Verlauf repräsentieren. Über diese Zuordnung wird die Zeitreihe ab dem Vorhersagezeitpunkt (letzter bekannter Zeitpunkt) prognostiziert. Die Fuzzy Prognose von Zeitreihen wird dabei in zwei Phasen unterteilt, eine Identifikationsphase und eine Defuzzifikationsphase [4][5].

Die Identifikationsphase enthält alle Algorithmen für eine unscharfe Zuordnung der Zeitreihe zu den bekannten Fuzzy Prototypen. Die bekannte Zeitreihe wird dabei nur im Bereich eines Zeitfensters betrachtet. Dieses Zeitfenster legt fest, wieviel Zeitpunkte der bekannten Zeitreihe in die Bestimmung der Sympathiewerte zu den entsprechenden Prototypen eingehen sollen. Dabei können die Zeitfensterlängen zu den verschiedenen Zeitpunkten auch variieren.

Als Resultat erhält man eine Menge von Sympathie- und Lagevektoren [5] des bekannten Zeitreihenabschnittes.

In der Defuzzifikationsphase werden diese Daten aus der Identifikationsphase für die vorherzusagenden Zeitpunkt über entsprechende Algorithmen in einen scharfen Objektvektor überführt, welcher dann den prognostizierten Vektor der Zeitreihe zu diesem Zeitpunkt bildet. Die prinzipielle Vorgehensweise erstellt für jeden Prototypen eine Prognose, welche für den vorherzusagenden Zeitpunkte über eine unscharfe Verknüpfung bezüglich der in der Identifikationsphase bestimmten Zugehörigkeiten in den scharfen Objektvektor überführt wird.

3. Eine Methode zur automatischen Klassifikation mittels Fuzzy Pattern Modellen

3.1 Grundlagen der Fuzzy Pattern Klassifikation (FPK)

Grundlage der Fuzzy Pattern Klassifikation (FPK) ist die formale Beschreibung der Klassen mittels einer parametrisierten Zugehörigkeitsfunktion - der Aizermanschen Potentialfunktion [1] - mit ihren Parametern a (Maximalwert der Zugehörigkeit - hier $a = 1$), der Randzugehörigkeit b , der Ausdehnung c , dem Abfall d der Zugehörigkeitsfunktion und der Lage u_0 . Dabei wird in der Regel zwischen dem links- und dem rechtsseitigen Verlauf unterschieden. Im mehrdimensionalen Fall kommt dazu noch eine Menge von Drehwinkeln, die eine „scharfe“ rotatorische Transformation der Objekte in den sogenannten Klassenraum [5] kennzeichnet.

Die Berechnungsformel der Zugehörigkeit eines Objektes im (p -dimensionalen) Klassenraum erfolgt mittels der allgemeinen mehrdimensionalen Aizermanschen Potentialfunktion:

$$\mu(\underline{u}) = \frac{a}{1 + \frac{1}{p} \sum_{i=1}^p \left(\frac{1}{b_i^{(lr)}} - 1 \right) \cdot \left(\frac{u_i}{c_i^{(lr)}} \right)^{d_i^{(lr)}}} \quad (1)$$

(\underline{u} ... in den Klassenraum transformiertes Objekt)

In Analogie zu dieser unscharfen Klassenbeschreibung werden für den Klassifikatoraufbau die einzelnen Objekte \underline{x}_k nicht als Meßvektoren, sondern als unscharfe Elementarklassen aufgefaßt. Die Beschreibung erfolgt mittels sogenannter elementarer Zugehörigkeitsfunktionen μ_k :

$$\mu_k(\underline{x}) = \left(1 + \frac{1}{p} \sum_{i=1}^p \left(\frac{1}{0,5} - 1 \right) \cdot \left(\frac{|x_i - x_{ki}|}{c_{ei}} \right)^2 \right)^{-1} = \left(1 + \frac{1}{p} \sum_{i=1}^p \left(\frac{|x_i - x_{ki}|}{c_{ei}} \right)^2 \right)^{-1} \quad (2)$$

Dabei werden die Parameter $d = 2$ und $b = 0,5$ gesetzt sowie $c_i = c_{ei}$. Der Parameter c_{ei} kennzeichnet dabei eine merkmalspezifische elementare Unschärfe.

Eine Menge/Klasse von Objekten mit derartigen elementaren Fuzzy-Beschreibungen kann mittels unscharfer Aggregation [1][5] zu einer Fuzzy Pattern Klasse verknüpft werden. Ausgangspunkt der Überlegungen für eine automatische Klassifikationsmethode auf Basis der Fuzzy Pattern Klassifikation ist die Tatsache, daß die bisher bekannten Klassifikationsverfahren Objekte nur als scharfe vektorielle Größen verarbeiten können. Im Ergebnis werden Klassen von Objekten gebildet, die ihrerseits wiederum scharfe Mengen beschreiben. Erst für die Aggregation dieser Objektgruppen zu einer Fuzzy Klasse werden unscharfe Informationen, wie die elementare Unschärfe, die Lage der Objekte zueinander und die Objektverteilung innerhalb einer klassifizierten Menge eingearbeitet. Diese Informationen sind jedoch für die Lage und die Ausprägung einer Fuzzy Pattern Klasse von großer Bedeutung.

3.2 Ausgewählte Methoden der automatischen Klassifikation

Insbesondere bei mehrdimensionalen Problemen stellt sich die Frage, welche Anforderungen von Seiten der Aggregation durch die mittels der automatischen Klassifikation erzeugten Objektstrukturen erfüllt werden müssen, damit eine sinnvolle Beschreibung der Objektmenge durch die FPK erfolgen kann. Fuzzy Pattern Klassen haben näherungsweise eine hyperellipsoidförmige Ausprägung, so daß sie Objektgruppen mit einer annähernd hyperellipsoiden Form am besten beschreiben können. Derartige Objektgruppen werden bei Bock [2] als kompakte Mengen oder kompakte Klassen bezeichnet².

Deshalb sollen hier nur automatische Klassifikationsverfahren betrachtet werden, die zu kompakten (hyperellipsoidförmigen) Punktwolken führen.

² Bock [2] definiert diese Klassen als Mengen, die „in ihrer (u. U. hypothetischen) räumlichen Darstellung eine runde oder ovale Form aufweisen“ (Seite 101). Für diese Art von Strukturen werden an verschiedensten Stellen Clusterverfahren aufgezeigt, die derartige Mengen erzeugen (bspw. S. 219 u. 227).

Die „klassischen“ Clusterverfahren basieren auf der Definition von Ähnlichkeits- bzw. Unähnlichkeitsmaßen. Diese wiederum beziehen sich i. Allg. auf Distanzen d_{kl} zwischen zwei Objekten O_k und O_l . Diese Distanzen, die gleichzeitig auch ein Unähnlichkeitsmaß zwischen zwei Objekten darstellen, müssen den verschiedenen Bedingungen genügen.³

Ein typisches Beispiel für Distanzmaße sind die gewichteten MINKOWSKI-Distanzen

$$d_{kl}^{(r)} = r \sqrt[r]{\sum_{i=1}^p g_i |x_{ki} - x_{li}|^r} \quad (3)$$

mit dem Spezialfall der euklidischen Distanz für ($r = 2$).

Bei der Definition von Distanzen D_{ij} von Objektmengen A_i und A_j (i.allg. disjunkt) wird auf Definitionen in Analogie zu den Objektdistanzen zwischen den einzelnen Mengenelementen zurückgegriffen. So werden Distanzen zwischen Objektmengen beispielsweise über die maximale, minimale oder mittlere Distanz ihrer Objekte bestimmt⁴. Bei der Berechnung von Distanzen zwischen Objektmengen und Einzelobjekten werden letztere als einelementige Mengen betrachtet.

Im Folgenden sollen einige Klassifikationsverfahren vorgestellt und deren Eigenschaften bezüglich der Erzeugung von kompakten Klassen eingeschätzt werden.

Hierarchische Verfahren

Sei $S = \{1, \dots, N\}$ eine Menge von N Objekten. Ein System \wp aus Teilmengen A, B, C, \dots von S heißt Hierarchie, wenn für zwei beliebige Mengen A, B aus \wp mit ($A \neq B$) höchstens eine der drei folgenden Eigenschaften gilt:

$$A \cap B = \emptyset \text{ oder } A \subset B \text{ oder } A \supset B \quad (4)$$

Die Mengen A, B, C, \dots heißen dann Klassen der Hierarchie \wp .

Ein Index zur Hierarchie \wp ist eine für alle Klassen $A \in \wp$ definierte, nichtnegative Funktion $h(A)$ derart, daß die Ungleichung

$$h(B) < h(A) \quad \forall A, B \in \wp : A \subset B \quad (5)$$

gilt. Außerdem sei $h(A) = 0$ genau dann, wenn A nur äquivalente Objekte enthält. Eine Hierarchie \wp , versehen mit einem Index $h(\cdot)$ heißt indizierte Hierarchie bzw. Dendrogramm. Die Funktion h wird dabei in der Regel durch eine Grenzdistanz repräsentiert.

Typische hierarchische Verfahren sind:

- *Single Linkage*, bei welcher alle Objekte miteinander verknüpft werden, deren Distanz kleiner als die vorgegebene Grenzdistanz ist (Klassendistanzen über die Definition des ähnlichsten Objektpaares)
- *Complete Linkage*, welche direkt auf die Klassendistanzen auf Basis des unähnlichsten Objektpaares basiert (Die Klassendistanz bestimmt sich aus dem maximalen Abstand zwischen den jeweiligen Objekten der Klassen.)
- *Average Linkage*, bei welcher die Verknüpfung auf Basis der mittleren Distanz ihrer Objekte erfolgt

Die genannten Verfahren weisen bezüglich der Kompaktheit der resultierenden Klassen erhebliche Unterschiede auf. So werden durch Single Linkage i. d. R. keine kompakten Strukturen erzeugt. Die beiden anderen Verfahren erzeugen kompakte Strukturen, haben jedoch den Nachteil, daß sie zum Einen zu gleichgroßen⁵ Klassen tendieren und zum Anderen bei einer Verschmelzung von größeren Klassen (bei großen Grenzdistanzen) die Kompaktheit der resultierenden Klasse nicht mehr gewährleisten können.

Große Vorteile der hierarchischen Verfahren liegen in der Anschaulichkeit der Dendrogramme bei einer „Erstclusterung“ ohne Vorkenntnisse und - daraus resultierend - in der Möglichkeit, diese Dendrogramme zur Festlegung einer geeigneten Klassenanzahl zu nutzen. Deshalb werden diese Verfahren oftmals als Start-Clusterverfahren gewählt, die anschließend mit einem weiteren (z. B. iterativen) Verfahren verbessert werden.

³ Bock [2] S. 24ff

⁴ Bock [2] S. 81ff

⁵ in Bezug auf ihre radiale Ausdehnung

Iterative Clusterverfahren

Bei iterativen Clusterverfahren wird i. Allg. auf eine Anfangsklassifikation - eine sogenannte Anfangspartition \wp^0 aufgesetzt. Bei der iterativen Bestimmung besserer Partitionen $\wp^1, \wp^2, \dots, \wp^v, \dots$ ($v = 0, 1, 2, \dots$) wird von einem vorher festgelegten Gütekriterium $g(\wp)$ ausgegangen. Dabei muß $g(\wp^{v+1}) < g(\wp^v)$ gelten.

Die gebräuchlichsten Methoden der iterativen Clustering sind die Austauschverfahren, bei denen bei Erfüllung eines bestimmten Kriteriums ein Objekt O_k von einer Klasse A_i in die Klasse A_j wechselt. Ein typisches Kriterium, ist das Varianzkriterium:

$$O_k \in A_i \rightarrow O_k \in A_j \Leftrightarrow \frac{n_j}{n_j + 1} \|x_k - \bar{x}_{A_j}\|^2 < \frac{n_i}{n_i + 1} \|x_k - \bar{x}_{A_i}\|^2. \quad (6)$$

Eine etwas anderer Ansatz wird durch die Iterierte-Minimal-Distanz-Partition verfolgt. Dort gilt für eine Objektmenge S und für die Klassen A_i der Partition v :

$$A_i^v = \left\{ O_k \mid O_k \in S \text{ und } D_{A_i^v, k} = \min_{1 \leq j \leq m} \left\{ D_{A_j^{v-1}, k} \right\} \right\} \quad (7)$$

Die Klassenanzahl m wird dabei vorgegeben. Wird jedoch noch eine Mindestanzahl an Objekten pro Klasse festgelegt, so kann das zur Auflösung einzelner Klassen führen und erlaubt damit eine Reduzierung der eigentlich festgelegten Klassenanzahl m . Bei der Iterierten-Minimal-Distanz-Partition ist es darüber hinaus möglich, gleich mehrere Objekte in einem Iterationsschritt zu tauschen.

Der Vorteil der iterativen Verfahren ist die Konstruktion von kompakten Klassen – die Nachteile bestehen in der Auffindung einer Anfangsklassifikation und der Festlegung einer Klassenanzahl m . Bestimmt man keine Anfangspartition, dann ist zumindest eine Menge von zentralen Objekten bzw. Kernpunkten entsprechend der Klassenanzahl m festzulegen.

Eine weitere Möglichkeit, Klassenstrukturen zu verbessern, ist der *rekursive Aufbau von Gruppen um Kerne*⁶. Dabei können große (nicht kompakte) Gruppierungen, wie sie beispielsweise bei einer Single Linkage Clustering entstehen, in kleinere kompakte Mengen zerlegt werden, ohne daß die ursprüngliche Klassenstruktur verloren geht. Nachteilig hierbei ist, daß oftmals eine Vielzahl von kleinen Klassen entsteht, die keine oder nur wenig Bedeutung für die Klassifikation besitzen bzw. für eine sinnvolle Klassenstruktur wieder zu größeren Klassen verschmolzen werden müßten.

Fuzzy Clusterverfahren

Das bekannteste und verbreitetste unscharfe Clusterverfahren ist der Fuzzy-c-Means Algorithmus [6]. Bei diesem Verfahren wird der Ansatz verfolgt, eine Optimalitätsaufgabe mit der Gütefunktion

$$J_{\text{FCM}}(\mathbf{U}, \mathbf{V}) = \sum_{k=1}^N \sum_{i=1}^c (\mu_{ik})^p \|x_k - v_i\|^2 \rightarrow \min_{\mu_{ki}} \quad (8)$$

unter der Nebenbedingung

$$\sum_{i=1}^c \mu_{ik} = 1 \quad \forall k = 1, \dots, N \quad (9)$$

mit den Parametern

- v_i ... Vektor des Klassenrepräsentanten
- μ_{ik} ... Zugehörigkeit des Objektes O_k zur Klasse A_i
- p ... freier Wichtungsparemeter (vorzugeben)
- c ... Klassenanzahl (vorzugeben)

zu lösen.

⁶ Bock [2], S.224ff

Dieser Ansatz führt über einen iterativen Algorithmus⁷ zu einer optimalen unscharfen Zuordnung der einzelnen Objekte zu den entsprechenden Klassen A_i , d.h., die Objekte werden nicht genau einer Klasse A_i , sondern mehreren Klassen mit unterschiedlichen Gewichtungen (Zugehörigkeiten) μ_{ik} zugeordnet.

Als nachteilig für eine automatische (fuzzy) Klassifikation stellen sich jedoch die folgende Aspekte heraus:

1. Es muß weiterhin ein Distanzmaß (Abstandsnorm) für alle Klassen vorgegeben werden.
2. Um die Optimalitätsaufgabe lösen zu können, muß die Nebenbedingung (15) erfüllt sein, die i. Allg. in der Fuzzy Theorie nicht angewandt wird, da sie zu einer Abhängigkeit der Klassen untereinander führt, die in der Fuzzy Theorie eigentlich nicht vorhanden ist.
3. Es muß die Anzahl der Klassen c vorgegeben werden.
4. Es muß eine Startkonfiguration (Zugehörigkeitsmatrix oder die Clusterzentren) vorgegeben sein, um den Iterationsprozeß starten zu können.

Diese 4 Punkte führen dazu, daß für die Clusteranalyse mittels Fuzzy-c-Means ein erhebliches Vorwissen notwendig ist, das in vielen Anwendungsfällen nicht vorliegt oder durch verschiedene Verfahren erst gewonnen werden muß. So können die Punkte 3. und 4. beispielsweise durch eine Startklassifikation oder durch eine Startpartition über die Bildung von Kernklassen bzw. Kernobjekten revidiert werden, wobei der Nachteil erhalten bleibt, daß Klassen nicht aufgelöst werden können.

Problematisch ist weiterhin, daß bei Nutzung von Fuzzy Pattern Strukturen trotzdem erst ein scharfer Klassifikator erzeugt werden muß, bevor eine unscharfe Aggregation erfolgen kann.

Verfahren zum Auffinden von zentralen Objekten

Die Verfahren zum Auffinden von zentralen Objekten bzw. von Kernpunkten stellen eine eigene Kategorie der automatischen Klassifikationsverfahren dar, da sie in der Regel nur grobe Schätzungen der Strukturen erlauben. Trotzdem sind sie auf Grund ihrer Einfachheit und der Möglichkeit einer objektiven Bereitstellung von Startstrukturen weit verbreitet. Die Idee beruht auf dem Ansatz, daß Kernobjekte bzw. Kernpunkte in ihrer Umgebung die größte Zahl von Objekten aufweisen müssen. Dazu wird in der Regel eine Grenzdistanz festgelegt, innerhalb der für jedes Objekt ausgezählt wird, wieviel weitere Objekte innerhalb dieser Grenzdistanz bezüglich des Kernpunktes liegen. Diese Zahl stellt ein Maß für die zentrale Lage eines Objektes dar⁸ und kann für eine Ordnung der Objekte unter diesem Aspekt dienen. Problematisch ist hierbei die Festlegung der Grenzdistanz, die gleichzeitig auch die Größe des zentralen Bereiches kennzeichnet, der einem zentralen Objekt zugeordnet wird. Man unterscheidet bei dieser Art der Clusterung verschiedene Ergebnisse:

- Kernpunkte – das sind Objekte mit „zentraler“ Bedeutung für Objekthäufungen
- Kernklassen – das sind Objekthäufungen um zentrale Punkte oder Kernpunkte
- Zentrale Punkte – das sind Repräsentanten von Objekthäufungen, die nicht notwendigerweise mit Objekten zusammenfallen

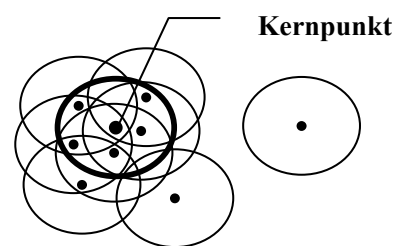


Abb. 6: Kernobjekt, Kernklasse

Für die Bestimmung einer Startlösung für den Fuzzy-c-Means Algorithmus wurde von Barone und Yager eine sogenannte Mountain-Method [7] entwickelt, bei der mögliche zentrale Punkte auf einem Gitter vorgegeben und anhand einer „mountain function“ einen Wert zugewiesen bekommen, der die Objektdichte um diesen Gitterpunkt wiedergibt. Dabei werden auf L_r -Distanzen verwendet, die über eine exponentielle Wichtung an jedem Gitterpunkt aufsummiert werden und somit ein „Dichte-Gebirge“ erzeugen. Die Gitterpunkte mit der höchsten Dichte werden dann sukzessive entfernt. Damit erhält man beispielsweise eine ausreichend gute Startlösung bezüglich der Klassenkerne für den Fuzzy-C-Means Algorithmus.

⁷ Bezdek [6], S.69ff

⁸ Oftmals wird hier auch von einer Dichte gesprochen. In Erweiterungen werden auch über Wichtungsfunktionen bezüglich der Distanzen zu den einzelnen Objekten sogenannte Objektdichten berechnet.

3.3 Methoden der automatischen Klassifikation in der Zeitreihenanalyse

Für die Erstellung von Fuzzy Prototypen gilt es vollständige Zeitverläufe über einem Zeitabschnitt anhand von Ähnlichkeiten zusammenzufassen. Dabei sind insbesondere die Zeitabhängigkeit und die Dynamik der Zeitreihen zu beachten.

Bestehende Vorgehensweisen beruhen dabei auf den folgenden zwei Ansätzen:

Konvertierung zu statischem Clusterproblem

Dieser Ansatz besteht in der Konstruktion eines erweiterten Merkmalsraumes durch eine orthogonale Aneinandersetzung der Merkmalsräume über den einzelnen Zeitpunkten. Damit wird es möglich, die herkömmlichen Klassifikationsverfahren anzuwenden. Darüber hinaus ist die Einarbeitung von zusätzlichen Merkmalen wie Momente, Gradienten, Varianzen usw. möglich. Ein großes Problem besteht jedoch darin, daß alle Merkmale gleichgewichtet nebeneinander stehen und der zeitliche Zusammenhang keine Rolle spielt. So können beispielsweise Zeitreihen, die über dem gesamten Zeitverlauf geringe Unterschiede zueinander aufweisen, innerhalb des Merkmalsraumes zu einer größeren Distanz führen als Zeitreihen, die sich in einem kleinen Zeitabschnitt signifikant unterscheiden.

Sequentielle Hierarchische Clusterung [9]

Ausgehend von der Methode der Single Linkage, Complete Linkage oder auch anderen hierarchischen Clustermethoden wird die automatische Klassifikation mit einer entsprechenden Grenzdistanz⁹ für jeden Zeitpunkt einzeln durchgeführt. Eine Gruppierung der Zeitreihen zueinander erfolgt dabei nur, wenn die „Ähnlichkeit“ für jeden Zeitpunkt erfüllt ist. Problematisch ist hierbei jedoch die Trennung von zueinander sehr ähnlichen Zeitreihen, die beispielsweise zu einem Zeitpunkt eine große Distanz aufweisen, die möglicherweise eine Fehlmessung oder eine Störung hervorgerufen hat.

Beide Methoden werden als Basis für eine Modellbildung in der Zeitreihenanalyse angewandt. Es bedarf jedoch in der Regel eines hohen Aufwandes, um Fehler zu beheben und die Klassenstrukturen zu optimieren.

3.4 Eine Methode auf Basis der Fuzzy Pattern Klassifikation

Ausgehend von den Überlegungen von Kapitel 3.2 bestand die Zielstellung darin, eine Clustermethode zu entwickeln, die folgende Eigenschaften erfüllt:

1. unscharfe Zuordnung der Objekte zu den Klassen
2. Bereitstellung einer unscharfen Klassenstruktur auf Basis der FPK
3. objektivierte Berechnung der Klassenanzahl
4. Betrachtung der Einzelobjekte als unscharfe Elementarklassen

Dazu gilt es ein Clusterverfahren innerhalb der Theorie der FPK zu entwickeln, welches auf bekannte Methoden sowohl der Clusteranalyse als auch der Fuzzy Pattern Klassifikation aufsetzt. Die grundlegende Idee war, ähnlich der Iterierten-Minimal-Distanz-Partition, ein Verfahren zu entwickeln, welches die Änderung einer Klassenzugehörigkeit nicht über ein Distanzmaß sondern über die unscharfe Zuordnung der einzelnen Objekte zu den Klassen (ihrem Zugehörigkeits- bzw. Sympathievektor) realisiert. Der Ansatz soll Klassenstrukturen erzeugen, wie sie als Ergebnis der Aggregation entstehen. Erste Gedanken zu einem solchem Verfahren sind unter [8] zu finden.

Analog zur Iterierten-Minimal-Distanz-Partition ist auch für diese Methode eine Startclustering notwendig, die im Prinzip der Kernklassen- bzw. Kernobjekt-Bestimmung¹⁰ ähnelt. Dabei werden mittels der elementaren Zugehörigkeitsfunktionen der Objekte die Kernobjekte oder auch Kernklassen bestimmt, wobei man Summenzugehörigkeiten eines Objektes zu allen anderen Objekten als Auswahlkriterium verwendet.

Beide (Teil-) Verfahren setzen direkt aufeinander auf, so daß man ein ganzheitliches unscharfes Klassifikationsverfahren erhält. Damit ist die Zielsetzung der direkten Bestimmung eines Fuzzy Pattern Klassifikators aus einer unscharf beschriebenen Objektstruktur (Elementarklassen) bei einer

⁹ absolut oder relativ zur bekannten Ausdehnung der Merkmale zu den einzelnen Zeitpunkten

¹⁰ Kapitel 3.2: *Verfahren zum Auffinden von zentralen Objekten*

variablen, objektiv zu bestimmenden Klassenanzahl (je nach vorgegebener elementarer Unschärfe und Objektstruktur) erfüllt.

Die vorzugebenden Parameter des Verfahrens sind die elementare Unschärfe c_e und gegebenenfalls eine scharfe Abweisungsschwelle, die es ermöglicht, Objekte mit einer sehr geringen Zugehörigkeit zu einer Klasse als nicht zugehörig zu definieren. Auf diese Weise ist es auch möglich, Objekte als Ausreißer zu kennzeichnen, vorausgesetzt, sie bilden keine eigene semantische Klasse.

Im Folgenden soll der Algorithmus dargestellt und anhand eines einfachen Beispiels illustriert werden.

3.4.1 Startverfahren der automatischen Fuzzy Klassifikation

Zielstellung dieses Startverfahrens ist die Bestimmung einer Startstruktur mittels der unscharf beschriebenen Objekte. Die Vorgehensweise besteht aus den folgenden Schritten:

Initialisierungsschritt:

- Festlegung einer elementaren Unschärfe c_e und damit der elementaren Zugehörigkeitsfunktion
- Festlegung einer Abweisungsschwelle

Wenn man Informationen über die einzelnen Merkmalsausprägungen hat, ist es möglich, eine elementare Unschärfe für jede Merkmalsrichtung anzugeben. Anderenfalls erweist es sich als günstig, eine relative elementare Unschärfe bezüglich aller Merkmalsausprägungen¹¹ festzulegen.

1. Schritt:

- Berechnung einer „Höhe“ für alle Objekte O_r

$$h_r = h(\underline{x}_r) = \sum_{k=1}^N \mu_1(\underline{x}_k) = \left(1 + \frac{1}{p} \sum_{i=1}^p \left(\frac{|x_{ki} - x_{ri}|}{c_{e_i}} \right)^2 \right)^{-1} \quad (10)$$

2. Schritt:

- Auswahl des Objektes O_R mit maximaler Höhe

$$\underline{x}_R = \left\{ \underline{x}_r : h(\underline{x}_r) = \max_k (h(\underline{x}_k)) \right\} \quad (11)$$

3. Schritt:

- Elimination aller Objekte O_k mit $\mu_R(\underline{x}_k) \geq$ Abweisungsschwelle
- weiter bei 1. Schritt

$$\mu_R(\underline{x}_k) \geq \text{Abweisungsschwelle} \quad (12)$$

Mit diesem Algorithmus erhält man eine Menge von Kernobjekten, die mittels der unscharfen Elementarbeschreibungen als Startklassifikation dienen kann. Eine Alternative dazu wäre, Klassen zu aggregieren, die aus den, dem Kernobjekt zugeordneten und im Schritt 3 eliminierten Objekten bestimmt wird. Damit entstehen unscharf beschriebene Kernklassen.

¹¹ Die Merkmalsausprägung wird über das Intervall definiert, in dem sich die entsprechenden Merkmalswerte der Objekte bewegen.

3.4.2 Iterationsverfahren der automatischen Fuzzy Klassifikation

Diese Vorgehensweise stellt das eigentliche iterative Fuzzy Klassifikationsverfahren auf Basis der FPK dar. Die Vorgehensweise besteht aus folgenden Schritten:

Initialisierungsschritt:

- Übernahme einer Startstruktur¹²
- Festlegung der elementaren Unschärfe c_e ¹³ sowie einer Abweisungsschwelle (optional)¹⁴
- Festlegung einer Mindestobjektanzahl¹⁵

1. Schritt:

- Aufbau des Fuzzy Klassifikators aus der aktuellen Klassenstruktur der Objekte

2. Schritt:

- Identifikation [1][5] aller Objekte
Berechnung eines Sympathievektors μ_k für jedes Objekt O_k bezüglich des Klassifikator, wobei das Objekt selbst aus dem Klassifikator entfernt wird (cross-validation).

3. Schritt:

- Abweisung von Objekten
Sind alle Sympathiewerte eines Objektes kleiner als die Abweisungsschwelle, dann erfolgt keine Klassenzuordnung - das Objekt wird abgewiesen

4. Schritt:

Es werden nur die Objekte betrachtet, deren Klassenzuordnung bezüglich des höchsten Sympathiewertes nicht mit der ursprünglichen Klasse übereinstimmt:

- existieren keine derartigen Objekte \Rightarrow 5. Schritt (ENDE)
- Berechnung des Überdeckungsgrades [1] für jedes Objekt O_k ,

$$g_k = \frac{\mu_k \text{ neue Klasse}}{\mu_k \text{ alte Klasse}} \quad (13)$$

- alle Objekte O_k mit $g_k \geq 1$ wechseln zur Klasse mit dem höchsten Sympathiewert; existieren keine derartigen Objekte (alle $g_k < 1$) dann wechselt das Objekt mit dem höchsten g_k
- weiter bei 1. Schritt

5. Schritt:

Ende des Iterationsverfahrens

Als Ergebnis dieser Klassifikation erhält man einen Fuzzy Pattern Klassifikator, der über alle Merkmale die Objektstruktur in einer unscharfen Weise wiedergibt. Für den Erhalt einer scharfen Objektklassifikation muß lediglich eine Identifikation [1][5] der dem Klassifikator zugrundeliegenden Objektmenge erfolgen und beispielsweise durch einen Maximumsentscheid bezüglich der Sympathiewerte eine Klassenzuordnung erfolgen.

¹² Alternativ zum Startverfahren der Fuzzy Klassifikation kann auch ein Ergebnis der Verfahren aus Kapitel 3.2 dienen.

¹³ Im Allgemeinen sollte hier die elementare Unschärfe des Startverfahrens übernommen werden.

¹⁴ Eine Festlegung der Abweisungsschwelle auf 0 ist dabei möglich und führt zu keiner Beschränkung

¹⁵ Eine Festlegung der Mindestobjektanzahl auf 1 führt dabei zu keiner Beschränkung.

3.4.3 Beispiele

In einem ersten Beispiel soll eine 3-Klassen-Struktur erkannt werden. Dabei besteht eine Gleichverteilung der Objekte innerhalb der Klassen. Das Ergebnis zeigt, bei einer Einstellung der relativen elementaren Unschärfe von 6%, einer Abweisungsschwelle von 0,1 und einer Mindestobjektzahl pro Klasse von 5, einen Fuzzy Pattern Klassifikator, der genau diese 3 Klassen beschreibt (Abb. 7 und 8).

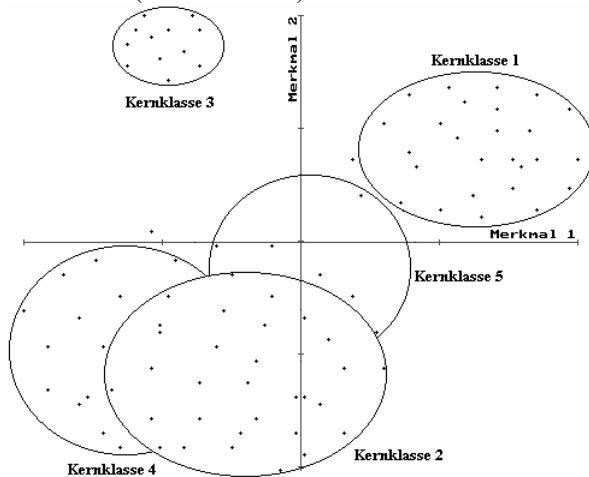


Abb. 7: Kernklassen des 3-Klassen-Problems

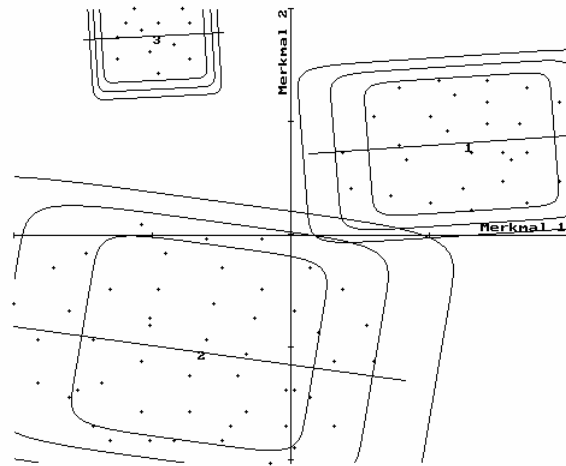


Abb. 8: Resultat des 3-Klassen-Problems

Ein weiteres Beispiel besteht aus zwei Klassen mit Objekthäufungen an einer Art Scheitelpunkten, die auch gleichzeitig die kürzeste Verbindung zwischen den Klassen darstellen. Für eine Einstellung einer relativen elementaren Unschärfe von 6%, einer Abweisungsschwelle von 0,1 und einer Mindestobjektzahl pro Klasse von 5 werden diese zwei Klassen durch jeweils eine entsprechende Fuzzy Klasse beschrieben (Abb. 9 und 10).

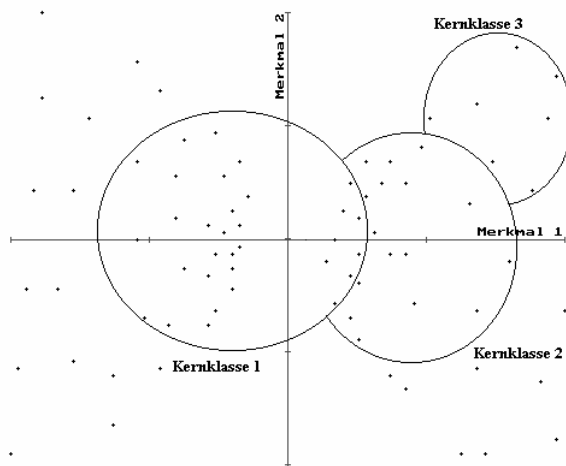


Abb. 9: Kernklassen des 2-Klassen-Problems

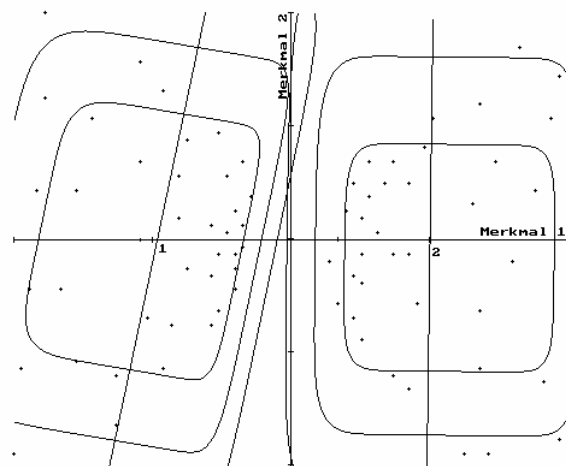


Abb. 10: Resultat des 2-Klassen-Problems

Die ersten beiden Beispiele stammen aus einer Serie von Testdatensätzen für Clustermethoden¹⁶ und können von Verfahren wie Complete Linkage oder Iterierter-Minimal-Distanz-Partition ebenfalls gut gelöst werden.

Für ein weiteres Beispiel wurden 4 Klassen erzeugt, deren Objekte eine Normalverteilung aufweisen. Für dieses Beispiel konnte die Fuzzy Clusterung bei einer Einstellung der relativen elementaren Unschärfe von 5%, einer Abweisungsschwelle von 0,1 und einer Mindestobjektzahl pro Klasse von 15 die entsprechenden 4 Fuzzy Klassen generieren (Abb. 11 und 12).

¹⁶ siehe auch Bocklisch [1]

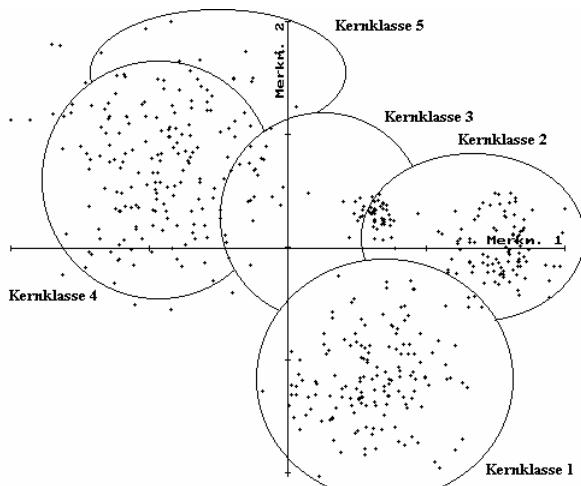


Abb. 11: Kernklassen des 4-Klassen-Problems

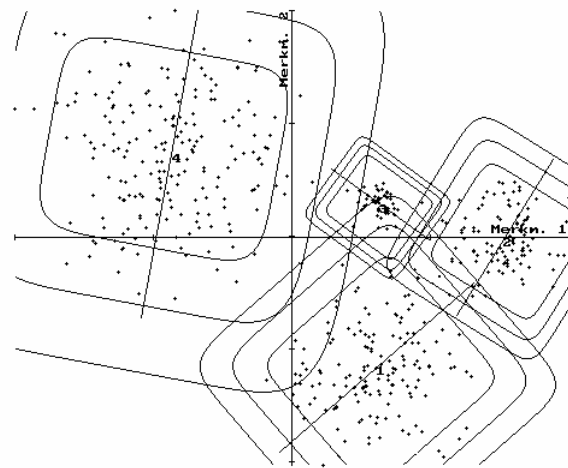


Abb. 12: Resultat des 4-Klassen-Problems

Bei Anwendung der Complete Linkage (Abb.13) und der Iterierten-Minimal-Distanz-Partition¹⁷ (Abb. 14) auf diesen Datensatz konnten nur unzureichende Ergebnisse festgestellt werden. Beide Verfahren tendieren zu gleich großen Objektmengen und die Complete Linkage kann durch die hierarchische Klassenverschmelzung die Kompaktheit nicht gewährleisten.

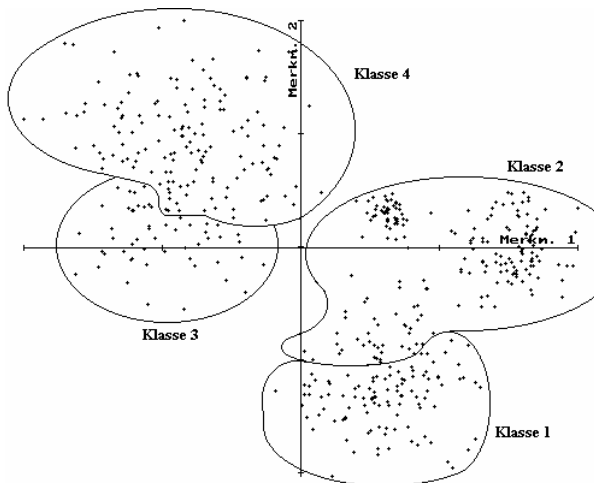


Abb. 13: Complete Linkage für das 4-Klassen-Problem

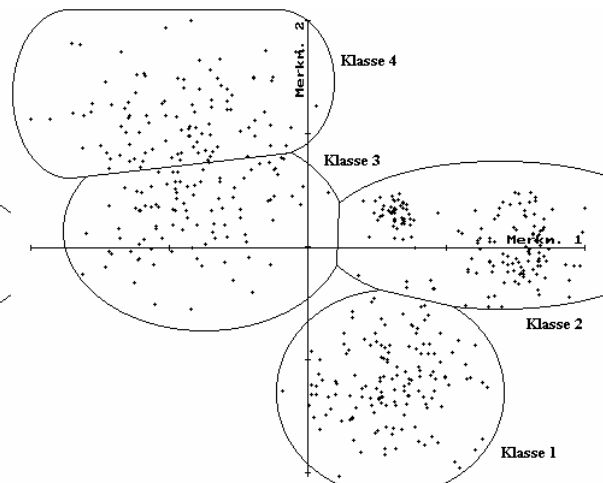


Abb. 14: Iterierte-Minimal-Distanz-Partition für das 4-Klassen-Problem

3.4.4 Die Fuzzy Klassifikationsmethode für Zeitreihen

Für die Anwendung der Fuzzy Klassifikationsmethode zur Erstellung eines quasi-statischen Modells kann der Algorithmus in der beschriebenen Weise verwendet werden.

Für die Problematik eines dynamischen Fuzzy Pattern Klassifikators müssen Anpassungen vorgenommen werden, da die Einzelklassen für diesen Fall komplette Fuzzy Prototypen darstellen. Die Rolle der Fuzzy Klassen werden von den Fuzzy Prototypen übernommen. Die Bestimmung der Zugehörigkeit (2. Schritt des Iterationsverfahrens) einer Zeitreihe zu einem Prototypen wird dabei über die resultierenden Sympathien der Einzelklassen zu den einzelnen Zeitpunkten bestimmt und über dem gesamten Zeitintervall mittels einem unscharfen UND-Operator zu einem resultierenden Zugehörigkeitswert zusammengefaßt. Auf dieser Basis werden auch die Überdeckungsgrade und die Abweisungsschwelle behandelt.

Als Beispiel soll aus der Energielaststudie [4] ein Klassifikationsmodell „Wochentag-Winter“ erstellt werden. Die Basis bilden hier 57 gemessene Zeitverläufe (Abb. 15).

¹⁷ Startklassifikation über Complete Linkage; die Klassenzahl wurde mit 4 vorgegeben

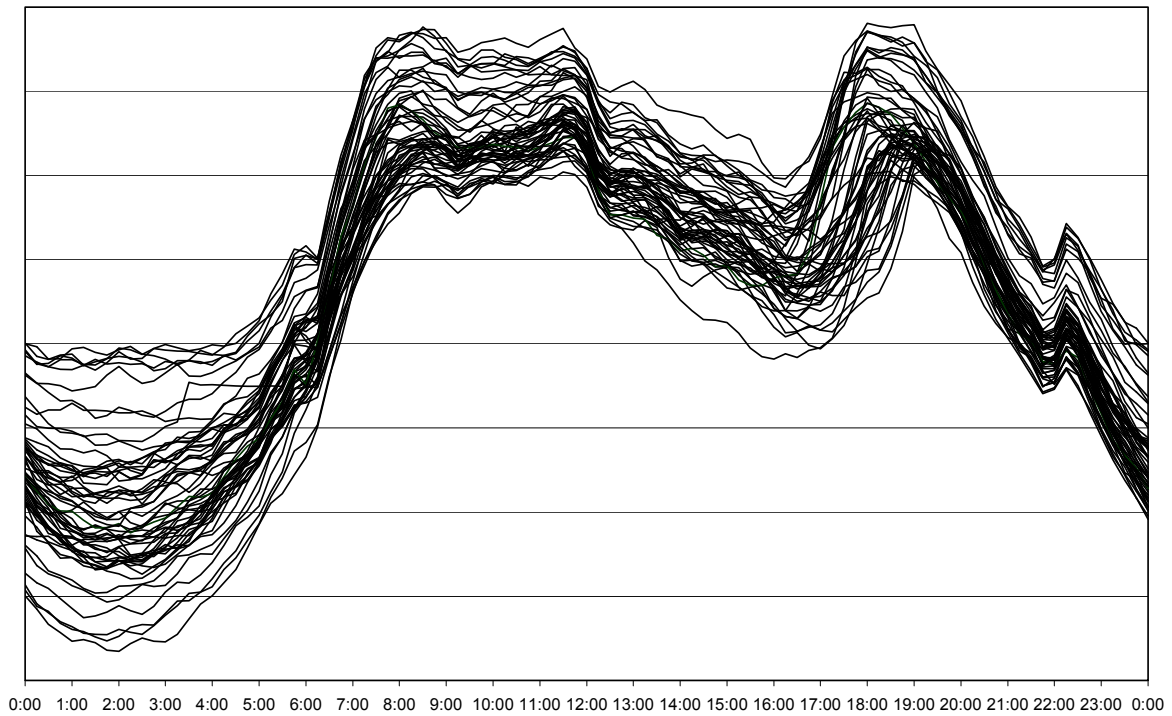


Abb. 15: Zeitverläufe für „Wochentag-Winter“

Mittels der Fuzzy Klassifikation für Zeitreihen konnten diese Zeitreihen für eine relative elementare Unschärfe von 5% einer Abweisungsschwelle von 0,1 und einer Mindestobjektzahl¹⁸ pro Klasse von 1 zu 5 Fuzzy Prototypen (Abb. 16) zusammengefasst werden. Dabei erfolgte die Verknüpfung der Zeitpunkte untereinander mittels eines unscharfen UND-Operators. 5 Zeitreihen wurden abgewiesen.

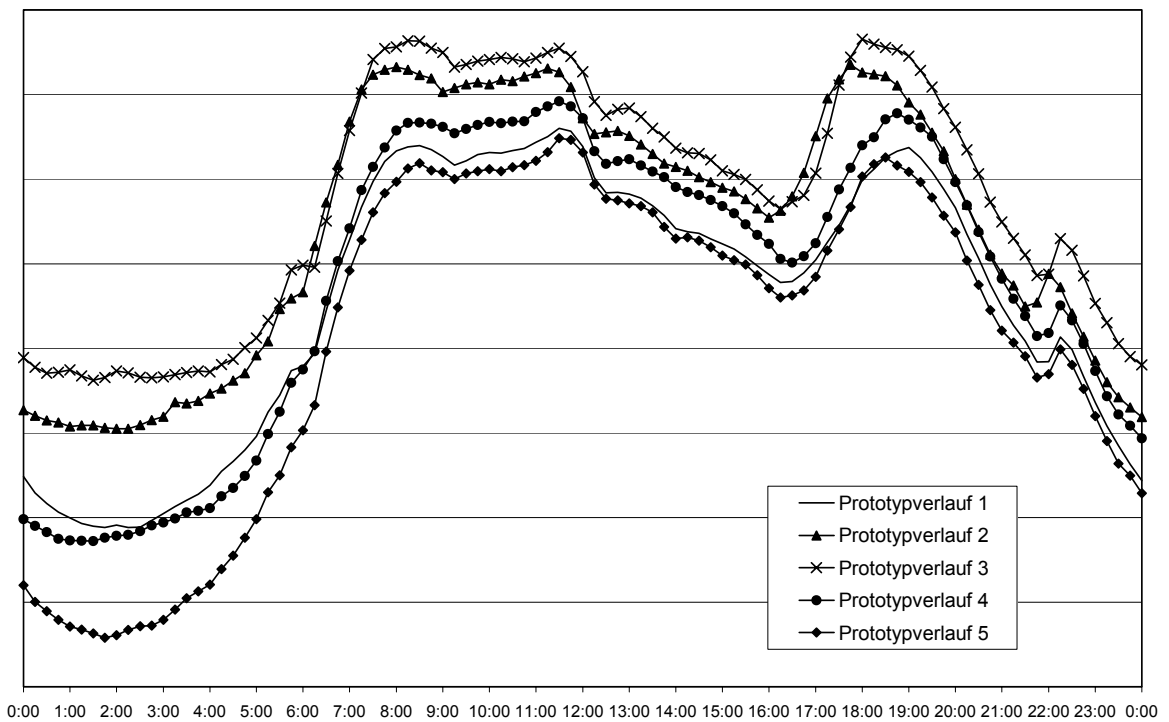


Abb. 16: resultierende Prototypen (5 Prototypverläufe)

¹⁸ hier repräsentiert ein Objekt einen vollständigen Tagesgang

Für die Bewertung des Prototypmodells wurden alle Zeitreihen anhand des Modells prognostiziert. Für die $\frac{1}{4}$ h, die $\frac{1}{2}$ h und die 2 h Prognose sind in der Tabelle 1 die Häufigkeiten, mit der die entsprechenden Fehlergrenzen eingehalten werden konnten, aufgeführt.

Fehler	1/4h	1h	2h
bis 0,5 %	49,2 %	37,1 %	28,2 %
bis 1,0 %	79,2 %	63,9 %	52,4 %
bis 1,5 %	92,6 %	79,8 %	68,6 %
bis 2,0 %	96,9 %	87,9 %	79,3 %
bis 2,5 %	98,6 %	92,7 %	85,7 %
bis 3,0 %	99,3 %	95,4 %	90,0 %
bis 3,5 %	99,6%	97,1 %	93,1 %
bis 4,0 %	99,8 %	98,0 %	94,9 %
bis 4,5 %	100,0 %	98,6 %	96,2 %
bis 5,0 %	100,0 %	99,1 %	97,4 %

Tab. 1: Güte des Prototypmodells „Wochentag-Winter“

Für die Energielaststudie [4] wurden für einen experten-basierten Klassifikator sehr ähnliche Ergebnisse erreicht. Die Abweichungen in den Häufigkeiten liegen zwischen $-1,5\%$ bis $3,5\%$, wobei die Verbesserungen vor allem bei der 2 h Prognose auftreten.

Literatur

- [1] S. F. Bocklisch, „Prozeßanalyse mit unscharfen Verfahren“, Verlag Technik Berlin, 1987
- [2] H. H. Bock, „Automatische Klassifikation“, Vandenhoeck & Ruprecht, Göttingen, 1974
- [3] M. Päßler, O. Amft, S.F. Bocklisch, W. Hiller, „Kurzeitprognose der Leistung von Windkraftanlagen mit Fuzzy Pattern Klassifikation“, Tagungsband der 5.Deutschen Windenergiekonferenz, DEWEK 2000
- [4] S. F. Bocklisch, M. Päßler, „Fuzzy Time Series Analysis“, Fuzzy Control - Theory and Practice, Editors: Rainer Hampel, Michael Wagenknecht, Nasredin Chaker, Physica-Verlag, Heidelberg, Series: Advances in Soft Computing, 2000
- [5] M. Päßler, „Mehrdimensionale Zeitreihenanalyse und -prognose mit Fuzzy Pattern Modellen“, Diplomarbeit, TU Chemnitz, Fakultät für Mathematik, 1998
- [6] J. C. Bezdek, „Pattern Recognition with Fuzzy Objective Function Algorithms“, Plenum Press, New York, 1981
- [7] J. M. Barone, R. R. Yager, „Mountain method-based fuzzy clustering: methodological considerations“, International Journal of General Systems, Band 23, Heft 4, 1995, S. 281-305
- [8] S. F. Bocklisch, W. Meyer, B. Straube, „Fuzzy Clustering and Classification for the Identification of Process Types in Technical and Biomedical Systems“, Kennwertermittlung und Modellbildung, Teil II, TH Karl-Marx-Stadt, 1979
- [9] L. I. Porges, „Zeitreihenmodellierung und Prognose mittels Fuzzy Pattern Klassifikation“, Diplomarbeit, TU Chemnitz, Lehrstuhl für Systemtheorie, 1995

Forschungsaufenthalte und Konferenzbesuche:

5. Deutsche Windenergiekonferenz 2000, Wilhelmshaven

2001 European Wind Energy Conference and Exhibition, Kopenhagen Dänemark

Publikationen:

M. Päßler, S. F. Bocklisch, „Fuzzy Time Series Analysis“
Fuzzy Control - Theory and Practice,
Editors: Rainer Hampel, Michael Wagenknecht, Nasredin Chaker, Physica-Verlag, Heidelberg
Series: Advances in Soft Computing, 2000

M. Päßler, O. Amft, S. F. Bocklisch, W. Hiller, „Kurzeitprognose der Leistung von Windkraftanlagen mit Fuzzy Pattern Klassifikation“
Tagungsband der 5. Deutschen Windenergiekonferenz, DEWEK 2000