

Eine Anwendung des Algorithmus von Ford und Fulkerson bei der Interpretation von Infrarotspektren

Martin Sonntag

Bergakademie Freiberg, Fachbereich Mathematik
Bernhard-von-Cotta Str. 2, D-09599 Freiberg

(received: October 1993)

1 Vorbemerkungen

Es wird der aus der Graphentheorie bekannte Algorithmus von Ford und Fulkerson zur Lösung des Maximalstromproblems in Transportnetzen vorgestellt. Dieser Algorithmus ist auf eine Vielzahl von Problemen anwendbar, unter anderem auch auf Zuordnungsprobleme verschiedenster Art. Eine solche Anwendung wird anhand eines einfachen Systems zur Interpretation von Infrarotspektren demonstriert.

Die Arbeit entstand während eines Forschungsaufenthalts an der Universität Bayreuth, der durch das Förderprogramm der Deutschen Akademie der Naturforscher – Leopoldina – ermöglicht wurde.

Die Idee zur Anwendung des Algorithmus von Ford und Fulkerson auf die Interpretation von Infrarotspektren verdanke ich Herrn Prof. Dr. R. Laue. Ihm und Herrn Prof. Dr. A. Kerber möchte ich für ihr Interesse an der Thematik und ihre ständige Diskussionsbereitschaft herzlich danken.

2 Graphentheoretische Grundlagen

Wir benutzen im folgenden graphentheoretische Standardterminologie, sämtliche hier nicht explizit erklärten Begriffe findet man in Harary [H], Sachs [S], Walther [W], Walther/Nägler [WN], Wilf [Wi] oder auch Carré [C]. Mit \mathbf{N}^+ , \mathbf{Q}^+ bzw. \mathbf{R}^+ bezeichnen wir die Menge der positiven natürlichen, rationalen bzw. reellen Zahlen. Ein *schlichter gerichteter Graph (Digraph)* $G = (V, E)$ besteht aus einer nichtleeren, endlichen Menge $V = \{v_1, v_2, \dots, v_n\}$ ($n \in \mathbf{N}^+$) von sogenannten *Knoten* und einer (endlichen) Menge $E \subseteq (V \times V) - \{(v, v) \mid v \in V\}$ von geordneten Paaren von Knoten, den *Bögen* (oder *Kanten*) von G .

Beispiel 1:

$$\begin{aligned} G &= (V, E) \quad \text{mit} \\ V &= \{v_1, v_2, v_3, v_4\}, \\ E &= \{(v_1, v_2), (v_1, v_3), (v_2, v_1), (v_3, v_2), (v_3, v_4)\}. \end{aligned}$$

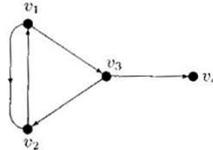


Abb. 1

Ein *Transportnetz* $N = (G = (V, E), c, q, s)$ besteht aus einem Digraphen G , zwei ausgezeichneten Knoten $q \in V$ (der *Quelle*) und $s \in V$ (der *Senke*) und einer sogenannten *Kapazitätsfunktion* $c: E \rightarrow \mathbf{R}^+ \cup \{0, \infty\}$, die jedem Bogen $(v, v') \in E$ eine nichtnegative *Kapazität* $c(v, v')$ zuordnet. (Wir vereinbaren, daß $x < \infty$ für jede reelle Zahl x gilt.)

Im folgenden sei $N = (G = (V, E), c, q, s)$ stets ein Transportnetz.

Definition. Eine Bogenfunktion $f: E \rightarrow \mathbf{R}$ heißt genau dann ein *Strom* der Stromstärke p auf N von der Quelle q zur Senke s , wenn die folgenden drei Bedingungen erfüllt sind:

- (1) Für alle Knoten $v \in V - \{q, s\}$ gilt die *Kirchhoffsche Strombedingung*

$$\sum_{\substack{v' \in V \\ (v, v') \in E}} f(v, v') - \sum_{\substack{v' \in V \\ (v', v) \in E}} f(v', v) = 0;$$

(2)
$$\sum_{\substack{v \in V \\ (q, v) \in E}} f(q, v) - \sum_{\substack{v \in V \\ (v, q) \in E}} f(v, q) = p;$$

(3)
$$\sum_{\substack{v \in V \\ (v, s) \in E}} f(v, s) - \sum_{\substack{v \in V \\ (s, v) \in E}} f(s, v) = p.$$

Verbal kann man die Kirchhoffsche Strombedingung (1) so formulieren:

Aus jedem Knoten (außer der Quelle q und der Senke s) fließt genausoviel heraus wie in ihn hineinfließt.

(2) und (3) beinhalten, daß aus der Quelle q genausoviel herausfließt, wie in die Senke s hineinfließt, nämlich gerade die Stromstärke p .

Das uns nun interessierende *Maximalstromproblem* besteht in der Suche nach einem mit den Kapazitäten *verträglichen* Strom f (d.h. $0 \leq f(v, v') \leq c(v, v')$ für alle $(v, v') \in E$) mit maximaler Stromstärke p .

Die Existenz eines mit den Kapazitäten verträglichen Stromes ist für beliebiges N gesichert, etwa ist der *Nullstrom* $f(v, v') = 0$ (für alle $(v, v') \in E$) ein solcher (mit

der Stromstärke $p = 0$). Da im allgemeinen jedoch auch unendliche Kapazitäten zugelassen sind, existiert nicht in jedem Falle ein Maximalstrom. Zur Formulierung eines entsprechenden Charakterisierungssatzes benötigen wir den Begriff des Schnittes:

Seien $A, B \subseteq V$ mit $A \cup B = V$, $A \cap B = \emptyset$, $q \in A$, $s \in B$.

Die Menge

$$S = S(A, B) := \{(v, v') \mid (v, v') \in E, v \in A, v' \in B\}$$

aller Bögen mit Startknoten in A und Zielknoten in B heißt ein *Schnitt* des Transportnetzes N .

Wir nennen

$$c_S := \sum_{(v, v') \in S} c(v, v')$$

die *Kapazität des Schnittes* S , wobei im Falle $S = \emptyset$ $c_S := 0$ gesetzt wird.

Dann gilt der

Satz von Ford und Fulkerson ([C], [S], [W], [Wi]).

Ein Transportnetz N besitzt genau dann einen Schnitt endlicher Kapazität, wenn für N ein Maximalstrom existiert. Besitzt N einen Schnitt endlicher Kapazität, so ist die minimale Kapazität aller möglichen Schnitte gleich der maximalen Stromstärke aller mit den Kapazitäten verträglichen Ströme.

3 Der Algorithmus von Ford und Fulkerson

Zur Bestimmung eines Maximalstromes in Transportnetzen existiert eine Vielzahl von Algorithmen; eine Darstellung einiger von ihnen findet man in [Wi], wo auch vergleichende Betrachtungen zu den unterschiedlichen Verfahren angestellt werden. In Abhängigkeit von den bei der Kapazitätsfunktion c neben ∞ auftretenden Zahlenwerten bezeichnen wir c als *reell*, *rational* bzw. *ganzzahlig*. In [Wi] wird anhand eines Gegenbeispiels gezeigt, daß – trotz Existenz eines endlichen Schnittes – der ursprüngliche Algorithmus von Ford und Fulkerson nicht für beliebige reelle Kapazitätsfunktionen c nach endlich vielen Schritten mit dem Erreichen eines Maximalstromes abbricht. Ist c jedoch ganzzahlig oder rational, so liefert der Algorithmus einen Maximalstrom, falls das betrachtete Transportnetz einen Schnitt endlicher Kapazität besitzt ([S]).

Da praktische Anwendungen in den seltensten Fällen für die Kapazitätsfunktion irrationale Werte erfordern und der Algorithmus von Ford und Fulkerson recht kurz und prägnant ist, wollen wir diesen hier vorstellen.

Sei $N = (G = (V, E), c, q, s)$ ein Transportnetz mit rationaler Kapazitätsfunktion c und f ein mit den Kapazitäten verträglicher (rationaler) Strom.

Der Algorithmus von Ford und Fulkerson besteht aus zwei Teilalgorithmen (zur Knotenmarkierung bzw. zur Stromverbesserung), die nacheinander sooft wie möglich abgearbeitet werden. Dabei wird der Strom iterativ verbessert. Besitzt das betrachtete Transportnetz einen Schnitt endlicher Kapazität, so bricht der Algorithmus nach endlich vielen Schritten bei Erreichen eines Maximalstromes ab.

Zur Markierung der Knoten von N benutzen wir eine Knotenfunktion $\mu: V \rightarrow \mathbf{Q}^+ \cup \{0, \infty\}$. Wir nennen einen Knoten $v \in V$ genau dann *markiert*, wenn $\mu(v) \neq 0$ ist. Man kann zeigen ([S]), daß der unten beschriebene Markierungsprozeß genau dann zur Markierung der Senke s führt, wenn der Strom f verbessert werden kann (und zwar um mindestens den Wert $\mu(s)$ der Markierung von s).

A: KNOTENMARKIERUNG

Schritt 1: Sei

$$\begin{aligned} \mu(q) &:= \infty && \text{und} \\ \mu(v) &:= 0 && \text{für alle } v \in V - \{q\}. \end{aligned}$$

Schritt 2: Sei $v \in V$ mit $\mu(v) > 0$. Wenn ein $v' \in V$ existiert mit

$$\begin{aligned} \mu(v') &= 0, \\ (v, v') &\in E \text{ und} \\ f(v, v') &< c(v, v'), \text{ so setze} \end{aligned}$$

$$\mu(v') := \min\{\mu(v), c(v, v') - f(v, v')\}.$$

Schritt 3: Sei $u' \in V$ mit $\mu(u') > 0$. Wenn ein $u \in V$ existiert mit

$$\begin{aligned} \mu(u) &= 0, \\ (u, u') &\in E \text{ und} \\ f(u, u') &> 0, \text{ so setze} \end{aligned}$$

$$\mu(u) := \min\{\mu(u'), f(u, u')\}.$$

Schritt 4: Wenn $\mu(s) = 0$ ist und es noch einen Knoten gibt, der mittels Schritt 2 oder Schritt 3 markiert werden kann, dann gehe zu Schritt 2.

B: STROMVERBESSERUNG

Fall 1: $\mu(s) \neq 0$.

Offensichtlich existiert eine Folge $W = (q = v_0, e_1, v_1, e_2, \dots, e_t, v_t = s)$ von paarweise verschiedenen Knoten $v_0, v_1, \dots, v_t \in V$ und Bögen $e_1, e_2, \dots, e_t \in E$, wobei sämtliche dieser Knoten markiert wurden und für $i = 1, 2, \dots, t$ gilt:

$$e_i = (v_{i-1}, v_i) \text{ oder } e_i = (v_i, v_{i-1}),$$

d.h.: beim Durchlaufen der Folge W von q nach s wird e_i im Sinne seiner Orientierung durchlaufen oder im entgegengesetzten Sinne.

Setze für $i = 1, 2, \dots, t$

$$f'(e_i) := \begin{cases} f(e_i) + \mu(s), & \text{falls } e_i = (v_{i-1}, v_i) \\ f(e_i) - \mu(s), & \text{falls } e_i = (v_i, v_{i-1}) \end{cases}$$

und

$$f'(e) := f(e) \text{ für alle } e \in E - \{e_1, e_2, \dots, e_t\}.$$

(Wie man sich überlegt, ist f' wieder ein Strom und zwar mit einer um $\mu(s)$ größeren Stromstärke als f .)

Setze

$$f := f'$$

und gehe zu **A**.

Fall 2: $\mu(s) = 0$.

Ende (f ist ein Maximalstrom).

Der interessierte Leser kann zu dem Algorithmus in [C], [S], [W], [WN] und [Wi] nachlesen.

4 Ein einfaches System zur Interpretation von Infrarotspektren

Neben der Lösung von Transportproblemen mittels des beschriebenen Algorithmus ist dessen Anwendung auf Zuordnungsprobleme von besonderer Bedeutung; ein bekanntes Beispiel dafür ist das der Tanzpaarung:

Bei einer Tanzveranstaltung sind m Damen und m Herren anwesend; von jedem der Anwesenden ist bekannt, mit welchen Tanzpartnern er gern den Abend verbringen würde. Gibt es eine (eindeutige) Zuordnung zwischen Damen und Herren, die den Wünschen aller gerecht wird?

Wenden wir uns nun jedoch dem für uns interessanteren Problem der Interpretation von Infrarotspektren zu. Dabei gehen wir im – zunächst betrachteten – einfachsten Fall von folgenden Prämissen aus:

Gegeben sind eine Menge $F = \{g_1, g_2, \dots, g_m\}$ von funktionellen Gruppen, eine Menge $I = \{i_1, i_2, \dots, i_n\}$ von Intervallen für die Lage der Banden (Intervalle für Wellenzahlen in cm^{-1}) und eine binäre Relation $R \subseteq F \times I$ mit der Eigenschaft

(E): $(g, i) \in R$ genau dann, wenn

i ein *charakteristisches Intervall* für die funktionelle

Gruppe g ist, d.h. bei Vorhandensein von g in der untersuchten chemischen Verbindung **muß** in i eine Bande vorliegen.

(Zur Einbeziehung von Intensitäten und Halbwertsbreiten siehe Abschnitt 5.)

Wir bilden zunächst den Digraphen $G = (V, E)$ mit

$$\begin{aligned} V &:= \{q, s\} \cup F \cup I && \text{und} \\ E &:= R \cup \{(g, g) \mid g \in F\} \cup \{(i, s) \mid i \in I\}. \end{aligned}$$

Nun führen wir folgende Kapazitätsfunktion c ein:

$$\begin{aligned} c(q, g) &:= \infty, && \text{für } g \in F; \\ c(g, i) &:= 1, && \text{für } g \in F, i \in I, (g, i) \in R. \end{aligned}$$

Für die restlichen Bögen hängt c von dem jeweils auszuwertenden Spektrum Sp ab; es sei für $i \in I$

$$c(i, s) := \begin{cases} \infty, & \text{falls } Sp \text{ im Intervall } i \text{ eine Bande besitzt} \\ 0, & \text{sonst.} \end{cases}$$

Damit ist $N = (G = (V, E), c, q, s)$ ein Transportnetz.

Beispiel 2:

$$\begin{aligned} F &= \{g_1, g_2, g_3, g_4\}; \\ I &= \{i_1, i_2, i_3, i_4, i_5\}; \\ R &= \{(g_1, i_1), (g_1, i_2), (g_1, i_3), (g_2, i_1), (g_2, i_2), (g_3, i_3), \\ &\quad (g_3, i_4), (g_4, i_5)\}; \end{aligned}$$

Banden wurden in i_1, i_2, i_4, i_5 festgestellt.

Damit erhalten wir das in Abb. 2 dargestellte Transportnetz (die an den Bögen in Klammern angegebenen Zahlen sind die Werte eines Maximalstromes, auf dessen Berechnung wir etwas später nochmals eingehen werden).

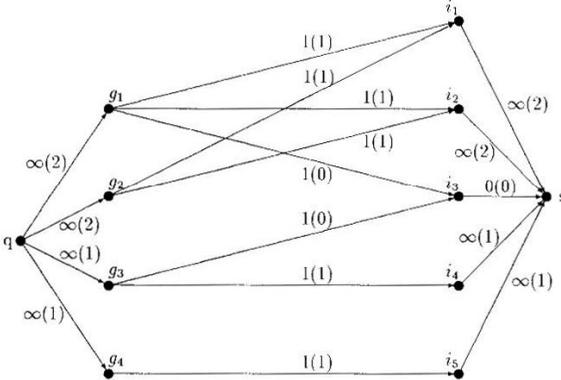


Abb. 2

Offensichtlich besitzt ein solches Transportnetz N - unabhängig vom auszuwertenden Spektrum - stets einen Schnitt endlicher Kapazität (etwa ist R ein solcher). Die Anwendung des Algorithmus von Ford und Fulkerson liefert daher einen Maximalstrom f . Da nach (E) eine funktionelle Gruppe $g \in F$ nur dann vorhanden sein kann, wenn alle ihrer charakteristischen Intervalle Banden enthalten, kann man aus f unmittelbar ablesen, ob g in der untersuchten Verbindung vorliegen kann (wir bezeichnen dann g kurz als eine *mögliche* funktionelle Gruppe) oder nicht: g ist genau dann eine mögliche funktionelle Gruppe, wenn gilt

$$(*) \quad f(q, g) = \sum_{\substack{i \in I \\ (g, i) \in R}} c(g, i).$$

Dabei ist

$$\sum_{\substack{i \in I \\ (g,i) \in E}} c(g,i) = \sum_{\substack{i \in I \\ (g,i) \in R}} c(g,i) = \text{Anzahl der für } g \text{ charakteristischen Intervalle}$$

und

$$f(g,g) = \sum_{\substack{i \in I \\ (g,i) \in E}} f(g,i).$$

Äquivalent zu (*) ist auch

$$\prod_{\substack{i \in I \\ (g,i) \in E}} f(g,i) > 0.$$

Es sei angemerkt, daß der oben angegebene Schnitt R im allgemeinen nicht minimal ist. Aufgrund der speziellen Struktur des Transportnetzes N existiert genau ein minimaler Schnitt, dieser hängt von dem auszuwertenden Spektrum ab und hat die Form

$$\{(g,i) \mid g \in F, i \in I, (g,i) \in E, f(g,i) > 0\} \cup \{(i,s) \mid i \in I, c(i,s) = 0\}.$$

Man erkennt, daß die Schnittkapazität - und damit auch die Stromstärke des Maximalstromes f - gleich der Anzahl der mit Banden belegten charakteristischen Intervalle aus I ist, wobei allerdings ein jedes $i \in I$ k -fach gezählt werden muß, wenn es für genau k funktionelle Gruppen aus F charakteristisches Intervall ist. Bedingt durch die sehr spezielle Struktur des oben konstruierten Transportnetzes N existiert in N sogar ein *eindeutig bestimmter* Maximalstrom f und dieser läßt sich durch eine erheblich vereinfachte "Variante" des Algorithmus von Ford und Fulkerson finden:

$$f(i,s) := \begin{cases} \sum_{\substack{g \in F \\ (g,i) \in R}} c(g,i), & \text{falls } c(i,s) = \infty \\ 0, & \text{sonst} \end{cases} \quad (i \in I);$$

$$f(g,i) := \begin{cases} 1, & \text{falls } c(i,s) = \infty \\ 0, & \text{sonst} \end{cases} \quad (g \in F, i \in I, (g,i) \in R);$$

$$f(g,g) := \sum_{\substack{i \in I \\ (g,i) \in R}} f(g,i) \quad (g \in F).$$

Als wesentliche Ursache für diese einfache Berechnungsmöglichkeit für f ist neben der Struktur des Digraphen vor allem die verwendete Kapazitätsfunktion zu nennen, die für Bögen der Form $(g,g), g \in F$, nur die Kapazität ∞ und für Bögen der Form $(i,s), i \in I$, nur die Kapazitäten 0 bzw. ∞ zuläßt.

Betrachten wir nochmals Beispiel 2, so erhalten wir nun leicht den in Abb. 2 angegebenen Maximalstrom f ; dieser hat offensichtlich die Stromstärke $\sum_{k=1}^4 f(g, g_k) = \sum_{i=1}^5 f(i, s) = 6$. Als mögliche funktionelle Gruppen ergeben sich g_2 und g_4 . Drei grundlegende Vorteile dieser graphentheoretischen Behandlung von Zuordnungsproblemen sind:

- (1) die anschauliche und übersichtliche Darstellung;
- (2) der hohe Allgemeinheitsgrad und die Robustheit dieses Zuganges;
- (3) die Flexibilität und leichte Erweiterungsfähigkeit des Modells.

Zu (2) ist zu bemerken, daß eventuell vorhandene Abhängigkeiten etwa der funktionellen Gruppen untereinander oder Überlappungen der charakteristischen Intervalle und ähnliches weder bei der Konstruktion des Transportnetzes noch bei der Abarbeitung des Algorithmus irgendwelche Schwierigkeiten verursachen oder gesondert berücksichtigt werden müssen. Versucht man hingegen, die erläuterte Spektrinterpretation mit Hilfe der Begriffsanalyse in einem Kontext (vgl. Wille/Gauter [WG], Kerber/Lex [KL]) zu erfassen, wobei die funktionellen Gruppen den Gegenständen und die charakteristischen Intervalle den Merkmalen entsprechen, so rufen etwa Überlappungen der Intervalle nur schwer zu beherrschende Probleme hervor. Zur Illustration von (3) überlegt man sich zunächst, wie einfach das Transportnetz um neue funktionelle Gruppen oder charakteristische Intervalle ergänzt werden kann. Erheblich interessanter ist jedoch die im nächsten Abschnitt behandelte Erweiterungsmöglichkeit.

5 Eine Erweiterung des einfachen Interpretationssystems

Da bei der Interpretation von Infrarotspektren neben der Lage der Banden (Wellenzahl) auch deren Intensität und Halbwertsbreite von grundlegender Bedeutung sind, ist die Einbeziehung dieser beiden Größen in die Betrachtungen notwendig. (In der Praxis ergibt sich hierbei eine Vielzahl von Problemen; wir weisen hier nur auf die unterschiedliche Wichtung der drei Merkmalsgruppen Wellenzahl, Intensität und Halbwertsbreite hin. Einige Ansatzpunkte zur Einbeziehung solcher und ähnlicher Überlegungen findet man in Abschnitt 6.)

Jedem geordneten Paar $(g, i) \in R$ bestehend aus einer funktionellen Gruppe $g \in F$ und einem für g charakteristischen Intervall $i \in I$ (vgl. (E) in 4.) kann man ein Intensitätsintervall $[x_{g,i}, x'_{g,i}]$ und ein Halbwertsbreitenintervall $[y_{g,i}, y'_{g,i}]$ zuordnen, die man aus einschlägigen Tabellen entnehmen kann. Eine Bande im charakteristischen Intervall i werden wir im folgenden genau dann als *passend* zu g bezeichnen, wenn sowohl ihre Intensität in $[x_{g,i}, x'_{g,i}]$ als auch ihre Halbwertsbreite in $[y_{g,i}, y'_{g,i}]$ liegen. Da in vielen Tabellenwerken die Intensitäten relativ unscharf rein verbal beschrieben werden (vgl. Pretsch/Clerc/Seibl/Simon [PCSS]), genügen für unsere Zwecke die Intensitätswerte 1, 2, 3, 4 und 5; für die Halbwertsbreiten (in cm^{-1}) werden wir jedoch (wie für die Lage der Banden) Intervalle angeben. Unsere Grundannahme ist ähnlich wie in 4. – nun, daß eine funktionelle Gruppe g nur dann in der untersuchten Verbindung vorhanden sein kann (wir nennen g dann wieder eine *mögliche* funktionelle Gruppe), wenn in jedem charakteristischen Intervall von g eine zu g passende Bande liegt.

Den Digraphen des zugehörigen Transportnetzes kann man aus dem in 4. benutzten Transportnetz erhalten, indem man jeden "Intervallknoten" $i \in I$ durch einen Digraphen der Form

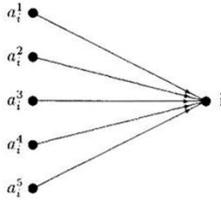


Abb. 3
und jeden Bogen $(g, i) \in R$ ($g \in F$, $i \in I$) durch

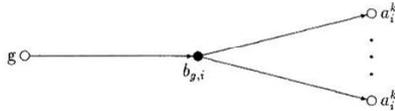


Abb. 4
ersetzt.
Dabei entspricht für $k \in \{1, 2, \dots, 5\}$ der Knoten a_i^k der Intensität k im Intervall i und $b_{g,i}$ ist das zur funktionellen Gruppe g und dem zugehörigen charakteristischen Intervall i gehörige Intervall für die Halbwertsbreite. $b_{g,i}$ und a_i^k werden genau dann durch einen Bogen verbunden, wenn für eine zu g passende Bande im Intervall i mit einer Halbwertsbreite in $b_{g,i}$ die Intensität k zulässig ist.
Genauer seien

$$F := \{g_1, g_2, \dots, g_m\}$$

eine Menge funktioneller Gruppen;

$$I := \{i_1, i_2, \dots, i_n\}$$

eine Menge von Intervallen für die Lage der Banden;

$$A := \{a_i^k \mid i \in I, k \in \{1, 2, \dots, 5\}\}$$

die Familie der zu den Intervallen gehörigen Intensitäten;

$$B := \{b_{g,i} \mid g \in F, i \in I, i \text{ charakteristisches Intervall von } g\}$$

die Familie der zu den funktionellen Gruppen und ihren charakteristischen Intervallen gehörigen Halbwertsbreitenintervalle.

Weiter seien

$$R_1 \subseteq F \times B$$

mit

$$R_1 := \{(g, b_{g,i}) \mid g \in F, i \in I, i \text{ charakteristisches Intervall von } g\}$$

und

$$R_2 \subseteq B \times A$$

mit

$$R_2 := \{(b_{g,i}, a_i^k) \mid g \in F, i \in I, (g, b_{g,i}) \in R_1, k \in \{1, 2, \dots, 5\}, \text{ für eine}$$

zu g passende Bande in i ist die Intensität k zulässig\}.

Damit können wir den gesuchten Digraphen $G = (V, E)$ angeben:

$$\begin{aligned} V &:= F \cup A \cup B \cup I \cup \{q, s\}; \\ E &:= \{(q, g) \mid g \in F\} \cup \\ &\quad R_1 \cup R_2 \cup \\ &\quad \{(a_i^k, i) \mid i \in I, k \in \{1, 2, \dots, 5\}\} \cup \\ &\quad \{(i, s) \mid i \in I\}. \end{aligned}$$

Mit der folgenden Kapazitätsfunktion c erhalten wir aus $G = (V, E)$ ein Transportnetz $N = (G, c, q, s)$:

$$\begin{aligned} c(q, g) &:= \infty && (g \in F); \\ c(g, b) &:= 1 && ((g, b) \in R_1); \\ c(b_{g,i}, a_i^k) &:= \begin{cases} 1, & \text{falls in } i \text{ eine Bande der Intensität } k \text{ mit einer Halbwertsbreite in } b_{g,i} \text{ vorliegt} \\ 0, & \text{sonst} \end{cases} && \begin{aligned} &(g \in F, i \in I, \\ &k \in \{1, 2, \dots, 5\}, \\ &(b_{g,i}, a_i^k) \in R_2); \end{aligned} \\ c(a_i^k, i) &:= \begin{cases} \infty, & \text{falls in } i \text{ eine Bande der Intensität } k \text{ vorliegt} \\ 0, & \text{sonst} \end{cases} && (i \in I, k \in \{1, 2, \dots, 5\}); \\ c(i, s) &:= \begin{cases} \infty, & \text{falls } i \text{ eine Bande enthält} \\ 0, & \text{sonst} \end{cases} && (i \in I). \end{aligned}$$

(Vgl. Beispiel 3, Abb. 5.)

Man überlegt sich, daß für die Bögen der Form $(b_{g,i}, a_i^k)$ für die Kapazitätsfunktion c anstelle des Wertes 1 jede beliebige positive natürliche Zahl oder ∞ verwendet werden kann, ohne das Endergebnis zu beeinflussen. Auch lassen sich zur Reduzierung der Größe des Transportnetzes alle Knoten a_i^k ($i \in I, k \in \{1, 2, \dots, 5\}$), die keinen

Vorgänger besitzen, zusammen mit dem jeweils zugehörigen Bogen (a_i^k, i) entfernen. Des weiteren kann man zwei Knoten $b_{g,i}$ und $b_{g',i}$ identifizieren, falls die beiden Halbwertsbreitenintervalle übereinstimmen und $b_{g,i}$ dieselben Nachfolger a_i^k, \dots, a_i^l wie $b_{g',i}$ hat ($g, g' \in F; i \in I; k, l \in \{1, 2, \dots, 5\}; k \leq l$). Dazu sind dann auch die Bogenmenge und die Kapazitätsfunktion entsprechend zu ändern.

Betrachten wir nun jedoch das von uns oben konstruierte Transportnetz $N = (G = (V, E), c, q, s)$ ohne diese Modifikationen, so erkennen wir, daß für dieses wieder eine vereinfachte Variante des Algorithmus von Ford und Fulkerson benutzbar ist, die allerdings im Vergleich zu dem in 4. benutzten Algorithmus etwas komplizierter ist (die Existenz eines endlichen Schnittes von N ist leicht zu sehen, so daß auch ein Maximalstrom existiert):

Sei $i \in I$ ein charakteristisches Intervall von $g \in F$, weiter seien $k, l \in \{1, 2, \dots, 5\}$ mit $k \leq l$ und $a_i^k, a_i^{k+1}, \dots, a_i^l$ die sämtlichen Nachfolger von $b_{g,i}$. Existiert kein $t \in \{k, k+1, \dots, l\}$ mit $c(b_{g,i}, a_i^t) = 1$, so setzen wir

$$f(b_{g,i}, a_i^k) := \dots := f(b_{g,i}, a_i^l) := 0,$$

anderenfalls wählen wir ein solches t und setzen

$$f(b_{g,i}, a_i^t) := 1 \text{ sowie}$$

$$f(b_{g,i}, a_i^{t'}) := 0 \text{ für alle } t' \in \{k, k+1, \dots, l\} - \{t\}.$$

Auf diese Weise läßt sich f für die Bögen $(b, a) \in R_2$ definieren und wir können f für alle restlichen Bögen so angeben, daß ein mit den Kapazitäten verträglicher Strom entsteht:

$$f(g, b_{g,i}) := \begin{cases} 1, & \text{falls ein } k \in \{1, 2, \dots, 5\} \\ & \text{existiert mit } f(b_{g,i}, a_i^k) = 1 \\ 0, & \text{sonst} \end{cases} \quad \begin{matrix} (g \in F, i \in I, \\ (g, b_{g,i}) \in R_1); \end{matrix}$$

$$f(g, g) := \sum_{(g, b_{g,i}) \in R_1} f(g, b_{g,i}) \quad (g \in F);$$

$$f(a_i^k, i) := \sum_{\substack{g \in F \\ (b_{g,i}, a_i^k) \in R_2}} f(b_{g,i}, a_i^k) \quad (i \in I, k \in \{1, 2, \dots, 5\});$$

$$f(i, s) := \sum_{k=1}^5 f(a_i^k, i) \quad (i \in I).$$

Man erkennt, daß der so konstruierte Strom maximal ist. Im ersten Teil des Algorithmus (Konstruktion von f für alle $(b, a) \in R_2$) findet man die Ursache dafür, daß in einem Transportnetz der hier betrachteten Art im allgemeinen mehrere Maximalströme existieren. Die Menge der möglichen funktionellen Gruppen ist jedoch unabhängig von dem jeweils berechneten Maximalstrom. Wenden wir den beschriebenen Algorithmus auf Beispiel 3 an, so erhalten wir als maximale Stromstärke 6 und als mögliche funktionelle Gruppen g_1 und g_2 .

Beispiel 3:

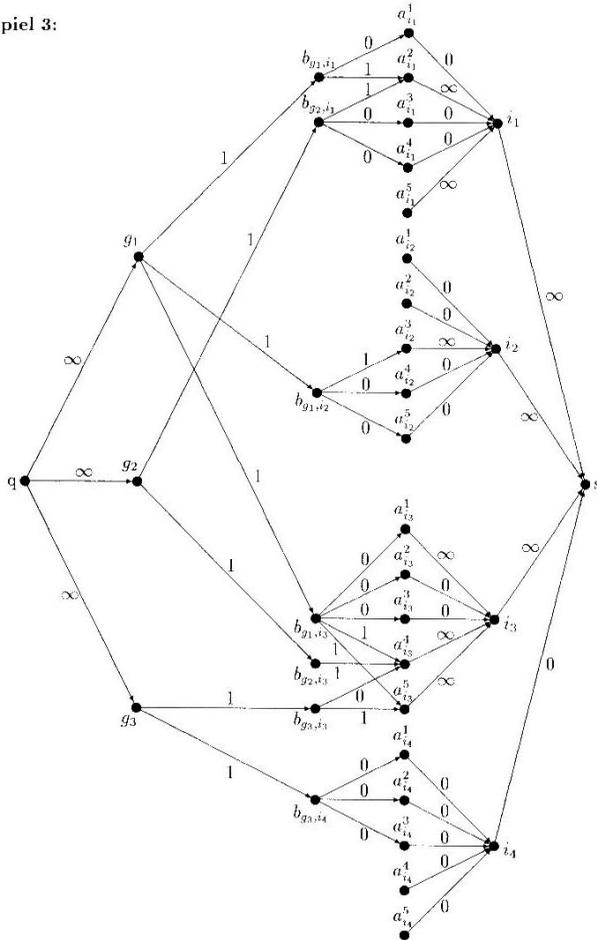


Abb. 5

6 Weitere Modifikationsmöglichkeiten

Neben der Aufnahme weiterer Spektralmerkmale in das Transportnetz (wie etwa im vorigen Abschnitt für Intensität und Halbwertsbreite demonstriert) bietet sich die Kapazitätsfunktion als wirksames Werkzeug an. Bisher wurden als Bogenkapazitäten ausschließlich 0,1 und ∞ benutzt. Der Algorithmus von Ford und Fulkerson gestattet aber beliebige rationale, nichtnegative Kapazitäten und ∞ . Damit hat man die Möglichkeit, die Bögen mit "Gewichten" aus $\mathbf{Q}^+ \cup \{0, \infty\}$ zu versehen und so das Verhalten des Transportnetzes auf die unterschiedlichste Art und Weise zu verändern. Dabei ist zu beachten, daß im Falle von Modifikationen der in 4. bzw. 5. untersuchten Transportnetze die dort beschriebenen vereinfachten Algorithmen im allgemeinen nicht mehr anwendbar bleiben. Der Algorithmus von Ford und Fulkerson in der unter 3. beschriebenen allgemeinen Form liefert jedoch in jedem Spezialfall einen Maximalstrom.

Betrachten wir noch einige weitere Modifikationsmöglichkeiten:

In Tabellenwerken zur Infrarot-Spektroskopie (etwa [PCSS]) findet man oft Hinweise auf Intervalle i , in denen unter gewissen Bedingungen (man liest auch: meist, oft, gelegentlich, ...) bei Vorhandensein einer bestimmten funktionellen Gruppe g eine zu g passende Bande liegt. Ein solches Intervall i ist nicht in dem unter 4. eingeführten Sinne charakteristisch für g . Wir können es aber durchaus in unser Transportnetz einbeziehen; der Einfachheit halber betrachten wir hier nur das in 4. untersuchte Transportnetz $N = (G = (V, E), c, q, s)$. Seien g und i wie eben erläutert, darüber hinaus besitze g genau m charakteristische Intervalle. Wir konstruieren aus N ein neues Transportnetz $N' = (G' = (V', E'), c', q, s)$:

$$\begin{aligned} V' &:= V \cup \{i\}; \\ E' &:= E \cup \{(g, i), (i, s)\}; \\ c'(g, i) &:= \alpha \quad \text{mit } 0 < \alpha < 1, \text{ fest;} \\ c'(i, s) &:= \begin{cases} \infty, & \text{falls das Spektrum in } i \text{ eine} \\ & \text{Bande besitzt} \\ 0, & \text{sonst;} \end{cases} \end{aligned}$$

für alle Bögen aus E erhält c' dieselben Werte wie c . Ist nun f ein Maximalstrom auf N' , so bedeutet dies, daß g im Falle $f(g, g) = m$ (d.h. i enthält keine Bande) mit einer gewissen Wahrscheinlichkeit $p_1 > 0$ als mögliche funktionelle Gruppe akzeptiert werden kann. Im Falle $f(g, g) = m + \alpha$ würde man g dann mit einer Wahrscheinlichkeit $p_2 > p_1$ als mögliche funktionelle Gruppe akzeptieren. Auf diese Weise ließe sich gegebenenfalls in der Menge der möglichen funktionellen Gruppen eine gewisse Rangfolge erhalten, die bei der Weiterverarbeitung dieser Menge – etwa mittels eines Strukturgenerators wie MOLGEN, GFNOA, ... – nützlich sein kann.

Eine weitere Möglichkeit wäre die Festlegung geeigneter endlicher Kapazitätswerte für die Bögen (g, g) , $g \in F$, welche etwa unterschiedliche Wahrscheinlichkeiten für das Vorliegen der funktionellen Gruppen $g \in F$ widerspiegeln, die aus Informationen

über Ausgangsstoffe, Reaktionsbedingungen o.ä. stammen. Damit sich diese Werte dann bei der Berechnung des Maximalstromes wunschgemäß auswirken, müssen eventuell auch die Kapazitäten der Bögen (g, i) bzw. (i, s) geeignet modifiziert werden ($g \in F, i \in I$).

Da solche Manipulationen stark von der angestrebten Zielstellung abhängen, wollen wir darauf nicht detaillierter eingehen.

Als letzte Bemerkung zur Modifikation des Algorithmus von Ford und Fulkerson soll an dieser Stelle jedoch noch auf die Verwendung einer Kostenfunktion für die Bögen des Transportnetzes hingewiesen werden. Zu diesem Zwecke sei $N = (G = (V, E), c, q, s)$ ein beliebiges Transportnetz und $z: E \rightarrow \mathbf{R}^+ \cup \{0\}$ eine Funktion, die jedem Bogen $e \in E$ eine nichtnegative reelle Zahl $z(e)$ zuordnet; $z(e)$ wird dabei interpretiert als die Kosten, die ein Strom der Größe 1 entlang des Bogens e verursacht. Eine interessante Frage ist nun die nach einem Maximalstrom f auf N , der bezüglich der Kostenfunktion z minimale Gesamtkosten

$$z_{Ges}(f) = \sum_{e \in E} f(e) \cdot z(e)$$

aufweist.

Die Anwendung auf Transportprobleme ist offensichtlich; ob die Benutzung dieses recht allgemeinen und leistungsfähigen Ansatzes für die Spektrinterpretation notwendig und sinnvoll ist, bleibt der Einschätzung des interessierten Lesers überlassen. Zum Maximalstromproblem mit Kostenfunktion verweisen wir auf [C].

7 Schlußbemerkung

Der vorgestellte Algorithmus von Ford und Fulkerson ist ein äußerst flexibles, erweiterungsfähiges und anschauliches Werkzeug zur Lösung gewisser Strom-, Transport- oder auch Zuordnungsprobleme (siehe etwa [C], [W], [WN], [Wi]). Die Interpretation von Infrarot-, ^{13}C -NMR-, ^1H -NMR- oder anderen Spektren ist nur ein Beispiel für seine Anwendung in der Chemie; eine Vielzahl – auch sehr komplexer – Zuordnungsprobleme in den verschiedensten Wissenschaftszweigen kann damit bearbeitet werden.

LITERATUR

- [C] Carré, B.: Graphs and networks;
Clarendon Press Oxford 1979
- [H] Harary, F.: Graphentheorie;
Oldenbourg Verlag München Wien 1974
- [KL] Kerber, A.; Lex, W.: Begriffsanalyse;
Maschinenskript 1993
- [PCSS] Pretsch, E.; Clerc, T.; Seibl, J.; Simon, W.:
Tabellen zur Strukturaufklärung organischer Verbindungen
Springer Verlag Berlin Heidelberg New York
London Paris Tokyo 1986
- [S] Sachs, H.: Einführung in die Theorie der endlichen Graphen, Teil 1;
BSB B.G. Teubner Verlagsgesellschaft Leipzig 1970
- [W] Walther, H.: Anwendungen der Graphentheorie;
Vieweg & Sohn Braunschweig Wiesbaden 1978
- [WN] Walther, H.; Nägler, G.: Graphen-Algorithmen-Programme;
Springer-Verlag Wien New York 1987
- [Wi] Wilf, H.S.: Algorithms and Complexity;
Prentice Hall New Jersey 1986
- [WG] Wille, R.; Ganter, B.: Mathematische Theorie der Formalen
Begriffsanalyse;
Maschinenskript 1990