Aufgabe 1 – *Arbeitsgruppen*

Zu einem Kongress werden Mitarbeiter von m Firmen gesandt. Dabei sendet die i-te Firma c_i viele Mitarbeiter. Während des Kongresses sollen die anwesenden Personen in bis zu r verschiedene Arbeitsgruppen mit jeweils höchstens 5 Mitgliedern aufgeteilt werden, wobei keine zwei Mitarbeiter derselben Firma in derselben Arbeitsgruppe sein sollen.

Wir sollen eine solche Aufteilung mit Hilfe von Fluss-Algorithmen finden.

(a) Definieren Sie dazu ein geeignetes Netzwerk $N = (V, A, c, s, t)$ und zeigen Sie, dass es genau dann eine mögliche Aufteilung wie oben gibt, wenn maxflow(N) eine gewisse (welche?) Eigenschaft hat.

(b) Angenommen, wir erhalten als Eingabe die Zahlen $m, r, c_1, \ldots, c_m$ sowie die Listen $L_1, \ldots, L_m$ mit den Namen aller Teilnehmer des Kongresses. Wir können davon ausgehen, dass keine zwei Teilnehmer den gleichen Namen haben. Geben Sie einen Algorithmus an, der dies als Eingabe verwendet, und als Ausgabe entweder r Listen, $G_1, ..., G_r$ gibt, wobei G_i die Namen aller Teilnehmer der i:ten Gruppe in einer gültigen Zuordnung enthält, oder antwortet „Zuordnung nicht möglich“.

(a) Die Aufgabe besteht darin, Mitarbeiter von Firmen auf Gruppen aufzuteilen mit den folgenden Kriterien:

   (i) Jede Firma i hat genau c_i Mitarbeiter auf die Gruppen verteilt.

   (ii) Jede Gruppe hat höchstens 5 Mitarbeiter.

   (iii) Keine Firma verteilt mehr als einen Mitarbeiter auf eine Gruppe.

   (iv) Es gibt höchstens r viele Gruppen.

Wir modellieren das Problem als Flussproblem auf einem Netzwerk N wie folgt: N hat eine Quelle s und eine Senke t. Darüber hinaus fügen wir für jede Firma i ($1 \leq i \leq m$) einen Knoten f_i und für jede der r (potentiell leeren) Arbeitsgruppen j einen Knoten a_j ein. Die Knoten f_i sind durch einer Kante (s, f_i) mit der Quelle verbunden, die Kapazität c_i hat (intuitiv wird der Fluss entlang dieser Kanten bestimmen, wie viele Mitarbeiter der Firma an Arbeitsgruppen teilnehmen können). Des Weiteren verbinden wir jede Firma f_i mit jeder Arbeitsgruppe a_j durch eine Kante (f_i, a_j) der Kapazität 1 (diese Kante stellt sicher, dass aus jeder Firma nur ein Mitarbeiter pro Arbeitsgruppe vertreten ist). Schlussendlich verbinden wir jede Arbeitsgruppe durch eine Kante (a_j, t) mit Kapazität 5 mit der Senke (diese Kante stellt sicher, dass jede Arbeitsgruppe höchstens 5 Teilnehmer hat).

Formal bedeutet dies, dass unser Netzwerk $N = (V, A, c, s, t)$ wie folgt definiert ist:

$V := \{s, f_1, \ldots, f_m, a_1, \ldots, a_r, t\}$
A $:= \{(s, f_i) \mid i \in [m]\} \cup \{(f_i, a_k) \mid i \in [m], k \in [r]\} \cup \{(a_k, t) \mid k \in [r]\}$

$$c(u, v) := \begin{cases} c_i, & \text{falls } u = s \text{ und } v = f_i, i \in [m] \\ 1, & \text{falls } u = f_i \text{ und } v = a_k, i \in [m], k \in [r] \\ 5, & \text{falls } u = a_k \text{ und } v = t, k \in [r]. \end{cases}$$

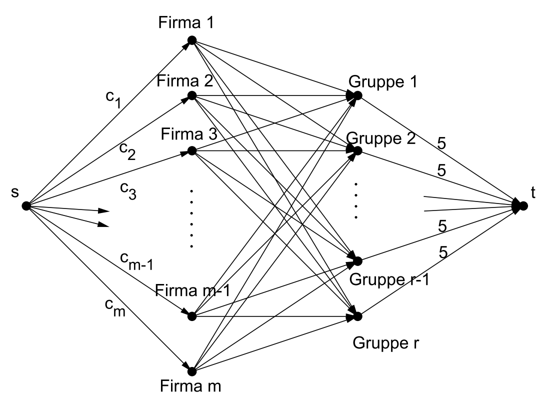

Abbildung 1: Die unbeschrifteten Kanten haben Kapazität 1.

Wir wollen im Folgenden zeigen, dass es eine zulässige Aufteilung gibt genau dann, wenn es in N einen maximalen Fluss mit Wert $\text{maxflow}(N) \geq c := \sum_{i=1}^{m} c_i$ gibt. Dazu müssen wir einerseits zeigen, wie wir aus einer zulässigen Aufteilung einen Fluss der Kapazität c konstruieren können und wie wir andererseits aus einem Fluss der Kapazität c eine solche Aufteilung konstruieren konnen.

Zunächst sehen wir, dass $\text{maxflow}(N) \leq c$, da die Summe aller von s ausgehenden Kapazitäten genau c ist. Daher gilt $\text{maxflow}(N) \geq c \iff \text{maxflow}(N) = c$.

Aufteilung $\implies$ Fluss: Nehmen wir zunächst an, dass es eine zulässige Aufteilung gibt (und fixieren wir eine solche Aufteilung). Dann konstruieren wir einen Fluss, indem wir entlang jeder Kante (s, f_i) den Fluss c_i schicken und von jedem Knoten f_i Fluss 1 entlang (f_i, a_j) genau dann, wenn ein Mitarbeiter von f_i an der Gruppe a_j teilnimmt - und Fluss 0 sonst. Schliesslich schicken wir entlang (a_j, t) den Fluss, der der Anzahl der Teilnehmer in Gruppe j entspricht. Man sieht leicht, dass dieser Fluss zulässig ist. Die Kapazität enlang der Kanten (f_i, a_j) wird nicht überschritten, da in jeder Gruppe höchstens ein Mitarbeiter pro Firma vertreten ist; die Kapazität entlang der Kanten (a_j, t) wird respektiert, da es pro Gruppe höchstens 5 Teilnehmer gibt. Aufgrund der Flusserhaltung $\text{inflow}(a_j) = \text{outflow}(a_j)$ gilt an jedem Firmen- und Arbeitsgruppenknoten per Konstruktion ($\text{inflow}(f_i) = c_i = \text{outflow}(f_i)$, da c_i Mitarbeiter von f_i auf die Gruppen verteilt werden. Schliesslich wollen wir noch zeigen, dass dieser Fluss maximal ist. Dafür genügt es, zu bemerken, dass $\text{val}(f) = \text{outflow}(s) = \sum_{i=1}^{m} c_i = c$.

Fluss $\implies$ Aufteilung: Sei nun f ein maximaler Fluss der Kapazität c. Bevor wir die Aufteilung konstruieren, machen wir zwei Beobachtungen: Aus Satz 3.12 folgt, dass wir annehmen können, dass dieser Fluss ganzzahlig ist, da alle Kapazitäten ganzzahlig sind. Insbesondere ist der Fluss entlang jeder Kante (f_i, a_j) entweder 0 oder 1. Des Weiteren wissen wir, dass entlang jeder Kante (s, f_i) der Fluss c_i fliesst, da jede Firma c_i Mitarbeiter zur Konferenz schickt. Wir weisen nun einen Mitarbeiter der Firma f_i der Gruppe a_j zu, genau dann wenn $f((f_i, a_i)) = 1$. Da $\text{inflow}(f_i) = c_i$ gilt auch $\text{outflow}(f_i) = c_i$ und deshalb werden c_i Mitarbeiter der Firma auf die Arbeitsgruppen aufgeteilt. Man sieht ebenfalls leicht, dass pro Arbeitsgruppe nur ein Mitarbeiter aus jeder Firma vertreten ist und dass in jeder Arbeitsgruppe höchstens 5 Teilnehmer sind (da $\text{inflow}(a_j) = \text{outflow}(a_j) \leq 5$). Damit haben wir gezeigt, dass aus der Existenz eines maximalen Flusses mit Wert c folgt, dass es eine zulässige Aufteilung gibt und den Beweis abgeschlossen.

(b) Wir konstruieren unser Netzwerk wie in a). Die Strategie ist nun, einen maximalen ganzzahligen Fluss f zu finden. Da alle Kapazitäten ganzzahlig sind, ist dies z.B. mittels Ford-Fulkerson möglich. Dann prüfen wir, ob $\text{val}(f) = c$.
Falls nein, dann geben wir "Zuordnung nicht möglich" aus.
Falls ja, dann nutzen wir die Richtung Fluss $\implies$ Aufteilung aus a), um eine gültige Aufteilung zu finden. Durch den aus a) erhaltenen Fluss f können wir für jede Arbeitsgruppe k alle Firmen finden als $F_k := \{i \in [m] | f(f_i, a_k) = 1\}$, die Mitarbeiter in die jeweilige Gruppe schicken. Umgekehrt können wir für jede Firma i die zugehörigen Arbeitsgruppen finden als $Z_i := \{k \in [r] | f(f_i, a_k) = 1\}$ mit $|Z_i| = c_i$. Sei nun $\phi_i : Z_i \to [c_i]$ eine beliebige bijektive Abbildung für jedes $i \in [m]$, z.B. die kanonisch kleinste solche Abbildung. ϕ_i bildet eine Arbeitsgruppe auf einen Mitarbeiter der Firma i ab. Nun können wir die Liste der Teilnehmer der Gruppe definieren als $G_k := (L_{i,\phi_i(k)})_{i \in F_k}$.

Die Korrektheit des Algorithmus folgt direkt aus Aufgabe a). Die Laufzeit des Algorithmus ist dominiert duch die Subroutine den maximalen Fluss zu finden, für Ford-Fulkerson also $\mathcal{O}(|A| \cdot |V| \cdot U)$ wobei $|A| \in \Theta(m \cdot r)$ die Anzahl der Kanten des konstruierten Netzwerks ist, $|V| \in \Theta(m+r)$ die Anzahl der Knoten und $U := \max(c_1, ..., c_m, 5)$ die maximale Kapazitat der Kanten. Die restliche "Buchhaltung" des Algorithmus (das Erstellen der Listen G_k) benötigt Zeit $\mathcal{O}(mr + c)$.

Insgesamt haben wir also Laufzeit $\mathcal{O}(|A| \cdot |V| \cdot U + mr + c) = \mathcal{O}(mr(m+r)U + c) = \mathcal{O}(mr(m+r)U)$, da $c \leq m \cdot U$.