8. Übungsblatt

Ausgabe: 27. Januar 2009
Abgabe: 10. Februar 2009

Die Bearbeitung in Zweiergruppen ist ausdrücklich erwünscht.

Correlation Clustering

Eine Alternative zum dichte-basierten Clusterungsproblem ist durch das *Correlation Clustering* Paradigma gegeben. Dazu sei $G = (V, E, \omega)$ ein gewichteter, ungerichteter Graph mit $\omega\colon E \to \mathbb{R}$. Es werden nun Clusterungen gesucht, die folgende Optimierungsproblem lösen:

$$\textsc{MaxAgr} \quad : \quad \max_{\mathcal{C}'} \left(\sum_{e \in E(\mathcal{C}')} \max(0, \omega(e)) - \sum_{e \in \overline{E(\mathcal{C}')}} \min(0, \omega(e)) \right)$$

$$\textsc{MinDis} \quad : \quad \min_{\mathcal{C}'} \left(- \sum_{e \in E(\mathcal{C}')} \min(0, \omega(e)) + \sum_{e \in \overline{E(\mathcal{C}')}} \max(0, \omega(e)) \right)$$

Problem 1 — 2+2+2 Punkte

(a) Beschreiben Sie anschaulich, was die beiden Problem MaxAgr und MinDis optimieren. Schreiben Sie die Zielfunktionen entsprechend um, ohne dabei max / min innerhalb zu verwenden.

(b) Zeigen Sie, dass eine Clusterung $\mathcal{C}$ genau dann die Zielfunktion von MaxAgr maximiert, wenn $\mathcal{C}$ die Zielfunktion von MinDis minimiert.

(c) Geben Sie einen einfachen 2-Approximationsalgorithmus für MaxAgr an. Ist dieser Algorithmus auch eine Approximationsalgorithmus für MinDis?

Problem 2 — 3 Punkte

Zeigen Sie, dass $\sum_{e \in E, \omega(e) > 0} \omega(e) - \sum_{e \in E, \omega(e) < 0}$ eine scharfe obere Schranke für den maximalen Wert von MaxAgr und 0 eine scharfe untere Schranke für den minimalen Wert von MinDis ist. Zeigen Sie, dass man in polynomieller Zeit entscheiden kann, ob es eine Clusterung $\mathcal{C}$ einer Instanz $I = \langle G = (V, E, \omega) \rangle$ gibt, die diesen Wert annimmt.

Problem 3 — 2+6+2 Punkte

Zeigen Sie, dass das Entscheidungsproblem zu MinDis $\mathcal{NP}$-vollständig für ungewichteten Graphen $G = (V, E, \omega)$ mit $\omega\colon E \to \{-1, +1\}$ ist, d.h. zu entscheiden, ob es eine Clusterung gibt, so dass der Wert der Zielfunktion höchstens K ist.

Bitte wenden!

Reduzieren Sie dabei von dem Problem Partition into Triangles (GT11): Gegeben ein Graph mit $n = 3k$ Knoten. Gibt es eine Partition der Knoten in k Menge $V_1, \dots V_k$, die jeweils drei Elemente enthalten, so dass $G[V_i]$ ein Dreieck ist. Gehen Sie dabei wie folgt vor:

(a) Sei $G = (V, E)$ ein beliebiger Graph, dann ist $H_G := (V, \binom{V}{2}, \omega)$ der zugehörige vollständige, gewichtete und ungerichtete Graph definiert durch:

$$\omega(\{u, v\}) := \begin{cases} +1 & \text{, falls } \{u, v\} \in E \\ -1 & \text{, sonst} \end{cases} .$$

Eine Clusterung $\mathcal{C}$ von H_G, die bezüglich MinDis optimal ist und bei der jeder Cluster $C \in \mathcal{C}$ höchstens drei Knoten hat, kann dazu genutzt werden zu entscheiden, ob G eine JA/NEIN-Instanz von Partition into Triangles ist.

(b) Sei $G = (V, E)$ ein beliebiger Graph, dann ist $H := H_{G^+}$, wobei der Graph $G^+ := (V', E')$ wie folgt definiert ist:

$$\begin{aligned} V_{\{u,v,w\}} &:= \left\{ c^{(i)}_{\{u,v,w\}} \mid 1 \le i \le n^6 \right\} \qquad \text{für alle } \{u, v, w\} \in \binom{V}{3}, \\ V' &:= V \uplus \biguplus_{\{u,v,w\} \in \binom{V}{3}} V_{\{u,v,w\}} \\ E_{\{u,v,w\}} &:= \left\{ \{x, y\} \mid x \in \{u, v, w\} \cup V_{\{u,v,w\}}, y \in V_{\{u,v,w\}}, x \neq y \right\} \\ &\qquad \text{für alle } \{u, v, w\} \in \binom{V}{3}, \\ E' &:= E \uplus \biguplus_{\{u,v,w\} \in \binom{V}{3}} E_{\{u,v,w\}} \end{aligned}$$

Zeigen Sie, dass aus einer optimalen Clusterung $\mathcal{C}$ von H bezüglich MinDis eine optimale Clusterung $\mathcal{C}'$ von H_G bauen läßt, in der jeder Cluster höchstens drei Knoten enthält.

(i) Zeigen Sie, $\mathcal{C}$ enthält $\binom{n}{3}$ Cluster und der Wert der Zielfunktion ist höchstens

$$n^7 \left(\binom{n}{2} - 1 \right) + \binom{n}{2} .$$

(ii) Jeder Cluster enthält genau eine Clique mit den Knoten $V_{\{u,v,w\}}$ für ein geeignetes 3-Tupel $\{u, v, w\}$.

(iii) Jeder Knoten $u \in V$ ist in einem Cluster, der eine Clique mit den Knoten $V_{\{u,v,w\}}$ für geeignete Knoten $v, w \in V$ enthält.

(c) Vervollständigen Sie die Argumentation für den $\mathcal{NP}$-Vollständigkeitsbeweis.