

# The complete optimal stars-clustering-tree problem

Tatiana Ayngorn

25. Januar 2010

## Inhaltsverzeichnis

<b>1</b>	<b>Einleitung</b>	<b>1</b>
1.1	Das <i>COS</i> und das <i>OS</i> . . . . .	1
<b>2</b>	<b>Anwendungen des <i>COS</i> in der Praxis</b>	<b>3</b>
2.1	Konstruktion von Kommunikationsnetzwerken . . . . .	3
2.2	Key-Management für sichere Gruppenkommunikation . . . . .	4
<b>3</b>	<b>Eigenschaften des <i>COS</i> und eine Hilfskonstruktion</b>	<b>4</b>
3.1	Hilfskonstruktion . . . . .	5
<b>4</b>	<b>Strukturtheorem für das <i>COS</i></b>	<b>6</b>
<b>5</b>	<b>Ein polynomieller Algorithmus für das <i>COS</i></b>	<b>9</b>
5.1	Kernidee des Algorithmus . . . . .	9
5.2	Algorithmus für das <i>COS</i> . . . . .	10
<b>6</b>	<b>Fazit</b>	<b>11</b>

## 1 Einleitung

Das Hauptproblem, das wir betrachten, ist das *complete optimal stars-clustering-tree problem*, abgekürzt *COS*. Es besteht darin, in einem vollständigen Graphen  $G = (V, E)$  mit gewichteten Kanten  $E$  und einer Knotenmenge  $V$  einen aufspannenden Baum  $T$  minimalen Gewichts zu finden, so dass jede Teilmenge der Knotenmenge einen vollständigen Stern in  $T$  induziert.

### 1.1 Das *COS* und das *OS*

Das *COS* ist ein Spezialfall eines allgemeineren Problems, des *optimal stars-clustering-tree problem*, abgekürzt *OS*. Die Voraussetzungen des *OS* sind: sei  $(V, \mathbf{S})$  ein Hypergraph, wobei  $V$  mit  $V = \bigcup \mathbf{S}$ ,  $|V| \geq 3$  die Knotenmenge

und  $\mathbf{S}$  eine einfache Familie von Teilmengen von  $V$  ist (d.h.,  $\mathbf{S}$  ist *einfach*, wenn für alle  $S_i, S_j, i \neq j$  gilt  $S_i \neq S_j$ ).

**Definition 1.1.** .

1. Ein **Stern** ist ein Baum, wo ein Knoten im Zentrum des Sterns ist und die restlichen Knoten Blätter sind, die mit dem Zentrum durch Kanten verbunden sind.
2. Sei  $G = (V, E)$  ein vollständiger Graph mit einer Gewichtsfunktion  $w : E \rightarrow \mathbb{R}$ . Ein aufspannender Baum  $T$  in  $G$  heißt ein **star-clustering-tree** bezüglich  $(V, \mathbf{S})$ , wenn jedes  $S_i \in \mathbf{S}$  einen Stern in  $T$  induziert. Das **Gewicht** von  $T = (V, E_T)$  ist  $w(t) = \sum \{w(e) : e \in E_T\}$ .
3. Jede Teilmenge  $S_i \in \mathbf{S}$ , die einen Stern in  $T$  induziert, wird **Cluster** genannt.
4. Sei  $T$  ein Wald mit der Knotenmenge  $V$ . Für jeden Knoten  $v \in V$  bezeichnen wir die Menge, die aus  $v$  und allen seinen Nachbarn in  $T$  besteht, mit  $N_T[v]$ .
5. Für jedes  $v \in V$  definieren wir einen **vollständigen Stern von  $T$ , zentriert in  $v$**  und induziert von  $N_T[v]$ . Wir sagen, dass eine Teilmenge  $S_i \in \mathbf{S}$  einen **vollständigen Stern in  $v$**  in  $T$  induziert, wenn  $N_T[v] = S_i$ .

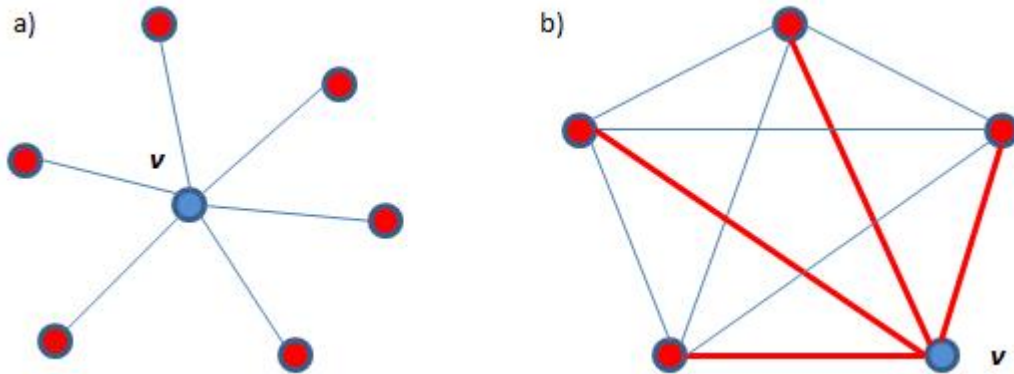


Abbildung 1: a) Ein Stern, der in  $v$  zentriert ist und von  $N_T[v]$  induziert wird; b) Ein vollständiger Stern in einem vollständigen Graphen  $K_5$ .

Das *OS*-Problem ist es, einen star-clustering-tree des minimalen Gewichts in  $G$  bezüglich  $(V, \mathbf{S})$  zu finden. Wir wissen aber nichts über die Komplexität des *OS*. Offensichtlich entspricht das *COS* dem *OS* bis auf eine zusätzliche Einschränkung, dass im *COS* jedes  $S_i \in \mathbf{S}$  einen vollständigen

Stern in dem Lösungsbaum  $T$  induziert. Es ist zu beachten, dass, wenn bei  $COS$   $S_i \neq S_j$  gilt,  $S_i$  und  $S_j$  kein gemeinsames Zentrum haben.

Für das  $COS$  führen wir die folgenden zwei Annahmen ein:

1. Die gegebene Menge  $\mathbf{S}$  von Teilmengen enthält keine einelementigen Mengen, d.h. es existiert kein  $S_i \in \mathbf{S}$  mit  $|S_i| = 1$ .
2. Der Durchschnittsgraph  $IG(\mathbf{S})$  ist zusammenhängend.

Wir brauchen die erste Annahme, da es sonst keinen Baum gäbe, der das Problem im trivialen Fall lösen kann. Die zweite Annahme ist notwendig, da wir einen polynomiellen Algorithmus entwickeln möchten, der das Problem löst. Andernfalls wäre die Komplexität des  $COS$  offen.

Die folgende Variante des  $COS$  ist NP-schwer. Das Problem heißt *complete minus-at-most-k OS* und entspricht dem  $OS$  bis auf eine zusätzliche Einschränkung, dass jede Teilmenge  $S_i \in \mathbf{S}$  einen vollständigen Stern, ausgenommen höchstens  $k$  Blätter, im Lösungsbaum  $T$  induziert.

**Definition 1.2. :**

*Eine Teilmenge  $S$  induziert einen vollständigen Stern, ausgenommen höchstens  $k$  Blätter, im Lösungsbaum  $T$ , wenn  $S$  einen Stern in  $T$  mit einem Zentrum in  $v \in S$  induziert, so dass der vollständige Stern in  $T$ , zentriert in  $v$ , höchstens  $k$  zusätzliche Blätter enthält, die nicht in  $S$  enthalten sind.*

Dieses Problem ist NP-schwer. Um es zu zeigen, kann man das Problem auf ein anderes polynomiell reduzieren, nämlich auf das Problem der Konstruktion eines minimal aufspannenden Baumes, wo der Grad jedes Knoten auf höchstens  $k$  beschränkt ist, welches bekanntermaßen NP-schwer<sup>1</sup> ist. Anders als bei dem  $COS$  haben wir nicht angenommen, dass der Durchschnittsgraph  $IG(\mathbf{S})$  zusammenhängend ist.

## 2 Anwendungen des $COS$ in der Parxis

Neben dem theoretischen Interesse für  $COS$  gibt es viele mögliche praktische Anwendungen dafür. Zwei Beispiele davon werden hier vorgestellt.

### 2.1 Konstruktion von Kommunikationsnetzwerken

*Gegeben:* ein vollständiger Graph, in dem jeder Knoten einen Kunden repräsentiert und jede Kante einen Anschluß  $e$ , der mit den Kosten (Gewicht)  $w(e)$  versehen werden kann. Außerdem gib es eine Menge von (nicht notwendig disjunkten) Teilmengen (Clusters) von Knoten, wo jede Teilmenge eine Gruppe von Kunden repräsentiert.

---

<sup>1</sup>Die NP-Schwierigkeit dieses Problems wurde von M.R. Garey und D.S. Johnson in [2] bewiesen.

*Problem:* finde einen kosten-(gewichts-)minimalen aufspannenden Baum im Kommunikationsnetzwerk, so dass jede Teilmenge von Knoten der gegebenen Menge einen vollständigen Stern im Lösungsbaum induziert.

Wird das Problem gelöst, so erhält das Netzwerk folgende Eigenschaften:

1. Die group-broadcast-Eigenschaft: es gibt einen einfachen Weg, eine Mitteilung an die Gruppe zu senden.
2. Die group-fault-tolerance-Eigenschaft: jede Gruppe von Kunden bleibt von Einflüssen der Netzwerkfehler, die außerhalb von der Gruppe vorkommen, ungestört.
3. Die group-privacy-Eigenschaft: Kunden, die keine Mitglieder einer Gruppe sind, können an der Kommunikation in der Gruppe nicht teilnehmen.

## 2.2 Key-Management für sichere Gruppenkommunikation

Wir betrachten verschiedene nicht disjunkte Gruppen von Users und wollen ein Kommunikationsnetzwerk konstruieren, welches jeder Gruppe eine sichere Gruppenkommunikation ermöglicht. Zunächst müssen wir für jede Gruppe ein Mitglied auswählen, welches die Schlüsselverwaltung (Key-Management) innerhalb von der Gruppe übernimmt. Dieser Verwalter muss ab und zu einen neuen Schlüssel an alle Gruppenmitglieder verteilen. Dann müssen wir einen aufspannenden Baum im Netzwerk konstruieren, so dass jede Gruppe einen vollständigen Stern induziert. Der Schlüsselverwalter soll aus Sicherheitsgründen im Zentrum des Sterns lokalisiert sein.

Da wir im Netzwerk so einen Baum von minimalem Gewicht konstruieren möchten, haben wir tatsächlich eine Instanz des *COS*-Problems.

## 3 Eigenschaften des *COS* und eine Hilfskonstruktion

In diesem Abschnitt werden einige Eigenschaften des *COS* präsentieren, sowie eine Hilfskonstruktion, die zum Strukturtheorem des *COS* führt, vorgestellt. Das Strukturtheorem spielt eine zentrale Rolle im polynomiellen Algorithmus für das *COS*, der später vorgeführt wird.

**Beobachtung 3.1.** *Angenommen, eine COS-Instanz (mit  $(V, \mathcal{S})$ ) hat eine zulässige Lösung  $T = (V, E_T)$ . Dann besitzt das COS folgende Eigenschaften:*

- a) *Für drei verschiedene Teilmengen  $S_i, S_j, S_k \in \mathcal{S}$  gilt:  $|S_i \cap S_j| \leq 2$  und  $|S_i \cap S_j \cap S_k| \leq 1$ .*

- b) Für zwei verschiedene Teilmengen  $S_i, S_j \in \mathbf{S}$  mit  $S_i \cap S_j = \{u, v\}$  gilt: die Kante  $(u, v)$  ist in jedem zulässigen Lösungsbaum enthalten, und die Menge der Zentren von  $S_i$  und  $S_j$  ist  $\{u, v\}$ .
- c) Für zwei verschiedene Teilmengen  $S_i, S_j \in \mathbf{S}$  mit  $S_i \cap S_j = \{u\}$  gibt es keinen zulässigen Lösungsbaum mit  $u$  als Zentrum von  $S_i$  oder  $S_j$ .
- d) Wenn  $(u, v) \in E_T$ , dann gibt es ein  $i$  mit  $\{u, v\} \subseteq S_i \in \mathbf{S}$
- e) Es existiert eine eindeutige Menge der Zentren von Teilmengen von  $\mathbf{S}$ , die zu  $T$  korrespondieren.

*Beweis.* **a)** und **e)** sind trivial. **d)** folgt aus der Tatsache, dass  $IG(\mathbf{S})$  zusammenhängend ist. **b)** offensichtlich gilt: wenn  $|S_i| = 2$  oder  $|S_j| = 2$ , dann folgt, dass  $(u, v)$  eine in jedem zulässigen Lösungsbaum enthaltende Kante ist. Wenn  $(u, v) \in E_T$ ,  $|S_i| \geq 3$  und  $|S_j| \geq 3$ , dann sind weder  $u$  noch  $v$  Zentren von  $S_i$  und  $S_j$ . Seien  $c_i$  und  $c_j$  Zentren von  $S_i$  und  $S_j$ , so dass  $\{c_i, c_j\} \cap \{u, v\} = \emptyset$ . Daraus folgt:  $(c_i, u), (c_i, v), (c_j, u), (c_j, v) \in E_T$ . Wir bekommen also einen Kreis, was ein Widerspruch zu der Baumeigenschaft ist. **c)** angenommen,  $u$  ist das Zentrum von  $S_i$  in  $T$ . Sei  $c_j$  das Zentrum von  $S_j$  in  $T$ . Da  $S_j$  einen Stern in  $T$  induziert, folgt:  $(c_j, u) \in E_T$ . Da aber  $c_j \neq u$  und  $c_j \in S_j$ , folgt  $c_j \notin S_i$  (da es gilt:  $S_i \cap S_j = \{u\}$ ), und somit induziert  $S_i$  keinen vollständigen Stern in  $T$ , was ein Widerspruch ist.  $\square$

**Notation 3.2.** Nach der Beobachtung 3.1.e) schreiben wir für eine eindeutige Menge  $C(\mathbf{S}, T)$  von Paaren von Teilmengen und ihren Zentren, die zu der zulässigen Lösung  $T$  korrespondieren  $C(\mathbf{S}, T) = \{(S_i, c_i) : c_i \in S_i \in \mathbf{S}\}$ , wobei  $c_i$  das Zentrum von  $S_i$  in  $T$  ist.

### 3.1 Hilfskonstruktion

- (i) Zunächst konstruieren wir  $A(\mathbf{S})$ , eine Hilfsmenge von  $\mathbf{S}$ , die aus (nicht-leeren) Durchschnitten von je zwei Teilmengen von  $\mathbf{S}$  und den elemententigen Mengen besteht. Wenn einer der folgenden Fälle eintritt, wird die Konstruktion gestoppt:
- $A(\mathbf{S})$  enthält mehr als  $n - 1$  Teilmengen von Größe zwei,
  - oder in  $A(\mathbf{S})$  existiert eine Teilmenge von Kardinalität größer als zwei,
  - oder es existiert eine Teilmenge von Größe zwei, welche in mehr als zwei Teilmengen von  $\mathbf{S}$  enthalten ist.
- (ii) Konstruktion eines Graphen  $GR$ : sei  $ER = (u, v) : \{u, v\} \in A(\mathbf{S})$ , d.h. eine Kante  $(u, v)$  ist in  $ER$ , wenn  $S_i \cap S_j = \{u, v\}$ . Sei  $VR$  die Menge aller Endknoten der Kanten in  $ER$ , dann definiere wir:  $GR := (VR, ER)$ . Wenn  $GR$  kein Wald ist, brechen wir die Konstruktion ab.

- (iii) Partition des Graphen  $GR$  in zwei Graphen,  $H_2$  und  $H_3$ .  $H_2$  ist ein Graph, der aus allen Zusammenhangskomponenten von  $GR$  besteht, die genau zwei Knoten haben.  $H_3$  ist ein Graph, der aus allen Zusammenhangskomponenten von  $GR$  besteht, die drei und mehr Knoten haben.
- (iv) Partition der Menge  $S$  in drei disjunkten Mengen,  $S_2$ ,  $S_3$  und  $S_f$ . Wir definieren die Mengen wie folgt:  $S_2 := \{S \in S : \exists e \in H_2, V(e) \subseteq S\}$  (d.h.  $H_2$  induziert eine Partition von  $V(H_2)$  in disjunkte Paare von Teilmengen),  $S_3 := \{S \in S : \exists e \in H_3, V(e) \subseteq S\}$  und  $S_f$  ist eine Menge, die alle anderen Teilmengen von  $S$  enthält, d.h.  $S_f = S \setminus (S_2 \cup S_3)$ .

**Lemma 3.3.** *Nehmen wir an, dass eine COS-Instanz (mit  $(V, S)$ ) einen zulässigen Lösungsbaum besitzt. Dann kann die Hilfskonstruktion durchgeführt werden, ohne gestoppt zu werden.*

*Beweis.* Nach Beobachtung 3.1.b) ist jede Kante von  $ER$  eine Kante in jedem zulässigen Lösungsbaum. Daher enthält  $A(S)$  nicht mehr als  $n-1$  Teilmengen von Größe zwei, denn ein Lösungsbaum enthält  $n-1$  Kanten. Der Graph  $GR$  ist ein Wald, denn wenn die Kanten von  $GR$  nicht aus einem Wald wären, könnten sie auch nicht ein Teil eines Lösungsbaumes sein. Nach der Beobachtung 3.1.a) gibt es keine Teilmenge von  $A(S)$  der Kardinalität größer als zwei und es gibt keine Teilmenge von  $A(S)$  von Größe zwei, die mehr als in zwei Teilmengen von  $S$  enthalten ist.  $\square$

## 4 Strukturtheorem für das COS

**Notation 4.1.** *Sei  $F(S_i)$  eine Teilmenge von  $S_i$ , die diejenigen Knoten von  $S_i$  enthält, die nicht in anderen Teilmengen von  $S$  enthalten sind. Ein Knoten in  $F(S_i)$  heißt freier Knoten von  $S_i$ .*

**Theorem 4.2 (Strukturtheorem).** *Nehmen wir an, dass eine COS-Instanz (mit  $(V, S)$ ) einen zulässigen Lösungsbaum besitzt. Dann kann die Hilfskonstruktion nach dem Lemma ohne Abbruch durchgeführt werden, und es gilt:*

- I.**
1. Für jedes  $S_i \in S_3$  existiert ein eindeutig bestimmter Knoten  $c_i \in S_i$ , so dass  $c_i$  das Zentrum von  $S_i$  in jedem zulässigen Lösungsbaum ist.
  2. Für  $S_i, S_j \in S_2$  in jedem zulässigen Lösungsbaum  $T$  mit  $S_i \cap S_j = \{u, v\}$  (d.h.  $(u, v) \in H_2$ ) ist entweder  $u$  das Zentrum von  $S_i$  und  $v$  das Zentrum von  $S_j$  oder umgekehrt.
  3. Für jedes  $S_i \in S_f$  in jedem zulässigen Lösungsbaum  $T$  ist das Zentrum von  $S_i$  einer von den freien Knoten von  $F(S_i)$ .
- II.** Für jede Teilmenge  $S_i \in S$  wählen wir das jeweilige Zentrum folgendermassen aus:

1. a) Wir betrachten jedes  $v \in \mathbf{H}_3$ , das kein Blatt von  $\mathbf{H}_3$  ist. Sei  $S_i \in \mathbf{S}_3$  mit  $S_i \supseteq N_{\mathbf{H}_3}[v]$ . Dann setzen wir  $v$  als Zentrum von  $S_i$ .
- b) Wir betrachten jedes  $u \in \mathbf{H}_3$ , das ein Blatt von  $\mathbf{H}_3$  ist. Sei  $S_i, S_j \in \mathbf{S}_3$  mit  $S_i \cap S_j = \{u, v\}$ , wobei wir  $v$  als Zentrum von  $S_i$  in a) gesetzt haben. Dann setzen wir  $u$  als Zentrum von  $S_j$ .
2. Für  $S_i, S_j \in \mathbf{S}_2$  mit  $S_i \cap S_j = \{u, v\}$  setzen wir den Knoten  $u$  als Zentrum von  $S_i$  und den anderen Knoten  $v$  als Zentrum von  $S_j$ .
3. Für jedes  $S_i \in \mathbf{S}_f$  wählen wir einen freien Knoten  $u \in F(S_i)$  zum Zentrum von  $S_i$ .

Anschließend verbinden wir jedes Zentrum  $c_i \in S_i$  per Kanten mit allen anderen Knoten in  $S_i$ . Wir bekommen einen zulässigen Lösungsbaum  $T^*$  für  $(V, \mathbf{S})$ . Außerdem gilt: wenn  $C$  eine Menge von Paaren  $(S_i, c_i)$  (mit  $c_i$  das Zentrum von  $S_i$ ) ist, dann ist  $T^*$  der eindeutige zulässige Lösungsbaum  $T$  mit  $C = C(\mathbf{S}, T)$

*Beweis. I.* 1. **(a) Fall 1:** Sei  $c \in \mathbf{H}_3$  kein Blatt.

Behauptung: es gibt genau eine Teilmenge  $S \in \mathbf{S}_3$  mit  $S \supseteq N_{\mathbf{H}_3}[c]$ .

Jede Kante in  $\mathbf{H}_3$  ist in jedem Durchschnitt zweier Teilmengen enthalten. Seien  $S_{i_1}, S_{j_1}, \dots, S_{i_p}, S_{j_p}$  alle Paare von Teilmengen von  $\mathbf{S}$ , so dass  $S_{i_k} \cap S_{j_k} = \{c, u_k\}$  für  $k = 1, \dots, p$ . Nach Beobachtung 3.1.b) folgt, dass  $c$  entweder das Zentrum von  $S_{i_k}$  oder von  $S_{j_k}$  ist, d.h.  $c$  ist das Zentrum einer der Teilmengen jedes Paares. In jedem zulässigen Lösungsbaum kann ein Knoten Zentrum höchstens einer Teilmenge sein, daher müssen alle Teilmengen, die  $c$  als Zentrum haben, gleich sein. Außerdem enthalten diese Teilmengen  $c$  und  $u_1, \dots, u_p$ . Es folgt die Behauptung.

$c$  hat mindestens zwei Nachbarn in  $\mathbf{H}_3$ ,  $u_i$  und  $u_j$  mit  $u_i \neq u_j$ . Sei  $S_i \in \mathbf{S}$  mit  $S \cap S_i = \{c, u_i\}$  und  $S_j \in \mathbf{S}$  mit  $S \cap S_j = \{c, u_j\}$ . Nach der Beobachtung 3.1.b) sind  $c$  und  $u_i$  Zentren von  $S$  und  $S_i$ , genauso sind  $c$  und  $u_j$  Zentren von  $S$  und  $S_j$ . Folglich muss  $c$  das Zentrum von  $S$  sein.

**(b) Fall 2:** Sei  $c \in \mathbf{H}_3$  ein Blatt.

Sei  $u$  ein Nachbar von  $c$  in  $\mathbf{H}_3$ . Es ist klar, dass  $u$  kein Blatt ist. Sei  $S_u$  eine eindeutige Menge mit dem Zentrum in  $u$ , wie es in I.1.(a) gezeigt wurde. Nach der Beobachtung 3.1.b) gilt:  $(c, u) \in \mathbf{H}_3$ . Für  $S$  mit  $S \cap S_u = \{c, u\}$  ist das Blatt  $c$  das eindeutige Zentrum von  $S$ .

2. Nach der Beobachtung 3.1.b) folgt, dass entweder  $u$  das Zentrum von  $S_i$  in  $T$  und  $v$  das Zentrum von  $S_j$  in  $T$  ist, oder umgekehrt.

OBdA sei  $u$  das Zentrum von  $S_i$  in  $T$  und  $v$  das Zentrum von  $S_j$  in  $T$ .

Behauptung: es gibt einen zulässigen Lösungsbaum  $\tilde{T}$ , so dass  $v$  das Zentrum von  $S_i$  in  $\tilde{T}$  und  $u$  das Zentrum von  $S_j$  in  $\tilde{T}$  ist, und alle anderen Zentren von  $\tilde{T}$  entsprechen denen von  $T$ .

Da  $u$  und  $v$  nur in  $S_i$  und  $S_j$  enthalten sind, ist es offensichtlich, dass das Vertauschen von  $u$  und  $v$  als Zentren keinen Einfluß auf die anderen Teilmengen nimmt. Die zwei Sterne, die in  $u$  und  $v$  zentriert sind, enthalten disjunkte Knoten, also entstehen keine Kreise. Es folgt die Behauptung.

3. Sei  $S_i \in \mathbf{S}_f$  und sei  $u$  das Zentrum von  $S_j$ .

Behauptung:  $u \in F(S_i)$  und  $\forall u \in F(S_i)$  mit  $v \neq u$  gibt es einen zulässigen Lösungsbaum  $\tilde{T}$ , so dass  $v$  das Zentrum von  $S_i$  in  $\tilde{T}$  ist, und dass alle anderen Zentren von  $\tilde{T}$  mit den Zentren von  $T$  übereinstimmen.

Die Beobachtung 3.1.c) impliziert, dass  $u$  in  $F(S_i)$  sein muss. Es ist offensichtlich, dass wir nach dem Verlegen des Zentrums von  $u$  nach  $v$  einen neuen Baum  $\tilde{T}$  bekommen, so dass  $S_i$  einen vollständigen Stern induziert. Alle anderen Teilmengen und ihre Zentren bleiben unverändert. Also ist  $\tilde{T}$  ein neuer zulässiger Lösungsbaum. Es folgt die Behauptung.

- II.** Sei  $T$  ein zulässiger Lösungsbaum für eine *COS*-Instanz mit  $(V, \mathbf{S})$ . Sei  $C(\mathbf{S}, T)$  eine eindeutige Menge von Zentren (wie in der Beobachtung 3.1.e) und sei  $C$  die Menge der in **(II)** gewählten Zentren. Wir konstruieren eine Folge von zulässigen Lösungsbäumen:  $T_0 = T$ ,  $T_1, \dots, T_k = T^*$ , die die folgende Eigenschaft erfüllt:

$$|C(\mathbf{S}, T_{r+1}) \setminus C| < |C(\mathbf{S}, T_r) \setminus C|, 0 \leq r \leq k-1 \quad (1)$$

**Schritt 0:** Aus dem Beweis von **(I.1.)** folgt, dass wir bei der Auswahl von Zentren für jedes  $S_i \in \mathbf{S}_3$  in **(II.)** das gleiche Zentrum bekommen, welches  $S_i$  in  $T_0$  hat. Es gilt:  $c_i$  ist genau dann das Zentrum von  $S_i \in \mathbf{S}_3$  in  $T_0$ , wenn  $c_i$  als Zentrum von  $S_i$  in **(II.1.)** gewählt wurde.

**Schritt 1:** Wir wiederholen den Wahlvorgang bis das Zentrum von jedem  $S_i \in \mathbf{S}_2$  mit dem von  $S_i$  aus dem Teil **(II.2)** übereinstimmt. Seien  $S_i, S_j \in \mathbf{S}_2$  mit  $S \cap S_j = \{u, v\}$ , so dass ihre Zentren in  $T_r$  nicht dieselben sind als in **(II.2)**. Dann stellen wir  $T_r$  auf  $T_{r+1}$  wie folgt um: wenn  $u$  das Zentrum von  $S_i$  und  $v$  das Zentrum von  $S_j$  in  $T$  ist, dann können wir nach dem Beweis von **(I.2.)**  $T_r$  verändern, indem wir  $v$  zum Zentrum von  $S_i$  und  $u$  zum Zentrum

von  $S_j$  wählen. Der Rest von  $T_r$  bleibt unverändert. Auf diese Weise bekommen wir einen neuen zulässigen Lösungsbaum  $T_{r+1}$ .

**Schritt 2:** Wir wiederholen den Wahlvorgang bis das Zentrum von jedem  $S_i \in \mathbf{S}_f$  in  $T_r$  das gleiche ist wie das Zentrum von  $S_i$  wie im Teil (II.3) Nach dem Beweis von (I.3.) können wir  $T_r$  so verändern, um einen neuen zulässigen Lösungsbaum  $T_{r+1}$  zu bekommen. Wir tauschen das Zentrum von  $S_i$  so aus, dass es mit dem in (II.3.) gewählten Zentrum übereinstimmt. Der Rest von  $T_r$  bleibt unverändert.

Wir sind also mit einem zulässigen Lösungsbaum  $T_0$  gestartet und haben gezeigt: wenn  $T_r$  ein zulässiger Lösungsbaum ist, dann ist auch  $T_{r+1}$  ein zulässiger Lösungsbaum, und die Eigenschaft (1) wird erfüllt. Daher hat die Folge höchstens  $|V|$  Bäume. Es ist offensichtlich, dass der letzte Baum  $T_k$  der Folge die Eigenschaft  $|C(\mathbf{S}, T_k) \setminus C| = 0$  besitzt, und daher ist  $T_k = T^*$  ein zulässiger Lösungsbaum.

Um die Eindeutigkeit zu zeigen, nehmen wir an, dass es zwei verschiedene zulässige Lösungsbäume  $T_1$  und  $T_2$  gibt. Nach Beobachtung 3.1.e) gibt es eine eindeutige Menge von Zentren  $C(\mathbf{S}, T_1)$  (sie korrespondiert mit  $T_1$ ) und eine eindeutige Menge von Zentren  $C(\mathbf{S}, T_2)$  (sie korrespondiert mit  $T_2$ ), so dass  $C = C(\mathbf{S}, T_1) = C(\mathbf{S}, T_2)$ . Angenommen, die Kante  $(u, v)$  liegt in  $T_1$ , aber  $(u, v) \notin T_2$ . Da  $(u, v) \in T_1$ , folgt nach Beobachtung 3.1.d):  $\exists S_i \in \mathbf{S} : u, v \in S_i$ . Sei  $w$  das Zentrum von  $S_i$  in  $T_1$ . Wenn  $w \neq u$  und  $w \neq v$ , dann bekommen wir einen Kreis in  $T_1$ . Daraus folgt, dass entweder  $u$  oder  $v$  das Zentrum von  $S_i$  in  $T_1$  und auch das Zentrum von  $S_i$  in  $T_2$  ist. Es folgt:  $(u, v) \in T_2$ , was ein Widerspruch zur Voraussetzung ist. □

## 5 Ein polynomieller Algorithmus für das *COS*

### 5.1 Kernidee des Algorithmus

Nach dem Strukturtheorem induziert jedes  $S_i \in \mathbf{S}_3$  denselben eindeutigen Stern in jedem zulässigen Lösungsbaum. Die Teilmengen von  $\mathbf{S}_2$  sind in Paare geteilt. Die Wahl von Zentren für jedes Paar ist unabhängig von der Wahl von Zentren für alle anderen Teilmengen von  $\mathbf{S}_2$  und von  $\mathbf{S}_f$ . Deswegen wählen wir aus zwei Alternativen die mit dem minimalen Gewicht aus. Für jede Teilmenge  $S_i \in \mathbf{S}_f$  wählen wir das Zentrum nur aus den freien Knoten von  $S_i$ ,  $F(S_i)$ , aus. Diese Wahl ist unabhängig von der Wahl von Zentren für alle anderen Teilmengen von  $\mathbf{S}_f$  und von  $\mathbf{S}_2$ . Deswegen wählen wir die Alternative mit dem minimalen Gewicht aus. Zusammengefasst, wir wählen die Zentren aus, die lokal optimal sind.

## 5.2 Algorithmus für das COS

**Algorithmus 1.** (Ein Algorithmus für das COS)

**Schritt 1 - Hilfskonstruktion:** Führe die Hilfskonstruktion durch.

- Wenn sie in (i) oder in (ii) gestoppt wird, dann existiert keine Lösung, und der Algorithmus stoppt.
- Betrachte die Menge  $A(\mathcal{S})$  und alle einelementige Mengen  $\{u\} \in A(\mathcal{S})$ . Wenn  $u$  in  $GR$  enthalten ist, dann existiert keine Lösung, und der Algorithmus stoppt.

**Schritt 2 - Wahl eines Zentrums für jede Teilmenge von  $A(\mathcal{S})$  :**

1. Für jede Teilmenge  $S_i \in \mathcal{S}_3$  wähle einen Knoten aus  $V(H_3)$ , wie im Strukturtheorem (II.1.) beschrieben wird, zum Zentrum von  $S_i$ .
2. Für jedes Paar von Teilmengen  $S_i, S_j \in \mathcal{S}_2$ , die von  $H_2$  induziert werden, so dass gilt:  $(u, v) \in S_i, S_j$ , gibt es zwei mögliche Alternativen:
  - $u$  ist das Zentrum von  $S_i$  und  $v$  das Zentrum von  $S_j$  mit Gewicht  $w(S_i, u) + w(S_j, v)$ , oder
  - umgekehrt mit dem Gewicht  $w(S_i, v) + w(S_j, u)$ .

Wähle die Alternative aus, bei der das Gewicht am kleinsten ist.

3. Für jede Teilmenge  $S_i \in \mathcal{S}_f$  bestimme die Menge freier Knoten  $F(S_i)$ .
  - Wenn  $F(S_i) = \emptyset$ , dann existiert keine Lösung, und der Algorithmus stoppt.
  - Anderenfalls gibt es andere mögliche Sterne von  $S_i$  zu beachten: berechne für jeden Knoten  $f_i \in F(S_i)$  das Gewicht  $w(S_i, f_i)$ .

Wähle eine Alternative mit dem minimalen Gewicht aus.

**Schritt 3 - Konstruktion eines Lösungsbaumes:** Nicht jede Teilmenge besitzt ein Zentrum. Verbinde in jeder Teilmenge jeden Knoten, der von dem Zentrum verschieden ist, mit dem Zentrum, wenn sie noch nicht miteinander verbunden sind. Man bekommt folglich einen zusammenhängenden Graphen  $T$ .

- Wenn  $T$  kein Baum ist, dann existiert keine Lösung, und der Algorithmus stoppt.
- Anderenfalls ist  $T$  der optimale Lösungsbaum.

Ende des Algorithmus.

**Theorem 5.1.** Sei  $(V, \mathbf{S})$  eine *COS*-Instanz, wobei  $|V| = n$  die Anzahl der Knoten und  $|\mathbf{S}| = m$  die Anzahl der Teilmengen von  $V$  ist. Dann gilt:

- a) Der Algorithmus 1. löst das *COS* korrekt.
- b) Die Komplexität des Algorithmus 1. ist  $O(m^2n + n^2)$ .

*Beweis.* a) Aus dem Teil (II) des Strukturtheorems folgt, dass der Algorithmus 1. korrekt abläuft.

b) In einer *COS*-Instanz kann ein Knoten das Zentrum von höchstens einer Teilmenge sein, daraus folgt:  $m \leq n$ . Im Schritt 1. des Algorithmus werden nur Schnitte von jeweils zwei Teilmengen berechnet, und daher wird  $A(\mathbf{S})$  in  $O(m^2n)$  Zeit konstruiert. Im Schritt 2.3. wird die Bestimmung bestmöglicher Zentren für alle Teilmengen von  $\mathbf{S}_f$  in  $O(n^2)$  Zeit durchgeführt. Im Schritt 3. wird lineare Laufzeit gebraucht, um einen zusammenhängenden Lösungsbaum zu konstruieren und um zu prüfen, ob diese Konstruktion ein Baum ist. Insgesamt beträgt die Komplexität des Algorithmus  $O(m^2n + n^2)$ .  $\square$

## 6 Fazit

In dieser Arbeit haben wir das Problem der Konstruktion eines kosten-(gewichts-)minimalen Gerüstes, eines stars-clustering-tree betrachtet. Wir haben ein Strukturtheorem und einen polynomiellen Algorithmus für das Problem betrachtet. Außerdem haben wir gezeigt, dass die Variante des Problems, in der jede Teilmenge einen vollständigen Stern bis auf höchstens  $k$  Blätter induziert, NP-schwer ist. Die Komplexität des *OS* ist bis jetzt noch offen, und es wäre interessant, dies zu lösen. Ein anderes interessantes Problem, das analog zu dem *COS* ist, wäre, ein gewichtsminimales Netzwerk zu konstruieren, so dass dieses Netzwerk nicht unbedingt ein Baum sein muss.

## Literatur

- [1] E. KORACH, M. STERN: The complete optimal stars-clustering-tree problem. *Discrete Applied Mathematics* 156 (2008) 444-450.
- [2] M.R. GAREY, D.S. JOHNSON: *Computers and Intractability*. Freeman, San Francisco, 1979.