Themen für Bachelorarbeiten

aus dem Bereich der

Kombinatorischen Optimierung

Dr. Eranda Dragoti-Cela
Institut für Diskrete Mathematik, TU Graz

März 2016

1. (VERGEBEN) Das klassische (un)kapazitierte Standortproblem ((un)capacitated facility location -(U)FLP)

   Der Input des unkapazitierten Standortproblems (UFLP) besteht aus einer endlichen Menge $\mathcal{F}$ von potenziell zu errichtenden Einrichtungen (Dienstleistungstätten, Spitäler, Schulen), $|\mathcal{F}| =: n_{\mathcal{F}}$, einer Menge von Kunden $\mathcal{C}$, $|\mathcal{C}| = n_{\mathcal{C}}$, sowie einer Abbildung $u\colon \mathcal{F} \to \mathbb{R}_+$, die die Errichtungskosten $u(i)$ für jede Einrichtung $i \in \mathcal{F}$ angibt, und einer Abbildung $c\colon \mathcal{F} \times \mathcal{C} \to \mathbb{R}_+$, die die Servicekosten $c(i,j)$ für die Betreuung/Bedienung von Kunde j durch die Einrichtung i angibt, für alle $i \in \mathcal{F}$, $j \in \mathcal{C}$. Das Ziel ist, eine Menge $F \subset \mathcal{F}$ von zu errichtenden Einrichtungen zu bestimmen, sowie eine Zuordnung $\phi\colon \mathcal{C} \to F$ der Kunden zu den errichteten Einrichtungen, sodass die Gesamtkosten $\sum_{i\in F} u(i) + \sum_{j\in\mathcal{C}} c(i,\phi(i))$ minimiert werden.

   Beim kapazitierten Standortproblem (CFLP) sind zusätzlich obere Grenzen b_i für die maximale Anzahl von Kunden, die von Einrichtung i betreut/bedient werden können, $i \in \mathcal{C}$.

   Beim metrischen UFLP (CFLP) werden die Servicekosten als Verbindungskosten angesehen, d.h. $c(i,j)$ sind die Kosten der Verbindung zwischen i und j, wobei $i,j \in \mathcal{F} \cup \mathcal{C}$, und es wird angenommen, dass die Verbindungskosten die Dreiecksungleichung erfüllen, d.h. $c(i,j) \le c(i,k) + c(k,j)$ gilt, für alle $i,j,k \in \mathcal{F} \cup \mathcal{C}$.

   UFLP und CFLP sind NP-schwer, und die metrischen Varianten ebenfalls. Beide Probleme sind jedoch mit garantierter Approximationsgüte approximierbar, die metrischen Varianten sogar mit konstanter Approximationsgüte. Dementsprechend viele Approximatiosnalgorithmen für diese Probleme wurden in der Literatur vorgestellt und analysiert [3, 13, 14, 18]. Die grundlegenden Ideen reichen von intutiv einfachen greedy Ansätzen bis zu komplexeren Ansätzen, die auf LP-Relaxationen von ganzzahligen linearen Formulierungen der Probleme basieren, und primal-duale- oder Rundungstechniken anwenden.

   Aus der Sicht der praktischen Lösbarkeit der Probleme wurden in der Literatur unzählige Heuristiken vorgestellt und analysiert; die Ansätze reichen von lokalen Suchverfahren bis zu Metaheuristiken und genetischen Algorithmen.

   Im Rahmen einer Bachelorarbeit könnten Approximationslagorithmen für die eine oder andere Variante des Problems behandelt werden, natürlich nach einer allgemeinen Einführung in dem das Problem und einige seiner Eigenschaften erläutert werden, inkl. ganzzahlige lineare Formulierungen und deren kontinuierlichen Relaxationen.

2. (VERGEBEN) Das Feuerbekämpfungproblem in Bäumen (the firefighting problem on trees (FPT))

Der Input des Feuerbekämpfungproblems (FP) bestehet aus einem Graphen $G = (V, E)$, in dem ein Knoten $r \in V$ als Wurzel gewählt wird. Das Problem ist in diskreten Zeitpunkten i, $i = 0, 1, \ldots, n$, $n := |V|$, definiert. Zum Zeitpunkt 0 fängt die Wurzel Feuer. Von einem zum Zeitpunkt i brennenden Knoten aus verbreitet sich das Feuer auf die noch nicht brennenden Nachbarn dieses Knoten in einem Zeitschritt, d.h. vom Zeitpunkt $i-1$ zum Zeitpunkt i. Es wird angenommen, dass die Feuerwehr zu jedem Zeitpunkt genau einen nicht brennenden Knoten vor Feuer schützen kann; der Schutz bleibt dann in Zukunft für immer erhalten. Spätestens nach n Zeitschritten ist der Prozess zu Ende, denn dann ist jeder Knoten, der noch nicht brennt, geschützt. Klarerweise hängt die Anzahl der am Ende des Prozesses brennenden Knoten von der Wahl der geschützten Knoten im Laufe des Prozesses ab. Gesucht ist eine Strategie für die Wahl des zu schützenden Knoten in jedem Zeitpunkt, sodass die Anzahl der nicht brennenden Knoten am Ende des Prozesses maximiert wird. Wenn der Graph G ein Baum ist dann spricht man von dem Feuerbekämpfungproblem in Bäumen (FTP).

Eine andere Variante des Problems ist das „Resource minimization for fire containment on trees" (RMFCT) in dem die Feuerwehr in jedem Zeitpunkt nicht nur einen sondern bis zu B Knoten vor Feuer schützen kann. Gesucht ist die kleinste Zahl $B \in \mathbb{N}$, sodass am Ende des Prozesses kein Blatt des Baums brennt.

Zum ersten Mal wurde das Feuerbekämpfungproblem (FP) von Hartnell in 1995 [12] vorgestellt. Das RMFCT wurde zum ersten Mal von Chalermsook and Chuzhoy in 2010 [4] vorgestellt.

Das FPT ist NP-schwer auch in Bäumen mit maximalem Knotengrad gleich 3, siehe [7], beim RMFCT ist es NP-schwer zu enscheiden ob der minimaler Wert von B gleich 1 ist, somit kann es keinen Approximationslagorithmus mit Approximationsgüte kleiner als 2 geben, falls $P \neq NP$, siehe [16]. Für beide Probleme gibt es in der Literatur jedoch einige Approximationsalgorithmen mit konstanter Gütegarantie, für das FPT gibt es sogar ein polynomielles Approximationschema (PTAS), siehe [2] und die darin enthaltenen Referenzen.

Weiters gibt es Ergbnisse über spezielle Graphenklassen, in denen das FP polynomiallösbar ist, etwa für ganz spezielle Bäume [7], oder Intervallgraphen, Splitgraphen [9]. Es gibt auch einen Überblicksartikel von Finbow und MacGillivray [8], veröffentlicht in 2009.

Im Rahmen einer Bachelorarbeit könnte eine allgemeine Beschreibung des Problems und unterschiedliche mathematische Formulierungen vorgestellt werden. Weiters könnten entweder Approximationsalgorithmen oder polynomiallösbare Spezialfälle behandelt werden.

3. Kombinatorische Optimierungsprobleme mit Budgetrestriktionen (combinatorial optimization with budget constraints)

   Viele klassische polynomial lösbare kombinatorische Optimierungsprobleme wie das minimale Spanbaumproblem, das maximale Matchingproblem, das kurzeste Wegeproblem, können als Spezialfälle des folgenden generischen Problems betrachtet werden. Sei $\mathcal{F}$ eine Menge von zulässigen Lösungen und $w \colon \mathcal{F} \to \mathbb{R}$ eine Abbildung, die jeder zulassigen Lösung $S \in \mathcal{F}$ ein Gewicht $w(S)$ zuordnet. Gesucht wird eine zulässige Lösung S^* mit maximalem Gewicht, i.e. $w(S^*) = \max\{w(S) \colon S \in \mathcal{F}\}$. Betrachten wir nun eine Modifikation des Problems in der den zulässigen Lösungen neben den Gewichten $w(S)$ auch Kosten $c(S)$, $S \in \mathcal{F}$, zugeordnet werden. Weiters gäbe es ein Budget B, das nicht überschritten werden darf, d.h. zulässig sind nur jene Lösungen S aus $\mathcal{F}$, für die $c(S) \leq B$ gilt. Das Optimierungsproblem lautet nun „Bestimme $S^* \in \mathcal{F}$, sodass $w(S^*) = \max\{w(S) \colon S \in \mathcal{F}, c(S) \leq B\}$". In diesem Fall spricht man von einem Optimierungsproblem mit einer Budgetrestriktion. Wenn es mehrere Kostenfuktions $c_i \colon \mathcal{S} \to \mathbb{R}$ und mehrere Budgetgrenzen B_i, $1 \leq i \leq k$, $k \in \mathbb{N}$, gibt, so spricht man von einem Optimierungsproblem mit k Budgetrestriktionen; gesucht wird nun ein $S^* \in \mathcal{F}$, sodass $w(S^*) = \max\{w(S)\colon S \in \mathcal{F}, c_i(S) \leq B_i, 1 \leq i \leq k\}$.

Alle oben genannten Probleme sind bereits NP-schwer für $k = 1$, es gibt jedoch polynomielle Approximationsschemas für alle drei Probleme in diesem Fall. Das polynomielle Approximationsschema für das maximale Matchingproblem mit einer Budgetrestriktion beruht auf die Lagrange Relaxation des Problems bei der die Budgetrestriktion relaxiert wird. Zwei Lösungen dieser Lagrange Relaxation werden geschickt miteinander kombiniert um eine approximative Lösung des ursprünglichen Problems zu erhalten. Als Nebenergebnis wird dabei auch ein altes kombinatorisches Puzzle gelöst, siehe [5].

Für $k \geq 2$ ist die Situation generell eine andere; das kürzeste Wege Problem und das Spannbaumproblem sind nicht approximierbar für $k \geq 2$, siehe [11]. Für das maximale matching Problem hingegen gibt es im Fall $k = 2$ ein polynomielles Approximationsschema [11].

Im Rahmen einer Bachelorarbeit könnten die PTAS für das maximale Matchingproblem mit einer Budgetrestriktion bzw. zwei Budgetrestriktionen behandelt werden.

4. (VERGEBEN) Optimierungsmodelle für die Aggregation unterschiedlicher Reihungen (ranking aggregation)

   Das Problem der Aggregation unterschiedlicher Reihungen tritt bei diversen Anwendungen in der (multikriteriellen) Entscheidungstheorie, in der Reihung von Webseiten, in Sport und in der künstlichen Intelligenz auf. Sehr allgemein formuliert besteht das Problem darin, aus einer Großzahl von einzelnen Reihungen eine optimale Endreihung zu generieren, die - den unterschiedlichen Kriterien entsprechend - die einzelnen Inputreihungen am besten repräsentiert. Ein mögliches Maß der Repräsentanz einer aggregierten Reihung sind die Abweichungen der aggregierten Reihung von den einzelnen Reihungen. Diese Abweichungen würden in eine Zielfunktion einfließen, die große Abweichung bestraft. Ein ideales Modell sollte auch unterschiedliche Präferenzen des Entscheidungsträgers bzw. unterschiedliche Vertrauensgrade der jeweiligen Reihungen, der Expertise deren Autoren entsprechend, berücksichtigen.

   Ein in der Literatur gut untersuchtes Modell für diese Problem modelliert die einzenen Reihungen als Total- oder Partialordnungen O_i, $1 \leq i \leq k$, in der endlichen Menge der zu reihenden Objekte. Der Abstand zweier Ordnungen O_i, O_j wird anhand der sogenannten „Kendall's tau distance“ $d(O_i, O_j)$ gemessen; $d(O_i, O_j)$ ist die Anzahl der Nicht-Übereinstimmungen in den Ordnungen O_i, O_j. Dabei ist eine Nicht-Übereinstimmung ein Paar von Objekten p, q, $p \neq q$, die von den zwei Ordnungen O_i, O_j unterschiedlich gereiht werden, d.h. es gilt entweder $p \prec_{O_i} q$ und $q \prec_{O_j} p$, oder $q \prec_{O_i} p$ und $p \prec_{O_j} q$. Es wird eine Ordnung O der Objekte gesucht, die $\sum_{i=1}^{k} d(O, O_i)$ minimiert. Die Bestimmung einer solchen Ordnung O ist NP-schwer für $k \geq 4$, siehe [19]. Das Problem ist von großer praktischer Relevanz, daher werden in der Literatur viele heuristische Ansätze beschrieben und verglichen, siehe [6].

   Das Problem kann auch als Optimierungsproblem in einem gerichteten gewichteten Graphen $G = (V, E)$ formuliert werden. Die zu reihenden Objekte stellen die Knoten des Graphen dar und die Kantengewichte $w(j, l)$ entsprechen dem Anteil der Ordnungen O_i, $1 \leq i \leq k$, für die $j \prec_{O_i} l$ gilt. Es wird nun eine Reihung (oder Permutation) σ der Knoten in V gesucht, die $\sum_{i,j\colon \sigma(i)<\sigma(j)} w(j, i)$ minimiert („scaled Kemeny aggregation“). Wenn O_i, $1 \leq i \leq k$, Partialordnungen sind, dann erfüllen die Gewichte w_{ij} die Dreiecksungleichen. Wenn O_i, $1 \leq i \leq k$, Totalordnungen sind, dann sind darüber hinuas auch die sogenannten Wahrscheinlichkeitsgleichungen $w(i, j) + w(j, i) = 1$, $i, j \in V$, $i \neq j$, erfüllt. In den letzten 10 Jahren wurden für dieses

Problem einige Approximationsalgorithmen mit konstanter Gütegarantie entwickelt, wenn die Gewichte die Dreiecksungleichungen und/oder die Wahrscheinlichkeitsgleichungen erfüllen, siehe [10] und die darinenthaltenen Referenzen. Für den Fall, dass beide Ungleichungen erfüllt sind, gibt es auch ein polynomielles Approximationsschema [15].

Im Rahmen einer Bachelorarbeit könnten unterschiedliche Varianten des Problems und verschiedene Modelle vorgestellt werden. Eine Bachelorarbeit mit einem theoretischen Fokus könnte sich darüber hinaus mit einigen Approximationsalgorithmen befassen. Eine Bachelorarbeit mit einer praktischen Ausrichtung könnte sich mit heuristischen Ansätzen zur praktischen Lösung des Problems, die in der Literatur vorgestellt worden sind, befassen. Die erfolgreichsten Ansätze und deren Ergenisse sollten vorgestellt werden.

5. Gebrochene Färbungprobleme in Graphen (fractional coloring)

   Gebrochene Färbungsprobleme (in English *fractional coloring*) gehören zur sogenannten gebrochenen Graphentheorie (in Englisch *fractional graph theory*) und verallgemeinen die Färbungprobleme aus der klassischen Graphentheorie. Eine b-fache Kantenfärbung (Knotenfärbung) eines Graphen ist eine Zuordnung von (Farb)Mengen mit Kardinalität b zu den Kanten (Knoten) des Graphen, sodass benachbarte Kanten (Knoten) disjunkte Farbmengen erhalten. Eine $a : b$-fache Kantenfärbung (Knotenfärbung) ist eine b-fache Kantenfärbung (Knotenfärbung) mit Farben aus einer Grundmenge von Farben mit Kardinalität a. Der b-fache chromatischer Index $\chi'_b(G)$ (die b-fache chromatische Zahl $\chi_b(G)$) eines Graphen G ist der kleinste Wert von a, für den eine b-fache Kantenfärbung (Knotenfärbung) existiert. Der gebrochene chromatische Index $\chi'_f(G)$ und die gebrochene chromatische Zahl $\chi_f(G)$) sind folgendermaßen definiert:

   $$\chi'_f(G) = \lim_{b\to\infty} \frac{\chi'_b(G)}{b} \quad \chi_f(G) = \lim_{b\to\infty} \frac{\chi_b(G)}{b}\,.$$

   Die gebrochenen Färbungprobleme lassen sich als lineare Optimierungsprobleme formulieren; die Anzahl der Restriktionen dieser linearen Probleme kann aber exponential groß im Vergleich zur Anzahl der Knoten bzw. Kanten des Graphen sein, was die effiziente Berechnung dieser Größen durch das Lösen des enstprechenden linearen Problems generell natürlich unmöglich macht. Während der gebrochenen chromatische Index effizient berechnet werden kann ist die Berechnung der gebrochenen chromatischen Zahl in beliebigen Graphen ein NP-schweres Problem, siehe Scheinerman und Ullman [17].

   Im Rahmen einer Bachelorarbeit Arbeit könnte eine allgeimene Einführung in die gebrochenen Färbungprobleme in Graphen, deren Eigenschaften und Anwendungen gegeben werden. Die Arbeit sollte auch den Zusammenhang der gebrochenen Färbungproblememen mit den dazugehörigen klassischen Analoga und auch die effiziente Berechnung des gebrochenen chromatischen Index erläutern.

6. Das Frequenzzuordnungsproblem (the frequency assignment problem) (FZP)

   Das Frequenzzuordnungsproblem in seinen vielen Varianten entspricht praktischen Fragestellungen in Telekommunikationsnetzwerken. Das grundlegende Problem kann in ein allgemeines Framework folgendermaßen beschrieben. Ein Telekommunikationsnetzwerk wird als Graph modelliert, dessen Knoten Kommunikationszellen (Antennen, Basisstationen udgl.) entsprechen. Die Kanten entsprechen den Verbindungen (Kanälen) zwischen den Kommunikationszellen. Jedem Knoten muss eine gewisse Anzahl von Frequenzen zugeordnet werden, damit der Knoten die ihm zugeordneten Verbindungen realisieren kann. In manchen Anwendungen gibt es für jeden Knoten eine im Voraus festgelegte Menge von Frequenzen, die diesem Knoten zugeordnet werden können. Abhängig von der Topologie des Graphen und den vorgenommenen Frequenzzuordnungen können zwischen bestimmten Verbindungspaaren kommunikationstörende Interferenzen eintreten. Die potentiellen Interferenzen werden oft anhand eines sogenannten Interferenzgraphen modelliert. In manchen Anwendungen muss das Eintreten von Interferenzen gänzlich ausgeschlossen werden, in Anderen wiederum wird das Eintreten von Interferenzen anhand von Straftermen in der Zielfunktion penalisiert. Das Ziel ist es eine Zuordnung von Frequenzen zu den Kommunikationszellen zu finden, die alle Anforderungen erfüllt und eine bestimmte Zielfunktion optimiert. Eine häufig verwendete Zielfunktion ist die Minimierung der Gesamtanzahl der verwendeteten Frequenzen. Eine andere in der Praxis relevante Zielfunktion ist die Minimierung der Differenz zwischen der höchsten und der niedrigsten Frequenz, die demselben Knoten zugeordnet werden.

In der Regel sind die Varianten des FZP NP-schwer und es besteht ein starker Zusammenhang mit Färbungsproblemen in Graphen. Einige Forschungsaspekte im Bereich der Frequenzzuordnungsprobleme sind: gemischt-ganzzahlige Formulierungen, die eine approximative Lösung des Problems ermöglichen, strukturelle Untersuchungen des dazugehörigen Polytops, die Ermittlung von unteren (oberen) Schranken, die Entwicklung von heuristischen Verfahren, die die Ermittlung von guten Lösungen für praktische Probleme ermöglichen.

Als allgemeine Orientierungshilfe und als Quelle von weiteren adequaten und modellspezifischen Referenzen kann die Arbeit von Aardal et al. [1] dienen.

Literatur

[1] K.I. Aardal, S.P.M. Van Hoesel, A.M.C.A. Koster, C. Mannino und A. Sassano, Models and solutions techniques for frequency assignment problems, 4OR, *Operations Research Quarterly* **1(4)**, 261-317, 2003.

[2] David Adjiashvili, Andrea Baggio, Rico Zenklusen, Firefighting on Trees Beyond Integrality Gaps `http://arxiv.org/abs/1601.00271`

[3] Hyung-Chan An, Mohit Singh, Ola Svensson, LP-Based Algorithms for Capacitated Facility Location, `http://arxiv.org/abs/1407.3263`

[4] P. Chalermsook and J. Chuzhoy, Resource minimization for fire containment, in *Proceedings of the 21st Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, 2010, 1334-1349.

[5] André Berger, Vincenzo Bonifaci, Fabrizio Grandoni, Guido Schäfer, Budgeted matching and budgeted matroid intersection via the gasoline puzzle, *Mathematical Programming, Ser. A* **128**, 2011, 355–372.

[6] Cynthia Dwork, Ravi Kumar, Moni Naor, and Dandapani Sivakumar, Rank aggregation methods for the web, in *Proceedings of the Tenth International Conference on World Wide Web*,ACM, 2001, 613–622.

[7] S. Finbow, A. King, G. MacGillivray, and R. Rizzi, The firefighter problem for graphs of maximum degree three, *Discrete Mathematics* **307(16)**, 2094–2105, 2007.

[8] S. Finbow and G. MacGillivray, The firefighter problem: a survey of results, directions and questions, *Australasian Journal of Combinatorics* **43**, 57–77, 2009.

[9] F. V. Fomin, P. Heggernes, and E. J. van Leeuwen, Making life easier for firefighters, in *Fun with Algorithms*, 177–188, 2012.

[10] Daniel Freund and David P. Williamson, Rank Aggregation: New Bounds for MCx http://arxiv.org/abs/1510.00738

[11] Fabrizio Grandoni, Rico Zenklusen, Optimization with More than One Budget http://arxiv.org/abs/1002.2147

[12] B. Hartnell, Firefighter! an application of domination, in *24th Manitoba Conference on Combinatorial Mathematics and Computing*, 1995.

[13] Kamal Jain, Mohammad Mahdian, Evangelos Markakis, Amin Saberi, Vijay V. Vazirani, Greedy Facility Location Algorithms Analyzed using Dual Fitting with Factor-Revealing LP `http://arxiv.org/abs/cs/0207028`

[14] Kamal Jain, Mohammad Mahdian, Amin Saberi, A new greedy approach for facility location problems, *STOC '02, 34th ACM Symposium on Theory of Computing*, 2012, 731–740.

[15] Claire Kenyon-Mathieu and Warren Schudy, How to rank with few errors, in *Proceedings of the 39th ACM Symposium on Theory of Computing*, ACM, 2007, 95–103.

[16] A. King and G. MacGillivray, The firefighter problem for cubic graphs, *Discrete Mathematics* **310(3)**, 614–621, 2010.

[17] E. Scheinerman, D. Ullman, *Fractional Graph Theory*, 2008.
`http://www.ams.jhu.edu/ers/books/fractional-graph-theory-a-rational-approach-to-the-theory-of-graphs`

[18] David Shmoys, Eva Tardos and Karen Aardal, Approximation algorithms for facility location problems, *STOC '97, 29th ACM Symposium on Theory of Computing*, 1997, pages 265–274.

[19] A. Van Zuylen and D. Williamson, Deterministic algorithms for rank aggregation and other ranking and clustering problems, in *Approximation and Online Algorithms, WAOA 2007*, Ch. Kaklamanis und M. Skutella, Hrsg., Spriner, 2007, Seiten