

Fortgeschrittene Netzwerk- und Graph-Algorithmen

Prof. Dr. Hanjo Täubig

Lehrstuhl für Effiziente Algorithmen
(Prof. Dr. Ernst W. Mayr)
Institut für Informatik
Technische Universität München

Wintersemester 2010/11



Dense ℓ -Subgraph

- Graph mit Durchschnittsgrad $\overline{\deg}(G)$ muss nicht unbedingt einen echten Teilgraphen mit gleichem Durchschnittsgrad enthalten.
- Maximum des Durchschnittsgrads eines Teilgraphs auf ℓ Knoten:

$$\gamma^*(G, \ell) = \max \{ \overline{\deg}(G[U]) : U \subseteq V \text{ und } |U| = \ell \}$$

Problem

Problem: **Dense- ℓ -Subgraph**

Eingabe: Graph G , Parameter $\ell \in \mathbb{N}$

Frage: Eine Knotenmenge der Kardinalität ℓ mit maximalem induzierten Durchschnittsgrad

Dense ℓ -Subgraph

- Optimierungsproblem **Dense ℓ -Subgraph** ist \mathcal{NP} -hart, denn Instanz $(G, \ell, \ell - 1)$ des zugehörigen Entscheidungsproblems entspricht der Suche nach einer Clique der Größe ℓ in G .

⇒ Wie sieht es mit Approximation aus?

Greedy-Approximation für Dense ℓ -Subgraph

Algorithmus 10 : Approximation eines ℓ -Subgraph mit hohem $\overline{\text{deg}}$

Input : Graph $G = (V, E)$ und
gerader Parameter $\ell \in \mathbb{N}$ (mit $|V| \geq \ell$)

Output : Menge von ℓ Knoten von G

Sortiere die Knoten in absteigender Reihenfolge ihrer Grade;

Sei H die Menge von $\frac{\ell}{2}$ Knoten von höchstem Grad;

Berechne $N_H(v) = |N(v) \cap H|$ für alle Knoten $v \in V \setminus H$;

Sortiere die Knoten in $V \setminus H$ in absteigender Reihenfolge der N_H -Werte;

Sei R die Menge von $\frac{\ell}{2}$ Knoten von $V \setminus H$ mit den höchsten N_H -Werten;

return $H \cup R$

Greedy-Approximation für **Dense** ℓ -Subgraph

Satz

Sei G ein Graph auf n Knoten und sei $\ell \in \mathbb{N}$ eine gerade natürliche Zahl mit $\ell \leq n$.

Sei $A(G, \ell)$ der Durchschnittsgrad des induzierten Teilgraphen, der vom vorstehenden Algorithmus ausgegeben wird.

Dann gilt:

$$\gamma^*(G, \ell) \leq \frac{2n}{\ell} \cdot A(G, \ell)$$

bzw.

$$A(G, \ell) \geq \frac{\ell}{2n} \cdot \gamma^*(G, \ell)$$

Greedy-Approximation für **Dense ℓ -Subgraph**

Beweis.

- Für Knotenteilmengen $U, U' \subseteq V$ sei $E(U, U')$ die Menge der Kanten mit einem Endpunkt in U und einem Endpunkt in U' .
- Sei $m_U = |E(G[U])|$
- Sei \deg_H der Durchschnittsgrad der $\frac{\ell}{2}$ Knoten von G mit höchstem Grad bezüglich G . Es gilt: $\deg_H \geq \gamma^*(G, \ell)$.
- Man erhält für die Anzahl der Kanten zwischen H und dem Rest $V \setminus H$:

$$|E(H, V \setminus H)| = \deg_H \cdot |H| - 2m_H = \frac{\deg_H \cdot \ell}{2} - 2m_H \geq 0$$

Greedy-Approximation für **Dense ℓ -Subgraph**

Beweis.

- Weil der Algorithmus greedy (gierig) arbeitet, muss der Anteil der Kanten nach R (i.Vgl. zu $V \setminus H$) mindestens so groß sein, wie der Anteil der Knoten:

$$\frac{|E(H, R)|}{|E(H, V \setminus H)|} \geq \frac{|R|}{|V \setminus H|} = \frac{\ell/2}{n - \ell/2} = \frac{\ell}{2n - \ell} > \frac{\ell}{2n}$$

- Also ist die Gesamtzahl der Kanten in $G[H \cup R]$ mindestens

$$\left(\frac{\deg_H \cdot \ell}{2} - 2m_H \right) \cdot \frac{\ell}{2n} + m_H \geq \frac{\deg_H \cdot \ell^2}{4n}$$



Approximation von **Dense** ℓ -Subgraph

- Die Approximationsgüte wird umso besser, je größer ℓ im Vergleich zu n ist.

- Es gibt andere Approximationsverfahren mit Güte $\mathcal{O}(\frac{n}{\ell})$, z.B. durch rekursives Löschen von Knoten mit kleinstem Grad.

Parametrisierte Dichte

- Schwelle für Dichte (density threshold) $\gamma: \mathbb{N} \rightarrow \mathbb{Q}_+$
 - ▶ γ soll in Polynomialzeit berechenbar sein und
 - ▶ $\forall \ell \in \mathbb{N}: \gamma(\ell) \leq \ell - 1$ soll gelten
- Knotenteilmenge U heißt genau dann **γ -dicht**, wenn

$$\overline{\text{deg}}(G[U]) \geq \gamma(|U|)$$

Problem

Problem: γ -dense **Subgraph**

Eingabe: Graph G , Parameter $\ell \in \mathbb{N}$

Frage: Existiert eine γ -dichte Knotenmenge der Kardinalität ℓ in G ?

Parametrisierte Dichte

- $\gamma(\ell) = \ell - 1$ entspricht **Clique**-Problem und ist \mathcal{NP} -vollständig
- $\gamma(\ell) = 0$ ist trivial, denn jede Menge U bestehend aus ℓ Knoten ist eine Lösung, die Antwort ist also immer "ja", wenn $\ell \leq |V|$
- Welche Funktionen $\gamma(\ell)$ erlauben noch Lösbarkeit in Polynomialzeit?

Satz

Sei γ eine Schwellwertfunktion für die Dichte.

- 1 Falls $\gamma = 2 + \mathcal{O}\left(\frac{1}{\ell}\right)$, dann ist γ -dense **Subgraph** in Polynomialzeit lösbar.
- 2 Falls $\gamma = 2 + \Omega\left(\frac{1}{\ell^{1-\varepsilon}}\right)$ für ein $\varepsilon > 0$, dann ist γ -dense **Subgraph** \mathcal{NP} -vollständig.

Parametrisierte Dichte

- Einen Teilgraph mit ℓ -Knoten und Durchschnittsgrad $\overline{\deg} \geq 2$ zu finden, geht also in Polynomialzeit.
- Aber einen Teilgraph mit ℓ -Knoten und Durchschnittsgrad $\overline{\deg} \geq 2 + \varepsilon$ (für $\varepsilon > 0$) zu finden, geht nicht in Polynomialzeit falls $\mathcal{P} \neq \mathcal{NP}$.
- k -Cores ließen sich in Linearzeit berechnen (sogar für alle k gleichzeitig)

⇒ riesiger **Komplexitätsunterschied** zwischen statistischer und struktureller Dichte!

Übersicht

- 1 Zusammenhang
 - Definitionen
 - Fundamentale Sätze

Zusammenhang in Graphen / Netzwerken

- beschäftigt sich mit der Stärke der Verbindung zwischen zwei Knoten in Bezug auf die Anzahl knoten- bzw. kantendisjunkter Wege
 - “Eine Kette ist nur so stark wie ihr schwächstes Glied.”
- ⇒ Wir suchen nach den schwächsten Elementen, die beim Entfernen die Verbindung zerstören.

Definition

Ein ungerichteter Graph heißt **zusammenhängend**, wenn es von jedem Knoten einen Pfad zu jedem anderen Knoten gibt.

Ein maximaler zusammenhängender induzierter Teilgraph wird als **Zusammenhangskomponente** bezeichnet.

Knoten-Zusammenhang

Definition

Ein ungerichteter Graph $G = (V, E)$ heißt **k -knotenzusammenhängend**, falls $|V| > k$ und für jede echte Knotenteilmenge $X \subset V$ mit $|X| < k$ der Graph $G - X$ zusammenhängend ist.

Der **Knotenzusammenhang** $\kappa(G)$ des Graphen G ist die größte natürliche Zahl k , für die G k -knotenzusammenhängend ist.

Bemerkungen:

- Jeder nicht-leere Graph ist 0-knotenzusammenhängend, da es keine Teilmenge X mit $|X| < 0$ gibt.
- Obwohl es wünschenswert wäre, dass die Bezeichnung “1-knotenzusammenhängend” gleichzusetzen ist mit der Bezeichnung “zusammenhängend”, wird üblicherweise der Graph bestehend aus nur einem einzelnen Knoten zwar als zusammenhängend, aber nicht 1-zusammenhängend bezeichnet.

Kanten-Zusammenhang und k -Komponenten

Definition

Ein ungerichteter Graph $G = (V, E)$ heißt **k -kantenzusammenhängend**, falls $|V| \geq 2$ und für jede Kanteilmenge $Y \subseteq E$ mit $|Y| < k$ der Graph $G - Y$ zusammenhängend ist.

Der **Kantenzusammenhang** $\lambda(G)$ des Graphen G ist die größte natürliche Zahl k , für die G k -kantenzusammenhängend ist.

Der Kantenzusammenhang eines unzusammenhängenden Graphen sowie des Graphen bestehend aus einem einzelnen Knoten ist 0.

Definition

Die maximalen k -fach knoten-/kanten-zusammenhängenden Teilgraphen werden als **k -Knoten-/Kanten-Zusammenhangskomponenten** bezeichnet.

Zusammenhang in gerichteten Graphen

Definition

Ein gerichteter Graph ist **stark zusammenhängend**, wenn es für jeden Knoten einen gerichteten Pfad zu jedem anderen Knoten gibt.

Ein maximaler stark zusammenhängender induzierter Teilgraph wird als **starke Zusammenhangskomponente** bezeichnet.

Knoten- und Kantenzusammenhang können auf gerichtete Graphen übertragen werden, indem man in der jeweiligen Definition fordert, dass $G - X$ bzw. $G - Y$ *stark zusammenhängend* ist.

Separatoren

Definition

Sei $G = (V, E)$ ein ungerichteter Graph.

Eine Knotenteilmenge $C \subset V$ heißt **Knoten-Separator**, wenn die Anzahl der Zusammenhangskomponenten in $G - C$ größer als in G ist.

Falls zwei Knoten s und t zwar in G in der gleichen Zusammenhangskomponente sind, aber nicht in $G - C$, dann bezeichnet man C als **s - t -Knoten-Separator**.

Kanten-Separatoren und **s - t -Kanten-Separatoren** sind analog definiert.

s - t -Separatoren können auch auf gerichtete Graphen übertragen werden: eine Knoten- bzw. Kantenmenge ist dann ein s - t -Separator, wenn es keinen gerichteten Pfad mehr von s nach t gibt, nachdem die Menge aus dem Graph entfernt wurde.

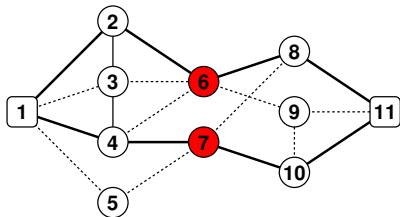
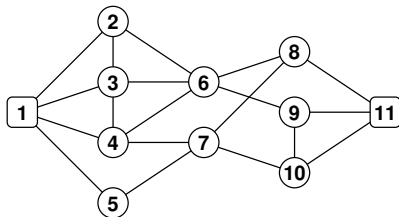
Disjunkte Pfade

Definition

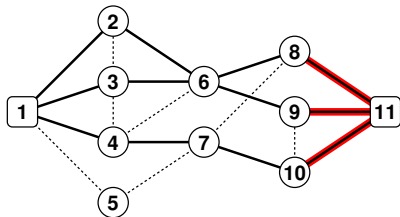
Zwei (gerichtete oder ungerichtete) Pfade von s nach t werden als **knotendisjunkte s - t -Pfade** bezeichnet, wenn sie keinen Knoten außer s und t gemeinsam haben.

Zwei Pfade werden als **kantendisjunkte Pfade** bezeichnet, wenn sie keine Kante gemeinsam haben.

Disjunkte s - t -Pfade



2 knotendisjunkte 1-11-Pfade



3 kantendisjunkte 1-11-Pfade

Lokaler Zusammenhang

Definition

Für zwei Knoten s und t eines Graphen G ist der **lokale (Knoten-)Zusammenhang** $\kappa_G(s, t)$ definiert als die minimale Anzahl von Knoten, die entfernt werden müssen, damit es keinen Weg mehr von s nach t gibt.

Für den Fall, dass zwischen s und t eine Kante existiert, können sie nicht durch das Löschen von Knoten separiert werden. Deshalb wird der lokale (Knoten-)Zusammenhang in diesem Fall $\kappa_G(s, t) = n - 1$ definiert. (Anderenfalls wäre höchstens $\kappa_G(s, t) = n - 2$ möglich.)

Der **lokale Kanten-Zusammenhang** zweier Knoten s und t ist entsprechend definiert als die minimale Anzahl von Kanten, die entfernt werden müssen, damit es keinen Weg mehr von s nach t gibt.

Lokaler Zusammenhang

Hinweis:

Für ungerichtete Graphen gilt $\kappa_G(s, t) = \kappa_G(t, s)$ und $\lambda_G(s, t) = \lambda_G(t, s)$, was für gerichtete Graphen im Allgemeinen nicht gilt.

Zweifachzusammenhang

Definition

Ein **Artikulationsknoten** ist ein Knoten, der beim Entfernen aus dem Graphen die Anzahl der Zusammenhangskomponenten erhöht.

Eine **Brücke** ist eine Kante, die beim Entfernen aus dem Graphen die Anzahl der Zusammenhangskomponenten erhöht.

Eine **Zweifachzusammenhangskomponente** ist ein maximaler 2-fach (knoten-)zusammenhängender Teilgraph.

Ein **Block** ist ein maximaler zusammenhängender Teilgraph, der keinen Artikulationsknoten enthält, d.h. die Menge aller Blocks eines Graphen besteht aus den isolierten Knoten, den Brücken, sowie den Zweifachzusammenhangskomponenten.

Block-Graph und CutPoint-Graph

Definition

Der **Block-Graph** $B(G)$ eines Graphen G hat jeweils einen Knoten für jeden Block von G (außer für isolierte Knoten), wobei zwei Knoten des Block-Graphen adjazent sind, wenn die entsprechenden Blöcke in G einen (Artikulations-)Knoten gemeinsam haben.

Der **CutPoint-Graph** $C(G)$ eines Graphen G hat jeweils einen Knoten für jeden Artikulationsknoten von G , wobei zwei Knoten des CutPoint-Graphen adjazent sind, wenn die entsprechenden Artikulationsknoten in G zum gleichen Block gehören.

Satz (Harary)

Für jeden Graphen gilt:

$$B(B(G)) = C(G) \quad \text{und} \quad B(C(G)) = C(B(G))$$

Block-CutPoint-Graph

Definition

Der **Block-CutPoint-Graph** eines Graphen G ist der bipartite Graph, dessen Knotenmenge aus je einem Knoten für jeden Artikulationsknoten von G und je einem Knoten für jeden Block von G besteht, wobei ein CutVertex-Knoten mit einem Block-Knoten genau dann durch eine Kante verbunden ist, wenn der Artikulationsknoten zu dem entsprechenden Block gehört.

Satz (Harary & Prins)

*Der Block-CutPoint-Graph eines zusammenhängenden Graphen ist ein **Baum**.*