Fortgeschrittene Netzwerk- und Graph-Algorithmen

Dr. Hanjo Täubig

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

Wintersemester 2009/10

Übersicht

Graphtraversierung

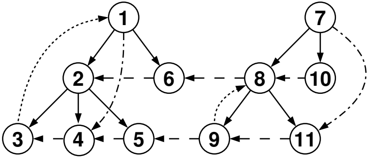
- Tiefensuche (depth first search, dfs)
  - Topologische Sortierung (bei DAGs)
  - Zweifachzusammenhangskomponenten
  - Starke Zusammenhangskomponenten
- Breitensuche (breadth first search, bfs)
  - schichtenweise Traversierung in aufsteigendem Abstand vom Ursprungsknoten

Traversierung / Kantentypen

DFS und BFS erzeugen Bäume bzw. Wälder und teilen die Kanten in folgende Klassen:

- Baumkanten (tree edges)
- Vorwärtskanten (forward edges)
- Rückwärtskanten (backward edges)
- Querkanten (cross edges)

Beispiel:



Artikulationsknoten und Blöcke

Definition

Ein Knoten v eines Graphen G heißt Artikulationsknoten (engl. *cut-vertex*), wenn sich die Anzahl der Zusammenhangskomponenten von G durch das Entfernen von v erhöht.

Definition

Die Zweifachzusammenhangskomponenten eines Graphen sind die maximalen Teilgraphen, die 2-fach zusammenhängend sind.

Ein Block ist ein maximaler zusammenhängender Teilgraph, der keinen Artikulationsknoten enthält. D.h. die Menge der Blöcke besteht aus den Zweifachzusammenhangskomponenten, den Brücken (engl. *cut edges*), sowie den isolierten Knoten.

Blöcke und DFS

Modifizierte DFS nach R. E. Tarjan:

- `num`[v]: DFS-Nummer von v
- `low`[v]: minimale Nummer `num`[w] eines Knotens w, der von v aus über beliebig viele (≥ 0) Baumkanten abwärts, evt. gefolgt von einer einzigen Rückwärtskante erreicht werden kann

- `low`[v]: Minimum von
  - `num`[v]
  - `low`[w], wobei w ein Kind von v im DFS-Baum ist
  - `num`[w], wobei $\{v, w\}$ eine Rückwärtskante ist

Blöcke und DFS

Lemma

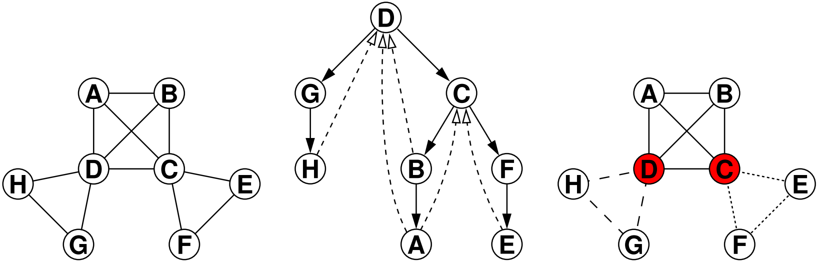
Sei $G = (V, E)$ ein ungerichteter, zusammenhängender Graph und T ein DFS-Baum in G.

Ein Knoten $a \in V$ ist genau dann ein Artikulationsknoten, wenn

- *a die Wurzel von T ist und mindestens 2 Kinder hat, oder*
- *a nicht die Wurzel von T ist und es ein Kind b von a mit `low[b]`$\geq$`num[a]` gibt.*

Blöcke und DFS

Die Kanten werden auf einem Stack gesammelt und nach der Erkennung eines Artikulationsknotens wird der gesamte Block abgepflückt.



Starke Zhk. und DFS

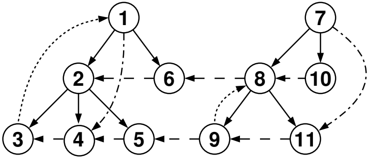
Modifizierte DFS nach R. E. Tarjan:

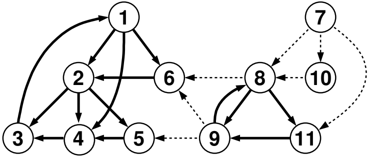
- `num[`v`]`: DFS-Nummer von v
- `low[`v`]`: minimale Nummer `num[`w`]` eines Knotens w, der von v aus über beliebig viele (≥ 0) Baumkanten abwärts, evt. gefolgt von einer einzigen Rückwärtskante oder einer Querkante zu einer ZHK, deren Wurzel echter Vorfahre von v ist, erreicht werden kann

- `low[`v`]`: Minimum von
  - `num[`v`]`
  - `low[`w`]`, wobei w ein Kind von v im DFS-Baum ist
  - `num[`w`]`, wobei $\{v, w\}$ eine Rückwärtskante ist
  - `num[`w`]`, wobei $\{v, w\}$ eine Querkante ist und die Wurzel der starken Zusammenhangskomponente von w ist Vorfahre von v

Starke Zusammenhangskomponenten

- Ein Knoten v ist genau dann Wurzel einer starken Zusammenhangskomponente, wenn `num[`v`]`=`low[`v`]`.





Übersicht

1. Grundlagen
2. Zentralitätsindizes
   - Beispiele
   - Grad- und Distanz-basierte Zentralitäten

Zentralitätsindizes

- Manche Knoten / Kanten in Netzwerken sind wichtiger, zentraler oder einflussreicher als andere.
- Zentralitätsmaße bzw. -indizes (kurz Zentralitäten) versuchen, diese Eigenschaften durch skalare Werte zu quantifizieren.
- Es gibt jedoch kein universelles Zentralitätsmaß, das jeder Anwendung gerecht wird.
  Das 'richtige' Maß hängt vom Kontext der Anwendung ab.

Beispiel: Leader election (1)

Bsp.: Wahl eines Klassensprechers

- Knoten entsprechen Personen
- Kante von Knoten A nach B, wenn Person A für Person B stimmt

- Person ist zentraler, je höher die Anzahl der erhaltenen Stimmen ist
  $\Rightarrow$ in-degree centrality

Beispiel: Leader election (2)

Bsp.: Wahl eines Klassensprechers

- Knoten entsprechen Personen
- Kante von Knoten A nach B, wenn Person A Person B überzeugt hat, für seinen/ihren Favoriten zustimmen
- Einfluss-Netzwerk

- Person ist zentraler, je mehr diese Person gebraucht wird, um die Meinung anderer zu transportieren
  $\Rightarrow$ betweenness centrality

Beispiel: Leader election (3)

Bsp.: Wahl eines Klassensprechers

- Knoten entsprechen Personen
- Kante von Knoten A nach B, wenn Person A mit Person B befreundet ist

- Person ist zentraler, je mehr Freunde diese Person hat und je zentraler diese Freunde sind
$\Rightarrow$ feedback centrality

Kantenzentralität

Genau wie die Zentralität von Knoten kann man auch die Wichtigkeit von Kanten betrachten
Beispiel: Internet

- Backbone: Verbindungen zwischen den Kontinenten gibt es wenige und sie müssen eine große Kapazität haben
- Kantenzentralität
  - Beteiligung einer Kante an kürzesten Wegen usw.
    $\Rightarrow$ betweenness edge centrality
  - Veränderung von Netzwerk-Parametern durch Löschen der Kante
    $\Rightarrow$ (edge) vitality (Bsp. flow betweenness vitality)

Definition: Zentralitätsindex

Abgesehen von der Intuition

- Wichtigkeit,
- Prestige,
- Einfluss,
- Kontrolle,
- Unentbehrlichkeit

gibt es keine allgemeingültige (formale) Definition von Zentralität.

Mindestanforderung:
Das Maß darf nur von der Struktur des Graphen abhängen.

Graph-Isomorphie

Definition (Isomorphie)

Zwei Graphen $G_1 = (V_1, E_1)$ und $G_2 = (V_2, E_2)$ sind isomorph ($G_1 \simeq G_2$), falls es eine Bijektion $\phi : V_1 \to V_2$ gibt, so dass

$$(u, v) \in E_1 \quad \Leftrightarrow \quad (\phi(u), \phi(v)) \in E_2$$

Definition (Struktureller Index)

Sei $G = (V, E)$ ein gewichteter (gerichteter oder ungerichteter) graph und sei X die Knotenmenge (V) oder die Kantenmenge (E). Eine reell-wertige Funktion s heißt *struktureller Index* genau dann, wenn die folgende Bedingung erfüllt ist:

$$\forall x \in X : \quad G \simeq H \Rightarrow s_G(x) = s_H(\phi(x))$$

Grad-Zentralität

Grad-Zentralität (degree centrality):

$$c_D(v) = \deg(v)$$

- lokales Maß,
  hängt nur von der direkten Nachbarschaft eines Knotens ab

Exzentrizität

Anwendungsbeispiel:
Positionierung eines Hospitals oder einer Feuerwehrstation

Ziel: Minimierung der maximal notwendigen Anfahrtzeit

Exzentrizität eines Knotens $v \in V$:

$$e(v) = \max_{w \in V}\{d(v, w)\}$$

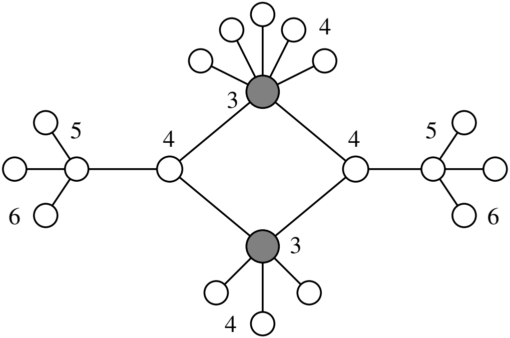
Zentralitätsmaß:

$$c_E(v) = \frac{1}{e(v)}$$

Optimaler Standort:

Knoten v mit minimalem Wert $e(v)$ (*Zentrum* von G)

Exzentrizität / Beispiel



(Quelle: Brandes/Erlebach (Eds.): Network Analysis)

Closeness

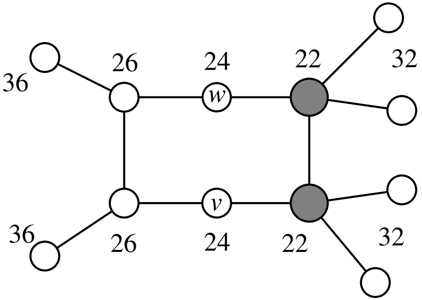
Anwendungsbeispiel:
Positionierung eines Einkaufszentrums
(minisum location / median / service facility location problem)

Ziel:
Minimierung der Summe der Entfernungen zu den anderen Knoten
(und damit auch der durchschnittlichen Entfernung)

Zentralitätsmaß:

$$c_C(v) = \frac{1}{\sum_{w \in V} d(v,w)}$$

Closeness / Beispiel



(Quelle: Brandes/Erlebach (Eds.): Network Analysis)

Graue Knoten sind am wichtigsten bezüglich Closeness,
v und *w* sind zentraler in Hinsicht auf Exzentrizität

Radiality

ähnlich zu Closeness

$$c_R(v) = \frac{\sum_{w \in V}(\Delta_G + 1 - d(v,w))}{n-1}$$

Δ_G: Durchmesser des Graphen
(größte Distanz zweier Knoten, nicht Länge des längsten Pfads!)