Algorithmische Methoden der Netzwerkanalyse

Marco Gaertler

Algorithmic Group
Faculty of Informatics
Universität Karlsruhe (TH)
Research University · founded 1825

25. November, 2008

- Zentralität
- Abstandszentralitäten

Strukturindex

Definition

Seien $\mathcal{K}$ eine unter Bildung von Zusammenhangskomponenten abgeschlossene Klasse nicht-isomorpher Multigraphen und $\mathbb{R}^*_{\geq 0}$ die Menge aller Vektoren über den nicht-negativen reellen Zahlen. Eine Funktion $s\colon \mathcal{K} \to \mathbb{R}^*_{\geq 0}$ heißt (Knoten- bzw. Kanten-)Strukturindex auf $\mathcal{K}$, falls:

- $s(G) \in \mathbb{R}^V_{\geq 0}$ bzw. $s(G) \in \mathbb{R}^E_{\geq 0}$ für alle $G = (V, E) \in \mathcal{K}$
- $s(G)_v = s(\alpha(G))_{\alpha(v)}$ bzw. $s(G)_e = s(\alpha(G))_{\alpha(e)}$ für alle $G = (V, E) \in \mathcal{K}$, $v \in V$, $e \in E$ und α Automorphismus von G
- $s(G)_v \cdot s(C)_w = s(C)_v \cdot s(G)_w$ bzw. $s(G)_e \cdot s(C)_f = s(C)_e \cdot s(G)_f$ für alle $G = (V, E) \in \mathcal{K}$, $C = (V_C, E_C)$ Zhk. und $v, w \in V_C$ bzw. $e, f \in E_C$

Zentrlität

Definition

Ein Strukturindex c auf einer unter Hinzufügen von Kanten abgeschlossenen Klasse $\mathcal{K}$ von Multigraphen heißt *(schwacher) (Knoten-)Zentrlität(sindex)*, falls eine der drei folgendenen Bedingungen für alle $G = (V, E) \in \mathcal{K}$ und $v, w \in V$ gilt ($G' := G + (v, w)$):

- $\bullet \to \circ$: $c(G)_v \geq c(G)_x \Longrightarrow c(G')_v \geq c(G')_x$ für alle $x \in V$
- $\circ \to \bullet$: $c(G)_w \geq c(G)_x \Longrightarrow c(G')_w \geq c(G')_x$ für alle $x \in V$
- $\bullet \to \bullet$: $c(G)_v + c(G)_w \geq c(G)_x \Longrightarrow$ $c(G')_v + c(G')_w \geq c(G')_x$ für alle $x \in V$

Normalisierung

Definition

Ein Knoten- bzw. Kanten-Strukturindex s auf $\mathcal{K}$ heißt *normiert*, falls

$$\sum_{v \in V} s(G)_v = 1 \qquad \text{bzw.} \qquad \sum_{e \in E} s(G)_v = 1$$

für alle $G = (V, E) \in \mathcal{K}$.

Gradzentralität

Beispiel

- Eingangsgrad ist eine $\circ \rightarrow \bullet$-Zentralität
- Ausgangsgrad ist eine $\bullet \rightarrow \circ$-Zentralität
- Knotengrad ist eine $\bullet \rightarrow \bullet$-Zentralität

offene Frage

Folgt aus $c \in \circ \rightarrow \bullet(\mathcal{K}) \cap \bullet \rightarrow \circ(\mathcal{K}) \cap \bullet \rightarrow \bullet(\mathcal{K})$, dass c monoton im Knotengrad ist?

Gradzentralität

Beispiel

- Eingangsgrad ist eine $\circ \to \bullet$-Zentralität
- Ausgangsgrad ist eine $\bullet \to \circ$-Zentralität
- Knotengrad ist eine $\bullet \to \bullet$-Zentralität

offene Frage

Folgt aus $c \in \circ \to \bullet(\mathcal{K}) \cap \bullet \to \circ(\mathcal{K}) \cap \bullet \to \bullet(\mathcal{K})$, dass c monoton im Knotengrad ist?

- Zentralität
- Abstandszentralitäten

Abstand

Definition

Sei $G = (V, E)$ ein Multigraph. Gibt es für zwei Knoten $s, t \in V$ einen Weg von s nach t, so heißt die kürzeste Länge eines (s, t)-Weges *Abstand* (*Distanz*), $d_G(s, t)$, von s nach t. Gibt es keinen (s, t)-Weg, so setzen wir $d_G(s, t) = \infty$.

Im folgenden sei $\mathcal{S}$ die Klasse aller stark zusammenhängender Multigraphen und $\mathcal{G}$ die Klasse alle Multigraphen.

Abstand

Definition

Sei $G = (V, E)$ ein Multigraph. Gibt es für zwei Knoten $s, t \in V$ einen Weg von s nach t, so heißt die kürzeste Länge eines (s, t)-Weges *Abstand* (*Distanz*), $d_G(s, t)$, von s nach t. Gibt es keinen (s, t)-Weg, so setzen wir $d_G(s, t) = \infty$.

Im folgenden sei $\mathcal{S}$ die Klasse aller stark zusammenhängender Multigraphen und $\mathcal{G}$ die Klasse alle Multigraphen.

Exzentrizität, Durchmesser, Radius

Definition

Für einen Multigraphen $G = (V, E)$ definieren wir:

$$\begin{aligned} e_G(v) &= \max\{d_G(v,w), d_G(w,v) \ : \ w \in V\} \\ &\quad \textit{quad} \text{ die } \textit{Exzentrizität} \text{ von } v \\ \mathrm{rad}(G) &= \min\{e_G(v) \ : \ v \in V\} \qquad \text{den } \textit{Radius} \text{ von } G \\ \mathrm{diam}(G) &= \max\{e_G(v) \ : \ v \in V\} \qquad \text{den } \textit{Durchmesser} \text{ von } G \ . \end{aligned}$$

Closeness

Definition

Die *Closeness*-Zentralität c_C ist definiert durch

$$c_C(G)_v := \frac{1}{\sum_{t \in V} d_G(v,t)}$$

für alle $G = (V, E) \in \mathcal{S}$, wobei $1/0 := 1$ gelte.

Closeness

$$c_C \overset{?}{\in} \bullet \longrightarrow \circ(\mathcal{S})$$

$$c_C \overset{?}{\in} \circ \longrightarrow \bullet(\mathcal{S})$$

$$c_C \overset{?}{\in} \bullet \longrightarrow \bullet(\mathcal{S})$$

Closeness

$$c_C \in \bullet \to \circ(\mathcal{S})$$

$$c_C \notin \circ \to \bullet(\mathcal{S})$$

$$c_C \notin \bullet \to \bullet(\mathcal{S})$$

Betweenness

Definition

Die *Betweenness*-Zentralität c_B ist definiert durch

$$c_C(G)_v := \sum_{s,t \in V} \frac{\sigma_G(s,t|v)}{\sigma_G(s,t)}$$

für alle $G = (V, E) \in \mathcal{G}$. Dabei bezeichne $\sigma_G(s,t)$ die Anzahl der kürzesten (s,t)-Wege und $\sigma_G(s,t|v)$ die Anzahl der kürzestens (s,t)-Wege, die v als inneren Knoten besitzen. Weiter gelte $0/0 := 0$.

Betweenness

$$c_B \stackrel{?}{\in} \bullet \longrightarrow \circ(\mathcal{G})$$

$$c_B \stackrel{?}{\in} \circ \longrightarrow \bullet(\mathcal{G})$$

$$c_B \stackrel{?}{\in} \bullet \longrightarrow \bullet(\mathcal{G})$$

Betweenness

$$c_B \notin \bullet \longrightarrow \circ(\mathcal{G})$$

$$c_B \notin \circ \longrightarrow \bullet(\mathcal{G})$$

$$c_B \in \bullet \longrightarrow \bullet(\mathcal{G})$$

Berechnung

Frage:

- Wie lassen sich Closeness und Betweenness berechnen?
- Wie schnell lassen sich Closeness und Betweenness berechnen?

Distanzmatrix

Lemma

Sei $G = (V, E)$ ein Multigraph, $A = A(G)$ seine Adjazenzmatrix und $A^k = (a_{s,t}^{(k)})_{s,t \in V}$ deren k-te Potenz. Für zwei Knoten $s, t \in V$ ist $a_{s,t}^{(k)}$ gerade die Anzahl aller gerichteten Kantenfolgen von s nach t der Länge k.

Breitensuche

- *Baumkante*, falls w nicht markiert ist
- *Rückwärtskante*, falls w markiert ist und $\mathrm{BFS}(w) < \mathrm{BFS}(v)$ gilt
- *Querkante*, falls w markiert ist und $\mathrm{BFS}(w) = \mathrm{BFS}(v)$ gilt
- *Vorwärtskante*, falls w markiert ist und $\mathrm{BFS}(w) > \mathrm{BFS}(v)$ gilt

Breitensuche

- *Baumkante*, falls w nicht markiert ist
- *Rückwärtskante*, falls w markiert ist und $\mathrm{BFS}(w) < \mathrm{BFS}(v)$ gilt
- *Querkante*, falls w markiert ist und $\mathrm{BFS}(w) = \mathrm{BFS}(v)$ gilt
- *Vorwärtskante*, falls w markiert ist und $\mathrm{BFS}(w) > \mathrm{BFS}(v)$ gilt