Universität Konstanz
Fachbereich Informatik & Informationswissenschaft
Prof. Dr. Ulrik Brandes / Dr. Jürgen Lerner / Bobo Nick

7. Übungsblatt

Ausgabe: 07.6.2010 **Abgabe:** 14.6.2010, 12 Uhr
Die Bearbeitung in Zweiergruppen ist ausdrücklich erwünscht.

Aufgabe 1: Maximum Likelihood Estimation — 5 Punkte

Durch $c : V \rightarrow \{A, B\}$ sei eine Aufteilung der Knoten in zwei diskunkte Klassen gegeben, $V = A \uplus B$. Damit lässt sich analog zu $\mathcal{G}(n, p)$ ein Zufallsgraphen-Modell $\mathcal{G}(n, c, p_1, p_2)$ definieren, wobei p_1 die Wahrscheinlichkeit für eine Kante zwischen zwei Knoten aus der gleichen Klasse und p_2 die Wahrscheinlichkeit für eine Kante zwischen zwei Knoten aus unterschiedlichen Klassen definiere.

Berechnen Sie für einen gegeben Graphen G_{obs} aus $\mathcal{G}(n, c, p_1, p_2)$ die wahrscheinlichsten zugrundeliegenden Parameter p_1 und p_2. Inwiefern ließe sich das Verhältnis von p_1 und p_2 für einen Test auf Homophilie im Netzwerk verwenden?

Aufgabe 2: Preferential Attachment, Erzeugungsalgorithmus — 4 Punkte

Modifizieren Sie den in der Vorlesung präsentierten Algorithmus zur Erzeugung von Multigraphen mit Bevorzugung so, dass die Auswahl des Zielknoten mit einer Wahrscheinlichkeit proportional zu einer beliebigen Gewichtung von Eingangs- und Ausgangsgrad vorgenommen werden kann.

Aufgabe 3: Preferential Attachment, Randverteilungen — 6 Punkte

Beweisen Sie, dass die Definition von Kantenwahrscheinlichkeiten im Modell für Multigraphen mit Bevorzugung nicht ausreicht, um *gleichzeitig* alle m Nachbarn eines neuen Knotens zu wählen, d.h. die Wahrscheinlichkeiten für bestimmte Nachbarschaften eines Knotens sind nicht wohldefiniert.

Tipp: Betrachten Sie die sechs möglichen Nachbarschaften im Fall, dass für $m = 2$ zum Graphen C_4 (dem einfachen Kreis mit 4 Knoten) ein weiterer Knoten v hinzukommt.

[Bitte wenden]

Aufgabe 4: ERGM, Kantenabhängigkeit **5 Punkte**

Sei $\mathcal{G}$ die Menge aller schleifenfreien, ungerichteten Graphen mit $n = 3$ Knoten. Aus der ERGM-Familie betrachten wir das Zufallsgraphenmodell $(\mathcal{G}, P)$, welches lediglich die Statistik $g_1 = s_2$ (Anzahl der 2-Sterne) mit zugehörigem Parameter $\theta_1 = \ln 2$ verwendet. Beweisen Sie (wie in der Vorlesung im Zusammenhang mit der Dreieck-Statistik bereits geschehen), dass durch die Verwendung der 2-Sterne-Statisik in diesem $(\mathcal{G}, P)$ jede Kante von jeder anderen abhängig ist.