Rolf Wanka
Alexander Raß

Erlangen, 23. Mai 2019

Übungen zur Vorlesung

Organic Computing

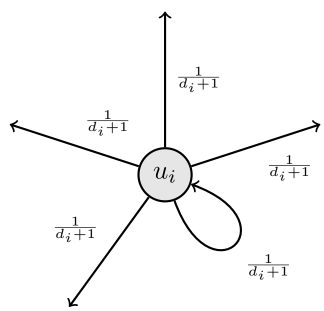
SS 2019

Blatt 4

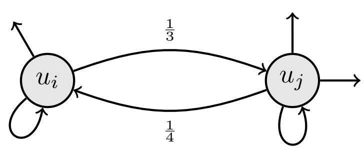
Am 31. Mai/5. Juni werden wieder Tafelübungen sein.

AUFGABE 13:

Sei $G = (V, E)$ ein ungerichteter Graph mit $n = |V|$ Knoten. Der **Grad** eines Knotens ist die Anzahl der Kanten, die diesen Knoten mit anderen Knoten verbindet. Für jeden Knoten u_i mit Grad d_i seien an den Kanten Wahrscheinlichkeiten angegeben, mit denen der Knoten verlassen wird. Alle Kanten und die Möglichkeit, den Knoten nicht zu verlassen, haben die gleiche Wahrscheinlichkeit. Es ergibt sich also folgendes Bild:



Für jede ungerichtete Kante gibt es also zwei Werte, einen für die Hinrichtung (den Knoten verlassend) und einen für die Rückrichtung (den Knoten betretend). Die folgende Grafik zeigt die Wahrscheinlichkeiten für eine Kante, die zwei Knoten mit unterschiedlichem Grad verbindet:



Einige der Knoten in G seien rot gefärbt, alle anderen blau.

(a) Modellieren sie die Wahrscheinlichkeiten, nach s Schritten auf den jeweiligen Knoten zu stehen. Geben Sie dazu eine Matrix M an, so dass die Wahrscheinlichkeiten durch Multiplikation mit M berechnet werden können.

(b) Zeigen Sie für $\vec{\pi} = \left(\frac{d_1+1}{2|E|+|V|}, \frac{d_2+1}{2|E|+|V|}, \ldots, \frac{d_n+1}{2|E|+|V|}\right)$: $\vec{\pi} \cdot M = \vec{\pi}$

(c) Betrachten Sie nun einen *regulären* Graphen, für dessen Kanten Wahrscheinlichkeiten, wie oben beschrieben, angegeben sind. Ein Graph ist **regulär**, wenn alle Knoten den gleichen Grad d haben. Jeder Knoten ist also mit genau d anderen Knoten durch eine Kante verbunden. Die Wahrscheinlichkeit für alle Kanten und jede Richtung ist also immer gleich:

$$p = \frac{1}{d+1}$$

Starten Sie bei Knoten u_1. In jedem Schritt folgen Sie einer der Kanten zu einem Folgeknoten u_{i+1} oder bleiben beim aktuellen Knoten u_i.

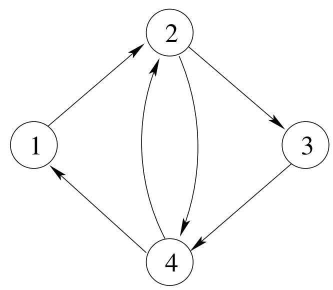
Dieses Vorgehen wiederholen Sie nun und merken sich, wie viele „rote“ und „blaue“ Knoten Sie schon besucht haben. Können Sie anhand dieser Summen die Anzahl der roten Knoten in dem Graphen G angeben? Falls ja, wie?

(d) Erinnern Sie sich an das erste Übungsblatt: Was hat das Vorgehen von (c) mit Eigenwerten zu tun?

AUFGABE 14:

Ziel des *PageRank-Algorithmus* ist es, die Relevanz von Internetseiten zu bewerten. Ähnlich wie in der vorherigen Aufgabe wird dazu ein ziellos durch das Internet surfender Benutzer simuliert, der auf jeder Seite mit gleicher Wahrscheinlichkeit einem dort platzierten Link folgt. Je häufiger der Benutzer auf einer bestimmten Seite vorbeikommt, als desto wichtiger wird sie eingestuft, und sie bekommt einen höheren PageRank.
In dieser Aufgabe sollen die Grundlagen des PageRank-Algorithmus erarbeitet werden.

(a) Betrachten Sie folgenden Netzgraphen mit den Internetseiten 1, 2, 3 und 4:

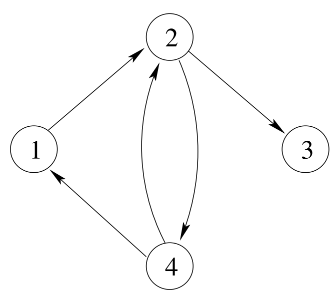


Der Vektor $\vec{\pi}^{(t)} = (\pi_1^{(t)}, \pi_2^{(t)}, \pi_3^{(t)}, \pi_4^{(t)})$ gebe die Wahrscheinlichkeiten an, dass sich der Benutzer zum Zeitpunkt t auf der jeweiligen Seite befindet. Angenommen, der Benutzer startet auf Seite 1, d.h. $\vec{\pi}^{(0)} = (1, 0, 0, 0)$. Wo befindet er sich nach zwei Schritten?

(b) Modellieren Sie das System mittels einer Matrix A, so dass gilt: $\vec{\pi}^{(t+1)} = \vec{\pi}^{(t)} \cdot A$

(c) Zu welchem Vektor $\vec{\pi}$ konvergiert das System? Wie sieht also die Rangfolge der Seiten aus?

(d) Im Internetgraph kann folgende Situation auftreten:



Was ist daran problematisch? Wie könnte dieses Problem gelöst werden? (Überlegen Sie sich, was Sie machen, wenn Sie auf einer Seite keinen oder keinen interessanten Link finden.) Berechnen Sie mittels Ihrer Methode eine Rangfolge für die Seiten 1, 2, 3 und 4.

Schlussbemerkung: Da für große Graphen, wie z.B. das Internet, der Vektor $\vec{\pi}$ nicht exakt berechnet werden kann, wird der PageRank iterativ bestimmt, unter Verwendung von $\vec{\pi}^{(t+1)} = \vec{\pi}^{(t)} \cdot A$ und erweitert um eine Strategie, die das Problem aus Aufgabenteil (d) vermeidet.