

## Peer-to-Peer-Systeme WS 2015/16

### Übungsblatt 2

Besprechung am 12. November 2015

#### Aufgabe 6

Stellen wir uns ein hierarchisches Peer-to-Peer-Overlay mit *drei* Hierarchiestufen vor: Es gibt „normale“ Peers, Superpeers und Super-Duper-Peers.

- (a) Nehmen wir an, dass ein Superpeer 100 normale Peers betreut und ein Super-Duper-Peer für 100 Superpeers zuständig ist. Wie viele Super-Duper-Peers wären in einem Netzwerk mit 999 000 Peers notwendig?
- (b) Überlege, wie in einem solchen Netzwerk die Bearbeitung von Suchanfragen erfolgen könnte. Welche Informationen sollte ein Superpeer speichern? Was sollte ein Super-Duper-Peer wissen? Wie würde eine Suche sinnvollerweise ablaufen?

#### Aufgabe 7

Betrachte ein hypothetisches hierarchisches Overlay ähnlich Gnutella 0.6. Wir nehmen an, dass darin jeder Superpeer mit genau fünf anderen Superpeers verbunden ist.

Nun setzt einer der Superpeers eine Suchanfrage im Superpeer-Overlay ab. Dabei setzt er Dynamic Querying ein. Ein erster Probe-Suchlauf über 10 Superpeers hat  $x$  Treffer geliefert ( $1 \leq x < 150$ ).

Mache einen Vorschlag für die Wahl der TTL bei einer anschließenden „vollen“ Suche (abhängig von  $x$ ), wenn diese neue Suche 150 Treffer erzielen soll.

## Aufgabe 8

Schreibe ein Programm (in einer Programmiersprache deiner Wahl), das für vorgegebene Parameter  $n, p$  Gilbert-Zufallsgraphen erzeugt.

Das Programm soll dann die durchschnittliche Pfadlänge des erzeugten Graphen bestimmen und ausgeben. Damit für die Bestimmung der durchschnittlichen Pfadlänge nicht alle Knotenpaare untersucht werden müssen, genügt es, wenn das Programm 100 Knotenpaare zufällig auswählt und diese Paare als Stichprobe für die Berechnung der Durchschnittspfadlänge verwendet.

Außerdem soll für die Graphen der sogenannte Clustering-Koeffizient bestimmt werden. Der Clustering-Koeffizient eines Graphen ist ein Maß für die Lokalität des Graphen, also wie „dicht“ die Umgebung eines Knotens typischerweise vernetzt ist. Um den Clustering-Koeffizienten eines Graphen zu bestimmen, berechnet man zunächst den Clustering-Koeffizienten seiner einzelner Knoten. Dafür betrachtet man alle Paare von Nachbarn eines gegebenen Knotens. Für jedes dieser Paare bestimmt man, ob die beiden Knoten selbst auch wieder durch eine Kante verbunden sind. Angenommen, der Knoten hat  $M$  mögliche Paare von Nachbarn, von denen  $m$  durch eine Kante verbunden sind. Dann ist der Clustering-Koeffizient des Knotens  $m/M$ . Der Clustering-Koeffizient eines Graphen ist der durchschnittliche Clustering-Koeffizient seiner Knoten; Knoten mit Grad kleiner 2 bleiben unberücksichtigt.

Nutze dein Programm und beantworte die folgenden Fragen:

- (a) Eine bekannte Aussage über Zufallsgraphen ist, dass die durchschnittlichen Pfadlängen nur sehr langsam mit der Knotenzahl wachsen. Die Theorie sagt: Bei konstantem durchschnittlichem Knotengrad wächst die durchschnittliche Pfadlänge logarithmisch mit der Knotenzahl. Wie musst Du  $p$  abhängig von  $n$  wählen, damit der Knotengrad bei steigender Größe des Graphen konstant bleibt?
- (b) Erstelle ein Diagramm, das die Abhängigkeit der durchschnittlichen Pfadlänge von der Knotenzahl zeigt und dabei den durchschnittlichen Knotengrad konstant hält. Wähle dafür geeignete Parameter(bereiche). Kannst Du in dem Diagramm das von der Theorie vorhergesagte logarithmische Wachstum nachweisen? (Hinweis: Für die visuelle Darstellung eines logarithmischen Wachstums lohnt es sich, über geeignete Achsenskalierungen im Diagramm nachzudenken!)
- (c) Untersuche mithilfe geeigneter Darstellung(en), wie der Clustering-Koeffizient von  $n$  und  $p$  abhängt. Kannst Du das beobachtete Verhalten erklären?