Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Distributed Computing

FS 2012

Prof. R. Wattenhofer / Philipp Brandes

Principles of Distributed Computing
Exercise 9

For this exercise, you may use the following inequalities, based on the *Chernoff bound*.[1]

Theorem 1 (Chernoff Bounds)
Let $X := \sum_{i=1}^{N} X_i$ be the sum of N independent $0-1$ random variables X_i.

Bound 1:

$$Pr\left[|X - \mathbb{E}[X]| \in O\left(\log n + \sqrt{\mathbb{E}[X]\log n}\right)\right] \quad > \quad 1 - \frac{1}{n^c} \quad \textit{for a selectable constant } c$$

Bound 2:

$$Pr[X \leq (1-\delta)\mathbb{E}[X]] \leq e^{-\mathbb{E}[x]\delta^2/2} \qquad \textit{for all } 0 < \delta \leq 1$$

1 Diameter of the Augmented Grid

Recall the network from the lecture where nodes were arranged in a grid and each node had an additional directed link to an uniformly and independently at random drawn node in the network (i.e., $\alpha = 0$). In the lecture, a proof of the fact that such a network has diameter $O(\log n)$ w.h.p. was sketched. We will now fill in the details.

a) Show that $O(n/\log n)$ many nodes are enough to guarantee with high probability that at least one of their random links connects to a given set of $\Omega(\log^2 n)$ nodes. Prove this (i) by direct calculation and (ii) using Chernoff's bound.

Hint: Use that $1 - p \leq e^{-p}$ for any p.

b) Suppose for some node set S we have that $|S| \in \Omega((\log n)^2) \cap o(n)$ and denote by H the set of nodes hit by their random links. Prove that H and together with its grid neighbors contains w.h.p. $(5 - o(1))|S|$ nodes!

Hint: Observe that *independently* of all previous random choices, each new link has at least a certain probability p of connecting to a node whose complete neighborhood has not been reached yet. Then use Chernoff's bound on the sum of $|S|$ many variables.

c) Infer from **b)** that starting from $\Omega(\log^2 n)$ nodes, with each hop the number of reached nodes w.h.p. more than doubles, as long as we have still $O(n/\log n)$ nodes (regardless of the constants in the O-notation).

Hint: Play with the constant c in the definition of w.h.p. and use the union bound.

d) Conclude that the diameter of the network is w.h.p. in $O(\log n)$.

[1] Chernoff-type and similar probability bounds are very powerful tools that allowed to design a plethora of randomized algorithms that *almost* guarantee success. Frequently this "almost" makes a huge difference in e.g., running time and/or approximation quality.

2 Greedy Routing in the Augmented Grid

Consider now $\alpha = 2$, i.e., the random link of node u connects it to node w with probability $d(u,w)^{-2}/\sum_{v \in V \setminus \{u\}} d(u,v)^{-2}$. In the lecture, we saw that with probability $\Omega(1/\log n)$, in each step we get to the next phase when we employ greedy routing. Hence, the expected number of steps is in $O(\log^2 n)$. Prove that the same bound on the number of steps holds w.h.p.!

3 Scale Free Networks

Different studies of the structures of social networks have reported that the degree distribution of the underlying connectivity graphs asymptotically follow a power law, i.e., the probability of a node in a social network to have degree k is given by:

$$Pr[k] = ck^{-\alpha} \quad \text{where } c \text{ is a normalization constant}$$

a) Is the diameter of two graphs with the same node-degree distribution equal (not necessarily power law graphs)?

b) Remember the the rumor game from the lecture: Two players choose a node on the graph, where they start their rumor. The player that is closer to a node in the graph can spread its rumor to the node. Winner is the player who can spread his rumor to more nodes. In a power law network, is it the optimal strategy to always choose the node with the highest degree?