Ph.D. course on Network Dynamics

Homework 3

To be discussed on Wednesday, November 30, 2011

Exercise 1. *Read Mitzenmacher's survey on power laws [1], available at http://www.eecs.harvard.edu/ michaelm/CS223/powerlaw.pdf.*

Exercise 2 (Stars in preferential attachment)**.** *This exercise was suggest by Zayide at the end of last class. A start graph is a tree in which all but one nodes are leaves (i.e., have degree* 1*). Consider the Albert-Barabasi preferential attachment model studied in class with* $m = 1$*. Prove that* $\mathcal{G}_t$ *is a star with probability* 2^{1-t}*, for all* $t \geq 1$*.*

Exercise 3. *Exercise 14.4 from [2] (this book, as well as all the other I ordered for the course can be borrowed from the Automatic Control Department's library).*

Exercise 4. *The following model does not involve a graph, but can be studied using the same mean-field method as for the Albert-Barabasi preferential attachment model.*

A famous surrealist author is known to compose text as follows. She starts with a random word. Suppose that t *words (not necessarily different) have already been written. The next word is chosen as follows:*

- with probability α*, it is a new word; - with probability* $1-\alpha$*, she chooses some* j *uniformly at random from the set of past instants* $\{1, \dots, t-1\}$ *and copies the* j*-th word that she has alerady written. Let* $n_i(t)$ *be the expected number of distinct words that appear exactly* i *times, after the first* t *words have been written.*

1. *Write down a recursion (in* t*) for the variables* $n_i(t)$*.*

2. *Assume (or, better, prove) that $n_i(t)/t$ converges to some $\beta_i \geq 0$ for all $i \geq 1$. Find equations that relate the β_i.*

3. *Show that β_i/β_{i+1} converges to 1 as i grows large, and that the β_i's correspond to a power law.*

Exercise 5 (Hoeffding-Azuma inequality). *Let $M_0, M_1, \ldots, M_t$ be a martingale such that*

$$|M_i - M_{i-1}| \leq c_i\,, \qquad \forall i = 1, \ldots, t\,. \tag{1}$$

1. *prove that, for $c \geq 0$*

$$e^x \leq \frac{\sinh(c)}{c} x + \cosh(c)\,, \qquad \forall x \in [-c, c]\,;$$

   (hint: use convexity)

2. *prove that, for $c \geq 0$*

$$\cosh(z) \leq \exp(z^2/2)\,, \qquad \forall z \in \mathbb{R}\,;$$

   (hint: use Taylor expansions)

3. *prove that, for all $\theta \geq 0$,*

$$\mathbb{E}[\exp(\theta(M_i - M_{i-1}))|M_0, \ldots, Mi - 1] \leq \exp(\theta^2 c_i^2/2)\,;$$

   (hint: use the martingale property, (1), and the previous two points)

4. *prove that, for all $\theta \geq 0$, and $\varepsilon \geq 0$,*

$$\mathbb{P}(M_t - M_0 \geq \varepsilon) \leq \exp(-\theta\varepsilon + \theta^2 \sum_{1 \leq i \leq t} c_i^2/2)\,;$$

5. *conclude that*

$$\mathbb{P}(M_t - M_0 \geq \varepsilon) \leq \exp(-\varepsilon^2/(2 \sum_{1 \leq i \leq t} c_i^2))\,, \qquad \forall \varepsilon \geq 0\,,$$

   which is the Hoeffding-Azuma inequality.

Exercise 6 (Hoeffding-Azuma inequality vs Chernoff bound)**.** *Observe that the Hoeffding-Azuma inequality applies to martingales with bounded increments, whereas the Chernoff bound applies to sequences of i.i.d. random variables, not necessarily with bounded support. Therefore, the ranges of applicability of the two are not comparable. Nevertheless, we can compare them in the following simple case.*

Let $X_1, X_2, \ldots, X_n$ *i.i.d. random variables with distribution Bernoulli*(p)*. Is the strongest upper bound on* $\mathbb{P}(\sum_{i=1}^n X_i \geq (p+\varepsilon)n)$ *provided by the Hoeffding-Azuma inequality or the Chernoff bound? Compare both also with Bernstein's inequality*

$$\mathbb{P}\left(\sum_{1\leq i\leq n} X_i \geq (p+\varepsilon)n\right) \leq \exp\left(-n\frac{\varepsilon^2}{2(1-p)(p+\varepsilon/3)}\right) .$$

References

[1] M. Mitzenmacher, *A brief history of generative models for power law and lognormal distributions*, Internet Mathematics **1** (2004), no. 2, 226–251.

[2] M. E. J. Newman, *Networks: an introduction*, Oxford University Press, 2010.