T-79.7003: Graphs and Networks		**Out: Nov. 8th, 2013**
	Homework 2	
Lecturer: Charalampos E. Tsourakakis		*In:Nov. 29th, 2013*

2.1 Concentration [15+10 Points]

(A) Let $X \sim NB(r,p)$ be distributed according to the negative binomial distribution with parameters $r \in \mathbb{N}$ and $p \in (0,1)$. Prove that

$$\boxed{\mathbf{Pr}\left[|\frac{r}{X} - p| \geq \epsilon p\right] \leq 2e^{-\frac{\epsilon^2 r}{3(1+\epsilon)}}.}$$

(B) Prove the following Chernoff bound. Let $X_1, \ldots, X_n$ be independent Poisson trials such that $\mathbf{Pr}[X_i = 1] = p_i, \mathbf{Pr}[X_i = 0] = 1 - p_i, i = 1, \ldots, n$. Let $X = X_1 + \ldots + X_n$, $\mu = \mathbb{E}[X]$. Then, for $0 < \epsilon \leq 1$,

$$\boxed{\mathbf{Pr}[X \geq (1+\epsilon)\mu] \leq e^{-\frac{\mu\epsilon^2}{3}}.}$$

2.2 Erdös-Rényi graphs [25 points]

Let $G = G(n,p), p = \frac{\log n + \omega}{n}$, $\omega \to \infty, \omega = o(\log n)$. Let $deg(x)$ be the degree of vertex x and $dist(x,y)$ be the shortest path distance between vertices x, y. Prove the following

$$\boxed{\mathbf{Pr}\left[\exists x, y \in [n]:\ \deg(x), \deg(y) \leq \log n/100 \text{ and } dist(x,y) \leq \frac{3\log n}{4\log\log n}\right] = o(1).}$$

2.3 How do search engines affect the Web? [50 points]

We consider a directed graph with growth at discrete time steps. The model has four parameters $\alpha, \beta, \delta_{in}, \delta_{out}$ Let $\{G_t\}_{t \geq 0}$ be the sequence of graphs generated according to the following rules.

- At time $t = 0$, let G_0 be a single vertex without edges.
- We form $G(t+1)$ from $G(t)$, $t \geq 0$ according to the following rules
  1. With probability α, add a new vertex v together with an edge from v to an existing vertex w, where w is chosen according to $d_{in} + \delta_{in}$.
  2. With probability β, add an edge from an existing vertex v to an existing vertex w, where v and w are chosen independently, v according to $d_{out} + \delta_{out}$ and w according to $d_{in} + \delta_{in}$.
  3. With probability $1 - \alpha - \beta$, add a new vertex w and an edge from an existing vertex v to w, where v is chosen according to $d_{out} + \delta_{out}$.

(A) [5 points] For which setting of the parameters $\alpha, \beta, \delta_{in}, \delta_{out}$ do we obtain the Barabási-Albert/Bollobás-Riordan model?

(B) [20 points] Implement the random graph model in the language of your preference. Set $t = 2000$. Simulate the model for various settings of the parameters. Plot the in-degree and the out-degree sequence in log-log scale for exactly three different settings of the parameters. For each setting of your choice, fit power law distributions to the in-degree and out-degree sequences and report the slopes.

(C) [25 points] Consider the following variation of the model, which introduces two extra parameters $0 < p < 1$ and $N \in \mathbb{Z}_+$. Whenever the model decides to take step (1) with probability α, we modify the model as follows. With probability p we choose v as before, namely together with an edge from v to an existing vertex w, where w is chosen according to $d_{in} + \delta_{in}$. With probability $1 - p$ we choose w from the N vertices with the highest in-degrees uniformly at random. Repeat (B) for this new model. What do you observe?