Hálózatok és a WWW Matematikája	**2011 ősz**

Feladatsor

Előadó: Benczúr András — 2011. október 29.

1. Egy uniform véletlen v csúcsból tegyünk egy d paraméterű geometriai eloszlás szerint hosszú véletlen sétát. Tekintsük a gráf csúcsain a séta *végpotja* által adott eloszlást. Geometriai eloszlás alatt azt értjük, hogy $\Pr(\text{a séta hossza } t) = d \cdot (1-d)^t$. Bizonyítsuk be, hogy a kapott eloszlás lesz a határértéke a PageRank-nek, azaz annak a bolyongásnak, amely egy ismeretlen eloszlásból indul és minden lépésben d valószínűséggel uniform véletlen eloszlásba, $(1-d)$ valószínűséggel pedig véletlen szomszédra ugrik.

2. Bizonyítsuk be, hogy összefüggő irányítatlan reguláris gráf esetén a bolyongás-mátrixnak sajátértéke a -1, ha a gráf páros, ellenkező esetben azonban az 1-en kívül nincs 1 abszolút értékű sajátérték.

3. Bizonyítsuk be, hogy véges erősen összefüggő gráf minden csúcsa nem-null perzisztens, azaz véges a visszatérés várható ideje.

4. Adjunk iteratív algoritmust, amely a $\text{PPR}_u^{(t+1)}(v)$ személyre szabott PageRank értékeket a $\text{PPR}_w^{(t)}(v)$ értékekből számítja. (Az órán látott algoritmus a $\text{PPR}_u^{(t)}(w)$-ket haszálja – tessék ellenőrizni!) Tételezzük fel, hogy a t-edik iteráció maximum hibája ϵ, azaz minden u, v csúcspárra $|\text{PPR}_u(v) - \text{PPR}_u^{(t+1)}(v)| < \epsilon$. Mekkora lesz a két algoritmusban a következő iteráció hibája? Melyik algoritmus jobb?

5. **A sortűz probléma.** Egy N processzorból álló lineáris hálózat minden eleme csak $O(1)$ memóriával rendelkezik, azaz többek között nem tudja megjegyezni vagy megszámolni N értékét. Mutassuk meg, hogyan koordinálhatók úgy, hogy a $O(N)$ lépés után egyszerre egy speciális állapotba kerüljenek. A celláknak egy közös órajelen és a bal oldalról indított start-jelen kívül más bemenete nincsen. (Alkalmazhatunk rekurziót!)

6. **Az SVD bizonyítás vége.** Az előadáson láttuk, hogy minden $n \times n$ és $m \times m$ méretű A és B mátrixokra létezik olyan u_1, σ_1 és v_1, hogy tetszőlegesen U_1 és V_1 unitér mátrixokká egészítve u_1-t és v_1-t, az UAV^T szorzatmátrix első sora és oszlopa $(\sigma_1, 0, \ldots, 0)$. Fejezzük be innen a bizonyítást indukcióval.

7. Bizonyítsuk be, hogy ha $\epsilon < c/2$, akkor legalább $\gamma(n/3)^2$ bit információ kell ahhoz, hogy egy n csúcsú gráfban meg tudjuk adni a közelítő $\hat{\text{Sim}}(u,v)$ értékeket úgy, hogy $u \in U$, $v \in V$, $|U| = |V| = n/3$ mellett
$$\text{Prob}(|\hat{\text{Sim}}(u,v) - \text{Sim}(u,v)| < \epsilon) > (1+\delta)/2.$$
Ötlet: Legyen $|U| = |V| = |X| = n/3$. A gráf irányított élei legyenek $x_i u_i$ $(i = 1, \ldots, n/3)$ és $x_i v_j$, ha $\text{bit}_{ij} = 1$.

8. Egy gráf c-expander, ha minden $|S| \leq n/2$ csúcshalmaznak legalább $c \cdot |S|$ különböző, $|S|$-en kívüli szomszédja van. Bizonyítsuk be, hogy egy ilyen gráfban bármely két csúcs távolsága legfeljebb $O(\log n)$. Határozzuk meg az $O()$-beli konstanst is!

9. Tegyük fel, hogy egy d-reguláris gráfban minden
$$|S| \leq \frac{n}{2} \cdot \frac{d+k}{d-k}$$
csúcshalmazból legalább $k \cdot |S|$ különböző él lép ki. Mutassuk meg, hogy a gráf c'-expander valamely c'-re. Adjuk meg a lehető legjobb c'-t.