Hálózatok és a WWW Matematikája	**2008 ősz**

Feladatsor

Előadó: Benczúr András 2008. november 17.

1. Definiáljuk a PageRank-et, mint annak a bolyongásnak a határértékét, amely egy ismeretlen eloszlásból indul és minden lépésben d valószínűséggel uniform véletlen eloszlásba, $(1-d)$ valószínűséggel pedig véletlen szomszédra ugrik. Bizonyítsuk be, hogy ha egy uniform véletlen csúcsból d paraméterű geometriai eloszlás szerint hosszú véletlen sétát teszünk, akkor az egyes csúcsokba pontosan a PageRank értéküknek megfelelő valószínűséggel érünk. Geometriai eloszlás alatt azt értjük, hogy $\Pr(\text{a séta hossza } t) = d \cdot (1-d)^t$.

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

4. Az előadáson látott exponenciális lecsengésű bolyongás felhasználásával konstruáljunk olyan gráfot, amelyben kevés csúcs hozzáadásával jelentősen módosítható a $d = 0$ értékhez tartozó elfajuló PageRank vektor. Becsüljük meg, hogy hány csúcs hozzáadásával mennyire tudjuk a PageRank értékeket az eredeti csúcsoktól „elszívni”.

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. Fejezzük ki a $\mathrm{Sim}(a, b)$ értékeket a gráf éleit megfordítva vett, pontosan k lépéses sétákból álló perszonalizált PageRank, az $\mathrm{RP}_u^{[k]}(a)$ értékek segítségével. *Útmutatás.*
   - Mi lesz $\sum_u \mathrm{RP}_u^{[k]}(a) \cdot \mathrm{RP}_u^{[k]}(b)$?
   - Indítsunk két egyidejű független véletlen sétát a-ból visszafelé. Legyen $\mathrm{SSim}^{(\ell)}(a) = \sum_t (1-c)^t \cdot \mathrm{Prob}$ (a két séta t lépés után megállítva azonos csúcsba ér és ezzel együtt legalább ℓ-szer találkozott). Határozzuk meg $\mathrm{SSim}^{(\ell+1)}(a)$-t az $\mathrm{SSim}^{(\ell)}(x)$-ek és az RP-k segítségével!
   - Szitaformula segítségével határozzuk meg az SSim-ekből $\sum_t (1-c)^t$ várható értékét, ahol t az első pillanat, hogy az a-ból induló két független egyidejű séta újra találkozik.
   - Határozzuk meg végezetül $\mathrm{Sim}(a, b)$-t.

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

8. Tegyük fel, hogy egy d-reguláris gráfban minden
$$|S| \le \frac{n}{2} \cdot \frac{d+k}{d-k}$$
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.