

# Bonusuppgift LANA, Block 2 Uppg. 7

Simon Sigurdhsson

27 april 2009

## 1 Formel för Google-matrisen $G$

$$G = (qA + (1 - q)A\bar{o}\bar{v}^T) * N \quad (1)$$

Där  $\bar{o}$  är en kolonnvektor med endast ettor och  $\bar{v}$  är en kolonnvektor genererad med matlab-kommandot `sum(A)/sum(sum(A))`, dvs. summan av alla kolonner i  $A$ , normerat för att summan av elementen i  $v$  ska bli ett.  $N$  är en diagonalmatris men elementen i diagonalen lika med ett genom summan av motsvarande kolonn i  $A$ . Denna genereras i matlab m.h.a. `diag(1./sum(A))`.

Detta ger en matris  $G$  som är en stokastisk matris oavsett värde på  $q$ .

## 2 Matriserna $A$ och $G$ för aktuellt nät

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \end{bmatrix} \quad (2)$$

$$G = G_1q + \frac{1}{33}G_2(1 - q) \quad (3)$$



### 3 Sidrankningar

Med  $q = 0,15$  fås följande rangordning:

<b>Sida</b>	1	2	3	4	5	6	7	8
<b>rank-värde</b>	0.2306	0.3423	0.3561	0.2399	0.2272	0.1960	0.1960	0.2286
<b>Rangordning</b>	8	3	2	6	10	13	12	9
<b>Sida</b>	9	10	11	12	13	14	15	
<b>rank-värde</b>	0.2097	0.1124	0.1130	0.3191	0.2366	0.4213	0.2448	
<b>Rangordning</b>	11	15	14	4	7	1	5	

#### 3.1 Förklaring av rangordning

Sidor som länkar till varandra (t.ex. sida 14 eller sida 2 och 3) får hög rang, eftersom de "hjälp" varandra. Rent matematisk är detta för att de båda ger varandra ett högre värde på  $p_i$ . Om de dessutom inte länkar till många andra sidor blir  $n_i$  litet och  $a_j$  alltså ännu större. Detta speglar det faktum att relevanta sidor ofta länkar till varandra, och motverkar effekten av sidor som masslänkar till sidor utan relevans.

#### 3.2 $q = 0$ och $q = 0,5$

För  $q = 0$  förstärks effekten av länkar markant (eftersom slumpmässiga hopp elimineras) och därmed rankas sidor som inte har ett balanserat antal inlänkar och utlänkar lågt. Detta syns t.ex. på sida 1, som rankas lägst med  $q = 0$ .

För  $q = 0,5$  får man en motsatt effekt. Sidor med många inlänkar rankas mycket högre eftersom den förstärkande effekten hos sidor som "hjälp" varandra minskas markant. Med  $q = 0,5$  rankas sida 10 högst.

#### 3.3 Ta bort sida 10

Inte mycket händer när sida 10 tas bort, den mest markanta skillnaden är att sidor som länkar till sida 10 får lite lägre ranking jämfört med de som inte gör det.