# Analiza poveznica

Središta i autoriteti na webu

Bruno Fabulić, Helena Marciuš, Dora Parmač

18. prosinca 2020.

# Sadržaj

# 1 Uvod

Budući živimo u svijetu informacija, Internet je dio našeg svakodnevnog žiota, te imamo pristup do više informacija nego što možemo procesuirati. Zbog toga nam je potrebna metoda za filtriranje informacija. Kao odgovor, javljaju se tehnike za pretraživanje.Njih možemo opisati kao proces koji nam omogućava pretraživanje veće kolekcije dokumenata za specifičnu informaciju koju nazivamo upit (eng. query).

Web je kolekcija dokumenata čije karakteristike ga čine posebno kompleksnim za pretraživanje. Pretraživači su suočeni s problemima kao što je sama veličina Weba, brzina kojom se mijenja i raste, te nedostatkom sustavne organizacije. Procjenjuje se da sadrži preko 45 milijardi stranica što ga čini najvećom kolekcijom dokumenata na svijetu, a to je samo indeksirani dio. Također, dinamičan je (godišnje nastaje milijarde novih dokumenata zbog čega je pretraživačima teško izračunati njihovu relevantnost za postavljeni upit) i nije sustavno organiziran jer svatko može stvoriti stranicu u bilo koju svrhu. Podatci konstantno nastaju i nestaju, a poveznice (eng. link) se stvaraju i pucaju jer odredišta postaju nedostupna.

Kao rješenja danog problema, pojavila se metoda analize poveznica. Stranice na webu povezane su hipervezama koju se web pretraživačima važan izvor informacija. Korisne su jer označavaju referencu stranice A na stranicu B što samo po sebi nije značajno korisno. No, ako pretpostavimo da autori stranice koriste poveznice za koje misle da će biti korisne čitatelju i ako su to poveznice na kvalitetne stranice koje obogaćuju sadržaj ili podupiru stajališta autora, onda na njih možemo gledati kao preporuku autora stranice A na stranicu B. Analiza poveznica se uspješno koristi za pronalaženje i indeksiranje novih stranica, rangiranje rezultata pretraživanja, te kategorizaciju web stranica.

## 2 HITS algoritam

Jedan od algoritama za rangiranje web stranice ovisno o upitu koje ćemo proučavati je Kleinbergov HITS algoritam.

Iz grafa poveznica G stvara se manji graf (eng. neighborhood graph) koji sadržava samo čvorove stranica koje odgovaraju upitu, proširen njihovim 'susjednim' stranicama. Susjedi u ovom slučaju znače set stranica koje poveznicama pokazuju na, ili na njih pokazuju dokumenti iz početnog seta koji odgovara upitu. Takav set stranica može biti jako velik pa se u praksi često ograničava na manji broj prethodnika. Rangiranje se potom vrši prema tom grafu uzimajući u obzir broj linkova vezanih uz čvor svake stranice iz početnog seta. Kao i prije problem je što svaka poveznica vrijedi jednako. Kod upita, algoritam prvo iterativno računa vrijednost središta (eng. hub score) i vrijednost autoriteta (eng. authority score) za svaki čvor iz grafa susjeda. Dokumenti se onda rangiraju prema obje vrijednosti.

Za dokumente s visokim autoritetom smatra se da bi trebali imati relevantni sadržaj. Dokumenti središta trebali bi imati poveznice na relevantne dokumente.Ideja je da bi dokumenti koji pokazuju na mnoge druge mogli biti dobra središta, a dokumenti na koje pokazuju mnogi drugi dokumenti dobri autoriteti.

# 3 WWW kao usmjereni graf

Svaku kolekciju povezanih web stranica možemo predstaviti usmjerenim grafom $G = (V, E)$. Web stranice predstavljamo vrhovima, tj. $V$ je skup web stranica, a poveznice između stranicama predstavljamo bridovima - za stranice $p_i$, $p_j \in V$, brid $e_{ij}$ je u $E$ ako postoji poveznica sa stranice $p_i$ na stranicu $p_j$. Kažemo da je stupanj izlaznosti (out - degree) vrha $p_i$ broj vrhova na koje $p_i$ pokazuje, a stupanj ulaznosti (in-degree) broj vrhova koji pokazuju na $p_i$.

Kažemo da je vrh i

- dobar izvor informacije o lokaciji kvalitetnog sadržaja ili dobro čvorište (hub) ako sadrži linkove na vrhove koju su dobri autoriteti u smislu sadržavanja kvalitetne informacije
- dobar autoritet (sadrže kvalitetnu informaciju o nekoj temi) ako na njega pokazuju dobri hubovi)

Dakle, svakom vrhu, odnosno svakoj stranici $p_i$ pridružujemo uređen par $(x_i, y_i)$ nenegativnih brojeva kao mjeru za "biti dobar autoritet" i "biti dobar "hub".

$$\overbrace{\bullet}^{hub} \rightarrow \underbrace{\bullet}_{autoritet} \leftarrow \overbrace{\bullet}^{hub}; \qquad \overbrace{\bullet}^{autoritet} \leftarrow \underbrace{\bullet}_{hub} \rightarrow \overbrace{\bullet}^{autoritet}$$

Težine $x_i \geq 0$ i $y_i \geq 0$ definirano rekurzivno s:

$$x_i = \sum_{j: e_{ji} \in E} y_j \qquad y_i = \sum_{j: e_{ij} \in E} x_j \tag{1}$$

Odnosno, autoritet-vrijednost web stranice je suma hub-vrijednosti svih stranica koje pokazuju na nju. Hub vrijednost web stranice je suma autoritet-vrijednosti web stranica na koje ona pokazuje. Intuitivno, smatramo da je web stranica dobar autoritet ako na nju pokazuju dobri hubovi. Analogno, smatramo da je web stranica dobar hub ako pokazuje na dobre autoritete.

Kako odrediti $x_i, y_i$ ?
Stavimo $x = (x_1 \dots x_n)^T, y = (y_1 \dots y_n)^T$ i

$$L = (L_{i,j})_{i,j=1}^n$$

$$L_{i,j} = \begin{cases} 1 & \text{ako } i \longrightarrow j \\ 0 & \text{inače} \end{cases}$$

Ako gornji princip iteriramo i generiramo nizove $x^{(k)} i y^{(k)} = Lx^{(k)}$ dobijemo $x^{(k+1)} = L^{(k)} y^{(k)}, y^{(k+1)} = Lx^{(k+1)} = LL^T y^{(k)}$.
Te relacije korigiramo normiranjem u

$$x^{(k+1)} = \frac{x^{(k+1)}}{||x^{(k+1)}||_2} \qquad y^{(k+1)} = \frac{y^{(k+1)}}{||y^{(k+1)}||_2}$$

Dakle, dobili smo dva niza generirana metodom potencija za $L^T L i L L^T$. Dobivene matrice imaju iste svojstvene vrijednosti, pa ako je $\lambda_1$ dominantna svojstvena vrijednost, onda imamo konvergenciju: postoje $x i y, ||x||_2 = ||y||_2 = 1$ tako da je

$$L^T L x = \lambda_1 x \qquad LL^T y = \lambda_1 y$$

## 3.1 Matrična formulacija

Za usmjereni graf $G = (V, E)$, definiramo matricu linkova:

$$L = \begin{cases} 1, & e_{ij} \in E \\ 0, & e_{ij} \notin E \end{cases}$$

Autoritet-vrijednosti $x_i$ čine vektor autoriteta $x = (x_1, x_2, \dots, x_n)$, a hub vrijednosti čine vektor hubova $y = (y_1, y_2, \dots, y_n)$. Sada se jednadžbe (1) mogu zapisati kao

$$x = L^T y \qquad y = Lx \tag{2}$$

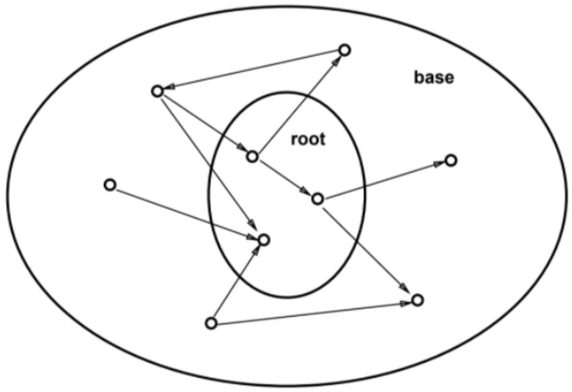
# 4 Podgraf fokusiran na upit

Pretpostavimo da je dan upit $\sigma$. Želimo pronaći podgraf grafa G na kojem ćemo izvesti algoritam. Mogli bi se ograničiti na skup svih stranica koje spominju upit $\sigma$. Međutim, broj takvih stranica može biti vrlo velik što bi uzrokovalo velikim "troškom" kod računanja. Također, kao što smo prije zaključili, najbolji autoriteti možda neće biti u tom skupu.
Htjeli bismo kolekciju $S_\sigma$ koja ima sljedeća svojstva

1. $S_\sigma$ je relativno mali skup
2. $S_\sigma$ sadrži većinom sadrži relevantne stranice
3. $S_\sigma$ sadrži većinu dobrih autoriteta

## 4.1 Konstrukcija $S_\sigma$

Za zadani upit $\sigma$, pronađemo relevantne web stranice i rangiramo ih (npr. broj pojavljivanja upita na stranici, PageRank) i odaberemo $k$ najbolje rangiranih stranica. Na taj način dobivamo skup $R_\sigma$ koji nazivamo početni skup("root set"). Taj skup zadovoljava 1. i 2. svojstvo, ali možda ne zadovoljava 3. svojstvo. Koristeći ovaj skup, konstruiramo skup $S_\sigma$ na sljedeći način: za svaku stranicu $p$ u $R_\sigma$, u $R_\sigma$ dodamo sve stranice na koje $p$ pokazuje i $d$ proizvoljnih stranica koje pokazuju na $p$.
Ograničenje na broj stranica koje pokazuju na $p$ osigurava da je skup $s_\sigma$ relativno mali (svojstvo (1)), a iz konstrukcije se vidi da je zadovoljeno svojstvo (2).
Pretpostavimo da je $q$ dobar autoritet za zadani upit. Iako se $q$ možda ne nalazi u $R_\sigma$, vrlo je vjerojatno da barem jedna od stranica iz $R_\sigma$ pokazuje na $q$ pa se stoga $q$ nalazi u $S_\sigma$, tj. zadovoljeno je i svojstvo (3).
Podgraf grafa $G$ induciran skupom $S_\sigma$ označavamo s $G[S_\sigma]$ i na njemu provodimo HITS algoritam.




Slika 1: Proširivanje root set-a u base set

## 5 Implementacija HITS algoritma

Iterativno računamo autoritet-vrijednosti i hub-vrijednosti. Sa $x^{(k)}$ i $y^{(k)}$ označimo vrijednost vektora $x$ odnosno $y$ u $k$-toj iteraciji. Iz jednadžbi (2) slijedi

$$cx^{(k+1)} = L^T L x^{(k)} \qquad cy^{(k+1)} = LL^T y^{(k)} \tag{3}$$

uz početni uvjet $x^{(0)} = y^{(0)} = (1, 1, \ldots, 1)$, gdje je $c$ konstanta takva da vrijedi $||x||_2 = 1$ i $||y||_2 = 1$. Matricu $L^T L$ nazivamo matrica autoriteta, a matricu $LL^T$ nazivamo matrica hubova. Uočimo da su ove matrice simetrične.

Možemo zapisati algoritam:

1. Ulaz: Matrica linkova $L$
2. Inicijaliziramo $x^{(0)} = (1, 1, \ldots, 1)$, $y^{(0)} = (1, 1, \ldots, 1)$
3. Računaj $k$-tu iteraciju
$$x^{(k+1)} = L^T L x^{(k)} \qquad y^{(k+1)} = LL^T y^{(k)}$$
4. Normiraj dobivene vrijednosti
$$x^{(k+1)} = \frac{x^{(k+1)}}{||x^{(k+1)}||_2} \qquad y^{(k+1)} = \frac{y^{(k+1)}}{||y^{(k+1)}||_2}$$
5. Provjeri kriterij zaustavljanja

- Ako kriterij zaustavljanja nije zadovoljen, idi na 3, korak
- Ako je kriterij zaustavljanja zadovoljen, završi

6. Izlaz: Vektori vrijednosti autoriteta, odnosno hubova, $x$ i $y$

$i$-ta komponenta vektora $x$ predstavlja autoritet-vrijednost web stranice $p_i$, a $i$-ta komponenta vektora $y$ predstavlja hub-vrijednost web stranice $p_i$.

### 5.1 Implementacija HITS-a u Octaveu

```
function [] = hits(L, W,  n)
C = zeros(n,n);
R = zeros(n,n);
k = 0;
eps = power (10, -5);
xp = ones(n, 1);
yp = ones(n, 1);
x = zeros(n,1);
y = zeros(n,1);
con = 1;
X = (L')*L;
Y = L*(L');

while con
 x = X*(xp);
 x/=norm(x);
 y = Y*(yp);
  y/=norm(y);
   if norm(x-xp) < eps && norm(y-yp) < eps
     con = 0;
   else
     xp=x;
     yp=y;
     ++k;
   endif
endwhile

[x_sort, ix] = sort(x, 'descend');
[y_sort, iy] = sort(y, 'descend');

disp("Prvih dvadeset autoriteta ");
for i = 1:20
  disp(W(ix(i)));
endfor
disp("Prvih dvadeset hubova ");
for i = 1:20
  disp(W(iy(i)));
endfor
```

# 6 Sličnost s bibliometrijskom analizom

Bibliometrija se bavi kvantitavnim proučavanjem pisanih dokumenata i pojava kao što su produktivnost autora, citiranje, disperzija članaka, učestalost riječi.
Ko-citacija dokumenata $a$ i $b$ je mjera sličnosti koja se definira kao frekvencija citiranja ta dva dokumenta, tj. broj dokumenata koji citiraju dokumente $a$ i $b$. Što je veći broj dokumenata koji citiraju $a$ i $b$, njihova je ko-citacija veća pa je veća vjerojatnost da su $a$ i $b$ semantički slični.
Slično, ko-referenca dokumenata $a$ i $b$ je mjera sličnosti koje se definira kao broj dokumenata na koje $a$ i $b$ imaju referencu. Ako $a$ i $b$ imaju veću ko-referencu, veća je vjerojatnost da $a$ i $b$ obrađuju sličnu temu.
Na sličan način možemo definirati mjere sličnosti između autoriteta, odnosno hubova.

## 6.1 Autoriteti i ko-citacije

Ako na web stranice $p_i$ i $p_j$ pokazuje veći broj stranica, vjerojatnije je da su one na neki način slične. Za web stranice $p_i$ i $p_j$ definiramo ko-citacije kao broj web stranica koje pokazuju na $p_i$i $p_j$. Matrično se ovo može zapisati kao

$$C_{ij} = \sum_k L_{ki}L_{kj} = \sum_k (L^T)_{ik}L_{kj} = (L^TL)_{ij} \tag{4}$$

uz $C_{ii} = 0$. Također vrijedi simetrija, tj. $C_{ij} = C_{ji}$.
Označimo sa $d_i$ stupanj ulaznosti stranice $p_i$. $d_i$ možemo računati kao

$$d_i = \sum_k L_{ki} = \sum_k L_{ki}L_{ki} = (L^TL)_{ii} \tag{5}$$

jer je $L_{ki} = L_{ki}^2$ jer je $L_{ki} = 1$ ili $L_{ki} = 0$.
Definiramo matricu stupnjeva ulaznosti $D$:

$$D = diag(d_i, d_2, \ldots, d_n) \tag{6}$$

Tada matrica autoriteta ima sljedeću strukturu:

$$L^TL = D + C \tag{7}$$

odnosno, matrica autoriteta je zbroj matrice ko-citacija i matrice stupnjeva ulaznosti.

### 6.2 Hubovi i ko-reference

Za web stranice $p_i$ i $p_j$ definiramo ko-reference kao broj web stranica na koje pokazuju $p_i$i $p_j$. Matrično se ovo može zapisati kao

$$R_{ij} = \sum_k L_{ik}L_{jk} = \sum_k L_{ik}(L^T)_{kj} = (LL^T)_{ij} \tag{8}$$

uz $R_{ii} = 0$. Također vrijedi simetrija, tj. $R_{ij} = R_{ji}$.
Označimo sa $o_i$ stupanj izlaznosti stranice $p_i$. $o_i$ možemo računati kao

$$o_i = \sum_k L_{ik} = \sum_k L_{ik}L_{ik} = (LL^T)_{ii} \tag{9}$$

jer je $L_{ik} = L_{ik}^2$ jer je $L_{ik} = 1$ ili $L_{ik} = 0$.
Definiramo matricu stupnjeva izlaznosti $O$:

$$O = diag(o_i, o_2, \ldots, o_n) \tag{10}$$

Tada matrica autoriteta ima sljedeću strukturu:

$$LL^T = O + R \tag{11}$$

odnosno, matrica hubova je zbroj matrice ko-referenca i matrica stupnjeva izlaznosti.

Također, imamo nejednakost:

$$max\{0, o_i + o_k - n\} \leq R_{ik} \leq min\{o_i, o_k\} \tag{12}$$

Direktna posljedica ove tvrdnje je $R_{ik} = 0$ ako $o_i = 0$ ili $o_k = 0$. Dakle, ako web stranica $p_i$ stupanj izlaznosti 0, tada je i-ti red $LL^T$ jednak 0. Iz jednadžbe (3), slijedi da hub score mora biti 0.

## 7 Konvergencija algoritma

Primijetimo da je $L^TL = (LL^T)^T$. Stoga, matrice $L^TL$ i $LL^T$ imaju iste svojstvene vrijednosti. Kako su to realne i simetrične matrice, njihove svojstvene vrijednosti su također realne.
Neka su $\lambda_1$, $\lambda_2$, $\ldots$, $\lambda_n$ svojstvene vrijednosti ovih matrica poredane padajuće po apsolutnoj vrijednosti, tj. vrijedi $|\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_n|$. Pretpostavimo da vrijedi $|\lambda_1| > |\lambda_2|$ Svojstvenu vrijednost $\lambda_1$ nazivamo dominantna svojstvena vrijednost.

**Teorem 7.1.** *Nizovi* $x^{(1)}, x^{(2)}, x^{(3)} \ldots$ *i* $y^{(1)}, y^{(2)}, y^{(3)} \ldots$ *konvergiraju limesima* $x^*$ *i* $y^*$*.*

*Dokaz.* Neka je $G = (V, E)$, gdje je $V = \{k_1 \ldots k_n$ i neka je A adjunktna matrica grafa G. Element na mjestu $(i, j)$ jednak je 1 ako je $(p_i, p_j)$ brid grafa $G$, a 0 inače. Lako se provjeri da operacije $I$ i $O$ mogu biti napisane kao $x \leftarrow A^Ty$ i $y \leftarrow Ax$, redom. Tada je $x_k$ jedinični vektor u smjeru $(A^TA)^{k-1}A^Tz$, a $y_k$ jedinični vektor u smjeru $(AA^T)^kz$.

Linearna algebra nam kaže da ako je $M$ simetrična $nxn$ matrica i vektor $v$ nije ortogonalan svojstvenom vektoru $\omega_1(M)$ tada jedinični vektor u smjeru $Mkv$ konvergira prema $\omega_1(M)$ kako se k povećava. Također, M ima samo nenegativne ulaze, pa glavni svojstveni vektor od $M$ ima samo nenegativne vrijednosti.

Posljedično, z nije ortogonalan $\omega_1(AA^T)$ pa tada niz $\{y_k\}$ konvergira prema $y^*$. Na sličan način se pokaže da ako je $\lambda_1 \neq 0$, tada $A^Tz$ nije ortogonalan na $\omega_1(AA^T)$. Iz toga slijedi da niz $\{x_k\}$ konvergira prema $x^*$. □

Dokaz ovog teorem daje sljedeći rezultat.

**Teorem 7.2.** $x^*$ *je dominantni svojstveni vektor matrice* $L^TL$*, a* $y^*$ *je dominantni svojstveni vektor matrice* $LL^T$*.*

## 8 Vjerojatnosna analiza

Analizirali smo strukturu autoriteta i hubova. Rezultati jednadžbe LALA pokazuju zanimljivu vezu između ko-citacija i stupnja ulaznosti: općenito, čvorovi s velikim izlaznim stupnjem će imati velike ko-citacije s ostalim čvorovima samo zato jer imaju više ulaznih linkova/bridova. Slično, veliku ko-referencu povezujemo s velikim izlaznim brojem linkova.

Ovakva intuicija se može precizirati ako pretpostavimo da (web) graf ima nasumično raspoređen graf fiksnog stupnja te koristeći vjerojatnosnu analizu na očekivanim vrijednostima ko-citacije i ko-reference. Općenito govoreći, web graf je nasumičan graf - milijuni pojedinaca i organizacija kreiraju web stranice za različite svrhe. Predlaže se da je web bolje opisan nasumičnim grafom fiksnog supnja, unutar kojeg su čvorovi stupnjeva $\{d_1 \dots d_n\}$ prvo dani, a bridovi su nasumično raspoređeni između čvorova koji podliježu ograničenjima stupnjeva čvorova. Stoga imamo sljedeću tvrdnju:

**Propozicija 8.1.** *Prosječna vrijednost ko-citacija dana je formulom*

$$\langle C_{ik} \rangle = \frac{d_i d_k}{n-1} \tag{13}$$

*Dokaz.* Pretpostavimo da je $d_i \geq d_k$. Tada iz jednadžbe (3) vidimo da ima barem $d_k$ elemenata različitih od 0, što je INNER PRODUCT i-tog retka i k-tog stupca matrice L.Promotrimo slučaj kada je q-ti red u k-tom stupcu jednak 1. Vjerojatnost da je vrijednost u i-tom stupcu jednaka 1 jest:

$$P(L_{qi} = 1) = \frac{C_{n-2}^{d_i - 1}}{C_{n-1}^{d_i}} = \frac{d_i}{n-1}$$

Ovdje je $C_{n-1}^{d_i}$ ukupan broj svih mogućih rasporeda $d_i$ jedinica u i-tom stupcu, a $C_{n-2}^{d_i-1}$ je ukupan broj svih mogućih rasporeda ako je jedinica u retku q. Dakle,

$$\langle C_{ik} \rangle = \sum_q \langle L_{qi}, L_{qk} \rangle = \sum_q^{d_k} \langle L_q \rangle = d_k * P(L_{qi} = 1)$$

odakle slijedi (13) □

Iz ove analize, vidimo da će vrh $i$ s velikim izlaznim stupnjem $d_i$ imati veliku ko-citaciju s drugim vrhovima, a ako to usporedimo s vrhom $j$ koji ima manji izlazni stupanj $d_j$ tj. ako je $d_i > d_j$, tada je:

$$\langle C_{ik} \rangle > \langle C_{jk} \rangle \qquad \forall k, k \neq i, k \neq j$$

Iz ove probabilističke jednadžbe, zaključujemo da je $C_{ik}$ veći od $C_{jk}$ većinu vremena, što nije nužno istina u svakom slučaju. Praktičnosti radi, kažemo da *u prosjeku* vrijedi $C_{ik} \gtrsim C_{jk}$.

Na sličan se način ova analiza može primijeniti za izlazni stupanj i ko-referencu hubova matrice $LL^T$. Imamo:

$$\langle R_{ik} \rangle = \frac{o_i o_k}{n-1}$$

Ako je $o_i > o_j$, tada $\langle R_{ik} \rangle > \langle R_{jk} \rangle$, odnosno kažemo da $R_{ik} \gtrsim R_{jk}$. vrijedi *u prosjeku.*

# 9 Average case analiza

Zbog dosadašnjih analiza, sada možemo zamijeniti matricu autoriteta njihovim prosječnim vrijednostima.

**Teorem 9.1.** *Matrica autoriteta $L^T L$ u prosječnom slučaju, uz uvjet $d_[i] + d_j < n+1$ za svaki $i, j$ ima sljedeće svojstvene vrijednosti i svojstvene vektore:*

1. *Za svojstvene vrijednosti vrijedi*

$$\lambda_1 > \hat{d}_1 > \lambda_2 > \hat{d}_2 > \cdots > \lambda_n > \hat{d}_n \tag{14}$$

2. *k-ti svojsteni vektor je*

$$\boldsymbol{u}_k = \left( \frac{d_1}{\lambda_k - \hat{d}_1}, \frac{d_2}{\lambda_k - \hat{d}_2}, \ldots, \frac{d_n}{\lambda_k - \hat{d}_n} \right)^T \tag{15}$$

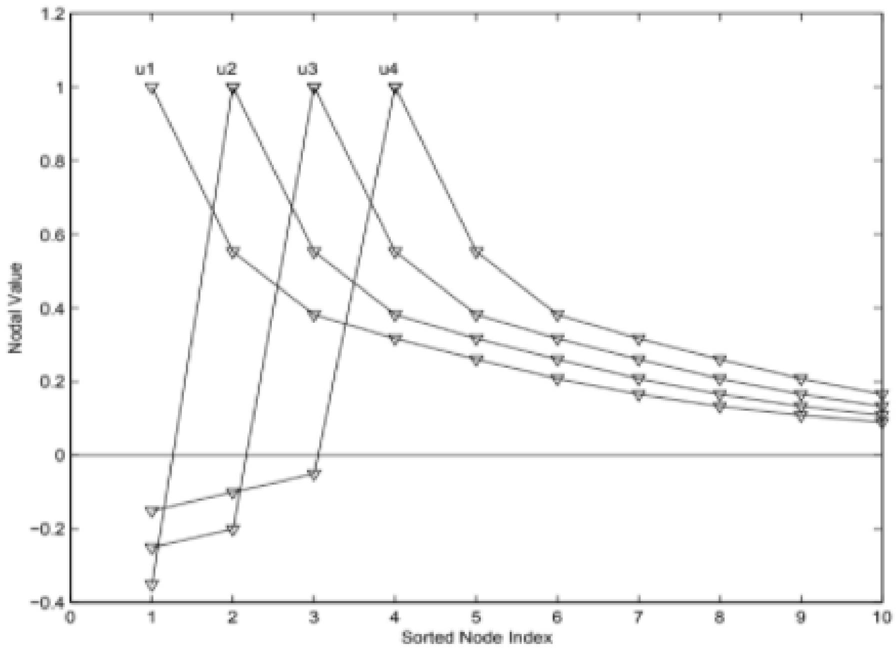
*Web stranice indeksiramo tako da je $d_1 > d_2 > \cdots > d_n$ i vrijedi $\hat{d}_i = d_i - \frac{d_i^2}{n-1}$.*

*Dokaz.* Koristeći izraz (13), prosječna matrica autoriteta je

$$\langle L^T L \rangle = \langle D \rangle + \langle C \rangle = diag(\hat{d}_1, \ldots, \hat{d}_n) + \frac{1}{n-1} dd^T \tag{16}$$

□

Analogan rezultat vrijedi za matricu hubova $LL^T$.




Slika 2: Svojstveni vektori jednadžbe

## 10 Svojstva HITS algoritma

Nekoliko zanimljivih rezultata slijedi direktno iz prethodnog teorema:

1. **Organiziranje web stranica.** Rangiranje autoriteta, je u prosjeku, identično kao rangiranje stranica pomoću ulaznih stupnjeva. Kako bi nam bilo to jasnije, imamo sljedeću tvrdnju:

   **Korolar 10.0.1.** *Elementi glavnog svojstvenog vektora $u_1$ monotono padaju.*

   *Dokaz.* Za svaki $i < j$:

   $$u_1(i) - u_1(j) = \frac{d_i}{\lambda_1 - \hat{d}_i} - \frac{d_j}{\lambda_1 - \hat{d}_j} = \frac{(d_i - d_j)[\lambda_1 - \frac{d_i d_j}{n-1}]}{(\lambda_1 - \hat{d}_i)(\lambda_1 - \hat{d}_j)} > 0$$

   □

   Iz ovoga zaključujemo da je rangiranje web stranica pomoću njihovih autority scorova isto kao i rangiranje prema ulaznim stupnjevima. Praktična primjena ovih rezultata jest da je jednostavno brojanje ulaznih bridova kao algoritam rangiranja učinkovito i efikasno.

2. **Jedinstvenost.** Ako je $d_1 > d_2$, tada je glavni svojstveni vektor $L^T L$ jedinstven i različit od drugog glavnog svojstvenog vektora.

3. **Konvergencija**. Konvergencija HITS algoritma može biti dosta brza. Početni vektor $x^{(0)} = (1 \dots 1)$ ima vrlo malo preklapanja s ostalim svojstvenim vektorima ($x^{(0)} * u_k, k > 1$) jer svi oni sadrže negativne vrijednosti čvora. Koristeći $L^T L = \lambda_1 u_1 u_1^T + \lambda_2 u_2 u_2^T + \dots$ nakon k iteracija, imamo:

   $$x^{(k)} = c_1 \lambda_1^k u_1 + c_2 \lambda_2^k u_2 + \dots$$

   gdje je $c_2 \ll c_1$ zbog malih preklapanja između $x^{(0)}$ i $u_2$

4. **Web zajednice.** HITS algoritam se uvijek koristio za identificiranje mnogih web zajednica koristeći različite svojstvene vrijednosti. Glavni svojstveni vektor definira dominantnu web zajednicu. Svaki sporedni svojstveni vektor definira dvije zajednice, jednu s nenegativnim $\{i | u_k(i) \geq 0\}$ i drugu s negativnim vrijednostima $\{i | u_k(i) < 0\}$