

## Mere pomembnosti posamezne enote

Mere pomembnosti (prestiža) ponavadi računamo samo za usmerjena omrežja, saj je smer bistvena lastnost povezave.

**Primer:** Ljudje, ki jih izbere za prijatelje veliko drugih, imajo v tej skupini poseben položaj – prestiž. Prestiž postane še posebno pomemben, če te izbire niso recipročne – npr. če se vsi otroci radi igrajo z najbolj popularnim otrokom, le-ta pa se ne igra rad z vsemi ampak le z določenimi.

### Vhodna stopnja točke

Kot smo že omenili je ena od mer za prestiž tudi vhodna (ali izhodna) stopnja. Glede na pomen relacije predstavljata podporo ali vpliv. V nadaljevanju se bomo omejili na relacijo pozitivne izbire (npr. *ima rad*, *glasuje za*).

Stopnja točke je zelo omejena mera pomembnosti, saj ima precej slabosti:

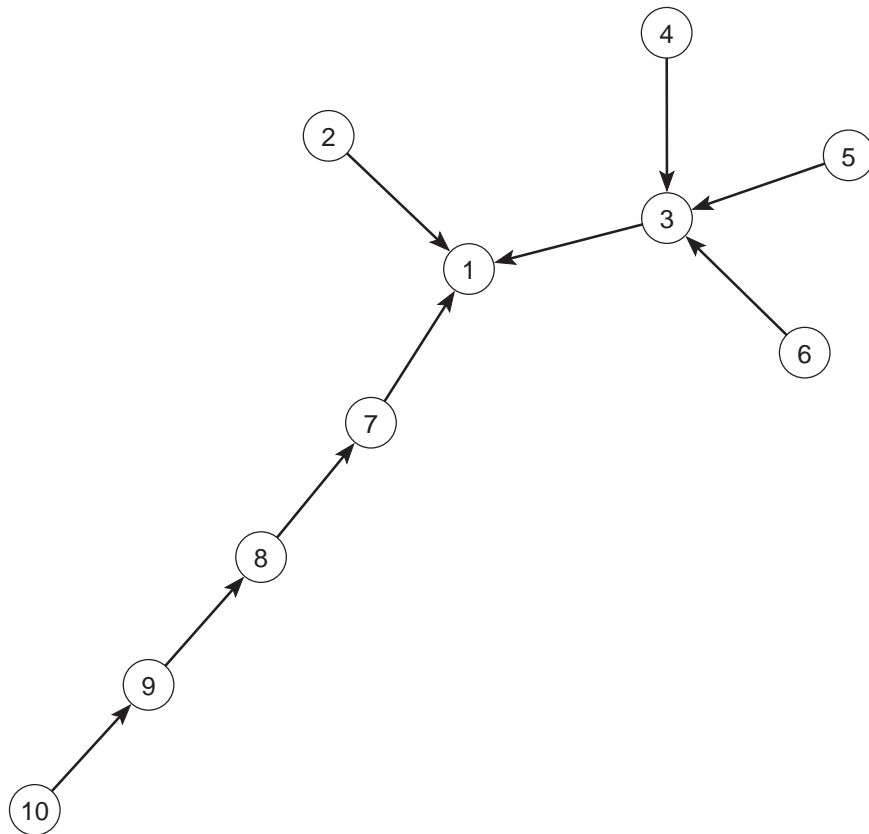
1. Stopnja točke upošteva samo neposredne izbire.
2. Pri stopnji točke ni pomembno ali je neka oseba izbrana od ljudi, ki so sami velikokrat izbrani (pomembni) ali pa so sami le redko izbrani (manj pomembni) – pri stopnji točke se upošteva samo *lokalna* ne pa tudi *globalna* struktura omrežja.

## Območje vpliva

Zaradi omenjenih slabosti vhodne stopnje točke je boljša mera *območje vpliva* (*influence domain*). Območje vpliva dane točke v usmerjenem omrežju je število (ali delež) vseh drugih točk, iz katerih lahko ob upoštevanju smeri povezav dosežemo izbrano točko. Delež točk v območju vpliva izračunamo tako, da število točk delimo z  $n - 1$  (število vseh preostalih točk).

Glede na definicijo območja vpliva je to mero smiselno računati le, če omrežje ni krepko povezano (sicer so v območju vpliva vsake točke vse druge točke v omrežju). Za krepko povezano omrežje, pa smo že omenili, da je mera pomembnosti lahko kar dostopnost (closeness).

Tudi območje vpliva ima kot mera pomembnosti nekaj slabosti: ne razlikuje med neposrednimi in posrednimi izbirami. Ponavadi smatramo neposredne izbire za bolj pomembne kot posredne izbire (izbire sestavljene iz daljše verige posrednikov) .



V območju vpliva točke 1 so vse ostale točke. Npr. izbire oseb 2, 3 in 7 so bolj pomembne za osebo 1 kot posredne izbire od oseb 4, 5, 6 in 8. Osebi 9 in 10 prispevata še občutno manj k vplivnosti osebe 1.

Večji pomen neposrednim izbiram damo z utežitvijo izbire z oddaljenostjo točke od izbrane osebe. Večja oddaljenost prispeva manj k celotni vplivnosti osebe. V ta namen izračunamo *povprečno oddaljenost točke* od vseh točk v njenem območju vpliva. Npr. povprečna oddaljenost točke 1 od točk v njenem območju vpliva je 2.0 ( $18/9$ ).

Na žalost pa ima tudi povprečna oddaljenost pomankljivost: Na sliki imata najmanjšo povprečno oddaljenost od točk v območju vpliva točki 3 in 9, saj sta neposredno izbrani od vseh točk v njunih območjih vpliva (4, 5 in 6, oziroma 10).

Če pa pogledamo sliko omrežja, pa ti dve točki gotovo nista bolj vplivni kot točka 1. Njuni majhni območji vpliva sta krivi za to, da imata nizko povprečno oddaljenost.

Za boljšo mero vplivnosti moramo pri povprečni oddaljenosti upoštevati še velikost območja vpliva.

## Bližina – Proximity prestige

*Bližino – Proximity Prestige (PP)* izbrane točke dobimo tako, da delež točk v območju vpliva izbrane točke delimo s povprečno oddaljenostjo te točke od vseh drugih točk v njenem območju vpliva.

Veliko območje vpliva in majhne razdalje povzročijo velik *PP*. Največji *PP* dosežemo, če je neka točka neposredno izbrana od vseh ostalih točk v omrežju (delež območja vpliva je 1, povprečna oddaljenost pa ravnотako).

Na sliki imajo vse robne točke (2, 4, 5, 6 in 10) prazno območje vpliva in zato  $PP = 0$ .

Območje vpliva točke 9 vsebuje samo točko 10 (delež je  $\frac{1}{9} = 0.11$ ). Povprečna oddaljenost je točke 9 je 1, tako da je *PP* za točko 9 enak  $PP = \frac{0.11}{1} = 0.11$ .

Točka 1 ima največje območje vpliva (9 točk = 100%).

Povprečna oddaljenost je 2.0, tako da je  $PP = \frac{1.00}{2.0} = 0.5$ .

Vse ostale točke imajo *PP* nižji od točke 1, a višji od točke 9.

Tabela prikazuje območje vpliva, relativno območje vpliva, povprečno oddaljenost in proximity prestige za vse točke s prejšnje slike.

| točka | $OV$ | $OV_r = \frac{OV}{9}$ | $\bar{d}$ | $PP = \frac{OV_r}{\bar{d}}$ |
|-------|------|-----------------------|-----------|-----------------------------|
| 1.    | 9    | 1.00                  | 2.00      | 0.500                       |
| 2.    | 0    | 0.00                  | ?         | 0.000                       |
| 3.    | 3    | 0.33                  | 1.00      | 0.333                       |
| 4.    | 0    | 0.00                  | ?         | 0.000                       |
| 5.    | 0    | 0.00                  | ?         | 0.000                       |
| 6.    | 0    | 0.00                  | ?         | 0.000                       |
| 7.    | 3    | 0.33                  | 2.00      | 0.167                       |
| 8.    | 2    | 0.22                  | 1.50      | 0.148                       |
| 9.    | 1    | 0.11                  | 1.00      | 0.111                       |
| 10.   | 0    | 0.00                  | ?         | 0.000                       |

Datoteka infl.net.

Poseben primer:

Če je omrežje krepko povezano, je Proximity Prestige kar enak dostopnosti glede na vhodne povezave.

## Primer

Za 22 držav imamo podatke o uvozu / izvozu med njimi.

Podatki so na datoteki [wass749a.net](http://wass749a.net) – Trade of basic manufactured goods between countries, Wasserman & Faust, 749.

Usmerjena povezava od države  $x$  do države  $y$  pomeni, da država  $x$  izvaža v državo  $y$ . Mere pomembnosti nam v tem primeru povejo katere države so pomembni izvozniki in katere pomembni uvozniki.

V tem primeru je najbolj smiselna mera pomembnosti stopnja točk (vhodna in izhodna).

Omrežje je krepko povezano, zato je Proximity Prestige kar enak dostopnosti glede na vhodne povezave ( $C_{Cin}$ ).

| Št. | Država | $C_{Dout}$ | $C_{Cout}$ | $C_{Din}$ | $C_{Cin}$ |
|-----|--------|------------|------------|-----------|-----------|
| 1.  | Alg    | 0.190      | 0.553      | 0.619     | 0.724     |
| 2.  | Arg    | 0.619      | 0.724      | 0.476     | 0.657     |
| 3.  | Bra    | 0.905      | 0.913      | 0.524     | 0.677     |
| 4.  | Chi    | 0.905      | 0.913      | 0.714     | 0.778     |
| 5.  | Cze    | 0.905      | 0.913      | 0.619     | 0.724     |
| 6.  | Ecu    | 0.095      | 0.525      | 0.429     | 0.636     |
| 7.  | Egy    | 0.429      | 0.636      | 0.571     | 0.700     |
| 8.  | Eth    | 0.095      | 0.525      | 0.476     | 0.656     |
| 9.  | Fin    | 0.953      | 0.955      | 0.714     | 0.778     |
| 10. | Hon    | 0.048      | 0.512      | 0.429     | 0.636     |
| 11. | Ind    | 0.667      | 0.750      | 0.667     | 0.750     |
| 12. | Isr    | 0.524      | 0.677      | 0.476     | 0.656     |
| 13. | Jap    | *1.000     | *1.000     | *0.810    | *0.840    |
| 14. | Mad    | 0.048      | 0.512      | 0.286     | 0.583     |
| 15. | NZ     | 0.524      | 0.677      | 0.667     | 0.750     |
| 16. | Pak    | 0.524      | 0.677      | 0.667     | 0.750     |
| 17. | Spa    | 0.952      | 0.955      | *0.810    | *0.840    |
| 18. | Swi    | *1.000     | *1.000     | 0.714     | 0.778     |
| 19. | Tai    | 0.619      | 0.724      | 0.714     | 0.778     |
| 20. | UK     | 0.952      | 0.955      | 0.762     | *0.808    |
| 21. | US     | *1.000     | *1.000     | *0.905    | *0.913    |
| 22. | Yug    | 0.810      | 0.840      | 0.714     | 0.778     |
|     | Usred. | 0.412      | 0.495      | 0.307     | 0.380     |



## Kazala in opisi

Poglejmo si še dve meri pomembnosti, ki sta še posebej uporabni v primeru omrežja povezav med predstavitvenimi stranmi na Internetu (usmerjeno omrežje).

V takem omrežju si lahko predstavljamo, da igra vsaka točka (predstavitvena stran) dvojno vlogo. Ponavadi stran nekaj opisuje (je opis – *authority* in se druge nanjo sklicujejo), lahko pa tudi kaže na eno ali več drugih strani v omrežju (je kazalo – *hub*). Zanimali nas bodo samo dobri opisi in dobra kazala.

Točka je dobro kazalo, če kaže na veliko dobrih opisov in je dober opis, če nanjo kaže veliko dobrih kazal.

Vsaki točki  $v$  priredimo dve uteži  $x_v, y_v \in [0, 1]$ , ki bosta določali, kako dober opis oz. kazalo je izbrana točka.

Pravimo, da je točka  $v$  boljši opis kot točka  $u$ , če je  $x_v > x_u$ .

Podobno velja tudi za kazala.

Uteži naračunamo iz sosednostne strukture omrežja, tako da rešimo problem lastnih vrednosti matrik  $AA^T$  (kazala) oziroma  $A^T A$  (opisi).

## Primer

Na svetovnem nogometnem prvenstvu v Parizu leta 1998 je sodelovalo 32 nogometnih reprezentanc. Nogometaši pogosto igrajo v nogometnih ligah drugih držav – v našem primeru smo poiskati vse države iz katerih so nogometaši igrali v ligah teh 32 držav in vse države v katerih ligah so igrali nogometaši iz teh 32 držav. Dobimo 35 različnih držav. Relacijo *igralcev iz države  $x$  igra v državi  $y$*  lahko predstavimo z usmerjenim omrežjem kjer vrednost  $a$  na povezavi med državama  $x$  in  $y$  pomeni, da igra  $a$  igralcev iz države  $x$  v državi  $y$  (sliki [football.htm](#) in [football.kin](#))

## Analiza omrežja

- *Analiza na osnovi stopenj*
  - obstajajo države, ki samo izvažajo igralce: Jugoslavija, Romunija, Norveška, Nigerija...
  - obstajajo države, ki samo uvažajo igralce: Španija, Francija, Turčija, Grčija, GBR...
  - iz največ držav (18) uvažata igralce Španija (sledita Italija in Nemčija s po 16 državami).
  - v največ držav izvažata svoje igralce Nigerija (10).
- *Analiza na osnovi kazal in opisov*

V našem primeru bi za opise proglasili države, ki imajo močne nogometne lige, za kazala pa države, ki imajo dobre igralce. Dobimo:

- *Kazala*: Jugoslavija, Norveška, Argentina, Romunija, Hrvaška, Danska...
- *Kazala in opisi*: Nizozemska, Škotska, Brazilija, Avstrija...
- *Opisi*: Španija, GBR, Italija, Nemčija, Francija, Turčija...

Države so izpisane po padajoči vrednosti.

Če primerjamo rezultate na osnovi stopenj z rezultati na osnovi kazal in opisov, se postavljata dve vprašanji:

1. *Zakaj se Nigerija ne pojavi med najpomembnejšimi kazali, čeprav izvaža v največ držav?*

Razloga sta dva:

- Pri stopnjah smo prešteli samo v koliko držav izvaža dana država, ne pa tudi koliko igralcev, medtem ko se pri kazalih in opisih to upošteva.

V tem primeru je torej bolj smiselno izračunati stopnje, pri katerih upoštevamo tudi vrednosti na povezavah, kar izračunamo z množenjem matrike sosednosti z vektorjem iz samih enic.

V našem primeru pomeni vhodna stopnja glede na vrednosti število uvoženih igralcev, izhodna stopnja glede na vrednosti pa število izvoženih igralcev.

Po številu uvoženih igralcev je vrstni red:

Španija (54), GBR (50), Italija (42), Nemčija 39,...

po številu izvoženih pa:

Norveška (22), Jugoslavija (21), Nigerija (21), Kamerun (20),...

Opazimo, da Nigerija ni več na prvem mestu.

- Še pomembnejši razlog pa je, da Nigerija izvaža majši delež svojih igralcev v najboljše lige kot Jugoslavija ali Norveška.

V najmočnejše lige (Španija, Italija, Nemčija, Francija in Velika Britanija) tako izvaža

- Nigerija samo 10 od svojih 21 izvoženih igralcev (48%);
- Jugoslavija kar 19 od svojih 20 izvoženih igralcev (95%);
- Norveška pa 17 od svojih 22 izvoženih igralcev (77%).

2. *Zakaj se Argentina nahaja med prvimi tremi kazali, čeprav izvaža zelo malo igralcev (samo 13)?*

Argentina izvaža teh 13 igralcev samo v dve državi, ki pa imata najmočnejši ligi in sicer izvaža 4 igralce v Španijo, 9 pa v Italijo.

## Mere pomembnosti v Pajku

Kako izračunamo običajne vhodne in izhodne stopnje smo spoznali v poglavju o središčnosti.

Stopnje, pri katerih upoštevamo tudi vrednosti na povezavah izračunamo z ukazom:

Net/Vector/Summing up Values of Lines/Input za utežene vhodne, oziroma

Net/Vector/Summing up Values of Lines/Output za utežene izhodne stopnje. Rezultat je nov vektor.

---

Za računanje območja vpliva in Proximity Prestige pokličemo:

Net/Partitions/Domain/Input. Dobimo naslednje rezultate:

- razbitje s številom točk v območju vpliva
- vektor z deležem točk v območju vpliva
- vektor s povprečnimi oddaljenostmi od točk v območju vpliva

Proximity Prestige dobimo tako, da prvi dobljeni vektor delimo z drugim (Vectors/Divide First by Second) – delež točk v območju vpliva delimo s povprečno oddaljenostjo od točk v območju vpliva.

Opise in kazala izračunamo v izbiri

Net/Vector/Important Vertices/1-Mode: Hubs/Authorities

Dobimo še vprašanje koliko kazal in koliko opisov naj se označi v razbitju. Rezultati operacije so trije:

- Razbitje, kjer vrednost 1 (rumena) pomeni, da je točka dober opis, vrednost 2 (zelena) pomeni, da je točka dober opis in dobro kazalo, vrednost 3 (rdeča) pa pomeni, da je točka dobro kazalo.
- Vektor z utežmi za kazala (*Hubs Weights*), večja vrednost pomeni boljše kazalo.
- Vektor z utežmi za opise (*Authority Weights*), večja vrednost pomeni boljši opis.

Pozor: algoritem predpostavlja, da vrednosti na povezavah predstavljajo podobnosti (večja vrednost pomeni bolj pomembno izbiro). V primeru, da je ravno obratno, moramo vse vrednosti postaviti na 1, sicer bodo rezultati napačni. To lahko storimo tako, da shranimo omrežje v obliki dvojiške matrike (Pajek Binary Matrix), kjer se vse neničelne vrednosti zamenjajo z 1 in ga nato še enkrat preberemo.

## Primerjava mer pomembnosti

Obravnavali smo naslednje mere pomembnosti:

- vhodna stopnja (*input degree*);
- vhodna dostopnost (*input closeness*), oz. bližina (*proximity prestige*), če omrežje ni krepko povezano;
- vmesnost (*betweenness*);
- opisi (*authorities*).

Nekatere od teh količin merijo pomembnost na zelo podoben način, nekatere pa so si precej različne. Za ugotavljanje podobnosti rezultatov lahko uporabimo *Pearsonov ko-relacijski koeficient*, ki ga v Pajku najdemo v Vectors/Info, še prej pa moramo izbrati oba vektorja, ki ju želimo primerjati.

Na splošno ugotovimo, da so si mere zelo podobne, razen vmesnosti, ki meri pomembnost (oziroma bolj središčnost) na precej drugačen način kot ostale mere.



## Vaje

1. V omrežju vajaklik.net (od koga bi si izposodil zapiske) izračunajte bližino ter kazala in opise.
2. V usmerjenem omrežju sampson.net (naklonjenost med Sampsonovimi menihi) izračunajte vse mere pomembnosti, ki jih je mogoče (smiselno) izračunati. Narišite ustrezno sliko.
3. V usmerjenem omrežju wass749a.net (izvoz med 22 državami) izračunajte vse mere središčnosti in pomembnosti, ki jih je mogoče (smiselno) izračunati.
4. V omrežju football.net (izvoz nogometašev med 35 državami) izračunajte vhodne in izhodne stopnje (brez upoštevanja vrednosti in z upoštevanjem vrednosti) ter kazala in opise.