

PRIKAZ OMREŽIJ S SVG

Andrej Mrvar

Fakulteta za družbene vede
Univerza v Ljubljani

Vladimir Batagelj

Fakulteta za matematiko in fiziko
Univerza v Ljubljani

Povzetek

V prispevku so opisani načini prikaza omrežij s SVG (Scalable Vector Graphics), ki so vgrajeni v program Pajek. Opisane možnosti so ponazorjene na primeru povezanosti branosti slovenskih dnevnikov in revij ter mednarodnega 'uvoza in izvoza' nogometašev.

Summary

NETWORK VISUALIZATION USING SVG

In the article the options based on SVG language (Scalable Vector Graphics) for network visualization are described. All these options are implemented in program Pajek. Their usage is illustrated by two examples: readership connections among Slovenian journals and magazines and international 'import and export' of football players.

1. UVOD

SVG - Scalable Vector Graphics je označevalni jezik za omrežne opise slik, ki temelji na XML (eXtensible Markup Language). Njegovo definicijo najdemo na W3c/SVG [10]. Načrtan je tako, da je usklajen z drugimi omrežnimi standardi: HTML, XML Namespace, XLink, XPointer, CSS 2, DOM 1, ECMA/JavaScript, Java, Unicode, SMIL 1.0, ... Omogočal naj bi v omrežne sestavke vključevati slike določene z opisom njihove zgradbe.

Pri njegovem nastanku so sodelovala vsa najpomembnejša podjetja s področja slikovnih storitev: Adobe, Corel, Macromedia, IBM, Sun, Microsoft, HP...

SVG pozna tri vrste slikovnih sestavin: poti sestavljene iz ravnih črt in krivulj, slike in besedila. Nad njimi je mogoče izvajati vse običajne slikovne transformacije. Poleg tega lahko slike opremimo še z animacijo. Do sestavin slike v SVGju (DOM – Document Object Model) lahko dostopamo z uporabo jezika JavaScript in na ta način dinamično izbiramo / spreminjamo prikaz v okviru spletnega pregledovalnika.

Za pregledovanje slik v obliki SVG potrebujemo poseben vstavek (plug-in), ki je prosto dostopen na Adobovem spletišču [6].

Sliko v SVG lahko nadalje urejamo s pravtako prostim programom WebDraw [9].

SVG poznajo tudi nekateri drugi programi, kot na primer Mayura [7] in zadnje izdaje programov Adobe Illustrator ter Corel Draw.

Jezik SVG ima glede na druge oblike opisa slik precej prednosti:

- možnosti povečave izbranih delov slike;
- kratke datoteke z opisom slike;
- besedila so ohranjena kot nizi znakov – možnost iskanja;
- neodvisnost od izhodnih naprav in vrst računalnika;
- dober nadzor nad barvami;
- sodejnost (interaktivnost) in sprotno ustvarjanje slik;
- žive slike (animacija).

Za primer si v opisu 1 pogledjmo opis slike omrežja s slike 1.

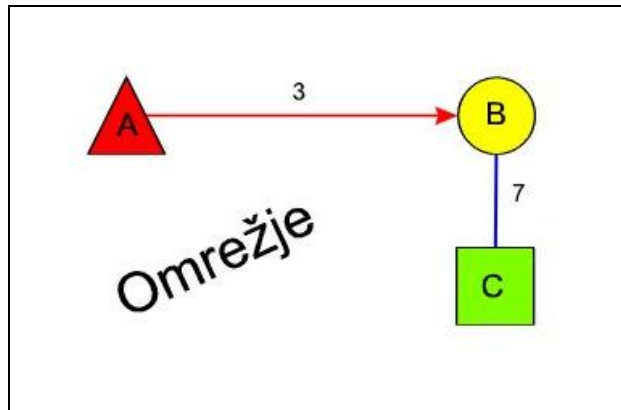
Več o SVG si preberite v Listu SIO [1].

2. GRAFI IN OMREŽJA

Osnova omrežja je graf. Graf $G = (V, L)$ je določen z množico točk V in množico (usmerjenih ali neusmerjenih) povezav L . Večkrat so na grafu poznane še dodatne funkcije, ki opisujejo lastnosti točk in/ali povezav. V tem primeru govorimo o omrežju $N = (V, L, F_V, F_L)$. Funkcija $f : V \rightarrow X$ iz F_V priredi posameznim točkam določen pomen. Podobno funkcija $g : L \rightarrow Y$ iz F_L priredi pomen povezavam.

Opis 1: Opis slike v SVG.

```
<?xml version="1.0" encoding="UTF-8"?>
<!-- CreationDate: 02-08-2001 -->
<!DOCTYPE svg [
  <!ENTITY pis "font-family:Helvetica;text-anchor:middle;">
  <!ENTITY vre "font-size:15;">
  <!ENTITY ozn "font-size:20;">
  <!ENTITY toc "stroke-width:1;stroke:black;">
  <!ENTITY pov "stroke-width:2;">
]>
<svg xml:space="preserve" width="470" height="283"
  viewBox="0 0 470 283">
  <rect width="470" height="283"
    style="fill:white;stroke:rgb(127,127,127);"/>
  <text x="192.52" y="78.30" style="&pis;&vre;">3</text>
<!-- Arc:      1      2 -->
  <g transform="translate(284.33 88.30)">
    <path d="M 0 0 L -2 5 L 10 0 L -2 -5 z"
      style="fill:red;stroke:red;"/></g>
  <line x1="89.21" y1="88.30" x2="284.33" y2="88.30"
    style="&pov;stroke:red;"></line>
  <text x="329.39" y="141.58" style="&pis;&vre;">7</text>
<!-- Edge:      2      3 -->
  <line x1="319.83" y1="88.30" x2="318.94" y2="194.70"
    style="&pov;stroke:blue;"></line>
<!-- Vertex:      1 -->
  <g transform="translate(89.21 88.30)">
    <path d="M 24 24 L 0 -24 L -24 24 z"
      style="&toc;fill:red;"/></g>
  <text x="89.21" y="94.96" style="&pis;&ozn;">A</text>
<!-- Vertex:      2 -->
  <ellipse cx="319.83" cy="88.30" rx="24" ry="24"
    style="&toc;fill:yellow;"/>
  <text x="319.83" y="94.96" style="&pis;&ozn;">B</text>
<!-- Vertex:      3 -->
  <g transform="translate(318.94 194.70)">
    <path d="M -24 -24 L 24 -24 L 24 24 L -24 24 z"
      style="&toc;fill:rgb(127,255,0);"/></g>
  <text x="318.94" y="201.37" style="&pis;&ozn;">C</text>
  <g transform="translate(89.21 215) rotate(335)">
    <text style="font-family:Helvetica;font-size:35;">
      Omre&#382;je</text></g>
</svg>
```



Slika 1: Slika omrežja.

Pajek [2] je programski paket, za Windows 32, ki omogoča analizo *velikih omrežij* (omrežij z več tisoč točkami). Program je prosto dostopen na spletu [8].

3. PRIKAZI OMREŽIJ V SVG

Opis slike v SVG lahko razdelimo na dele. Ena od lastnosti, ki jo lahko pripišemo posameznemu delu je tudi vidnost (*visibility*). Z JavaScriptom lahko vidnost spreminjamo (vključujemo / izključujemo).

V zadnjem času so bile v program Pajek dodane tudi nekatere možnosti za pripravo slik omrežij v SVG, ki izkoriščajo možnosti, ki jih jezik SVG omogoča (izbira delov slik glede na različne vrste točk in povezav). V črnobeli in tiskani različici članka se večine teh možnosti ne da prikazati. Lahko jih le nakažemo z različnimi pogledi na sliko. Ustrezne spletne strani si lahko ogledate na <http://vlado.fmf.uni-lj.si/pub/networks/pajek/examples/>

Ena od takih možnosti so *prelivi* – zvezno prehajanje barve ozadja iz ene barve v drugo (*linear* ali *radial gradients*).

Pri izpisu v SVG zaenkrat Pajek ponuja naslednje možnosti:

- Najenostavnejši je izpis omrežja z vsemi točkami in povezavami brez možnosti izbir posameznih delov slike.
- Izpis z možnostjo vključevanja in izključevanja oznak točk in povezav, prikaz samo usmerjenih ali samo neusmerjenih povezav.
- Izpis z upoštevanjem trenutnega razbitja (točke razdelimo v skupine glede na neke lastnosti) ali dveh razbitij. V primeru, da dveh razbitij ne izberemo, določa trenutno razbitje tako skupine kot tudi barve točk. Če pa izberemo dve razbitji, določa prvo razbitje skupine drugo pa barve točk. Pri izpisu glede na podano razbitje poznamo tri možnosti:
 - Možnost vključevanja in izključevanja posameznih skupin iz razbitja in povezav med skupinami.

- Možnost vključevanja in izključevanja posameznih skupin iz razbitja. Povezave med skupinami se rišejo kot polpovezave.
- Prikazovanje gnezdenih razbitij – kadarkoli vključimo izbrano skupino, se vključijo tudi vsi višje skupine in izključijo vse nižje skupine od izbrane (primerno za prikazovanje jeder).
- Izpis z upoštevanjem vrednosti na povezavah. Vnesemo lahko meje razredov ali število razredov v katero želimo razvrstiti vse povezave. Glede na dobljene meje dobimo podmnožice povezav (in točk, ki so njihova krajišča). Dobljene podmnožice lahko prikazujemo ali skrivamo v okviru spletnega pregledovalnika. Na voljo je več možnosti:
 - Možnost prikazovanja in skrivanja povezav z izbranimi vrednostmi in pripadajočih točk.
 - Možnost prikazovanja in skrivanja povezav z izbranimi ali višjimi vrednostmi in pripadajočih točk (gnezdenje).

Pri tem se za slikovno predstavitev vrednosti povezav lahko uporabijo:

- različne barve povezav,
- različne sivine povezav,
- različne debeline povezav.

4. PRIMER: SLOVENSKI DNEVNIKI IN REVIIJE

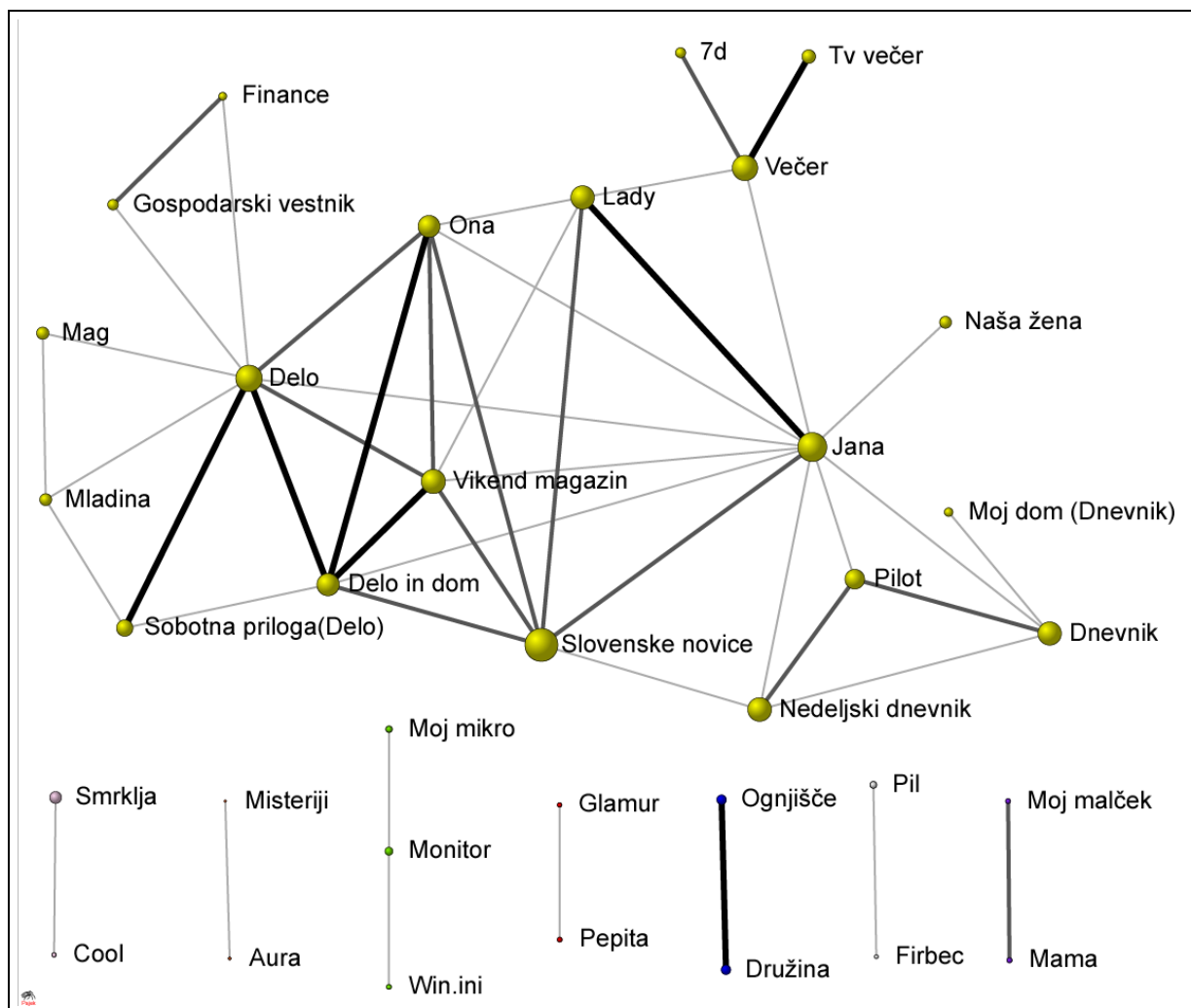
V letih 1999 in 2000 je v anketah Cati Centra [4] več kot 100000 anketirancev odgovarjalo na vprašanje, katere dnevnike in revije berejo. Našteli so skupaj 124 različnih revij. Dobljene podatke lahko predstavimo kot dvovrstno omrežje:

	Delo	Dnevnik	Sl.novice	...
Anketiranec1	X		X	...
Anketiranec2		X		...
Anketiranec3	X			...
.....

Iz dvovrstnega omrežja bralec / revija smo zgenerirali običajno omrežje, kjer so enote revije. Neusmerjena povezava z vrednostjo a med revijama pomeni, da obe reviji hkrati bere a anketirancev. Zanka pri posamezni reviji pa pomeni skupno število anketirancev, ki berejo to revijo.

Dobimo matriko A :

	Delo	Dnevnik	Sl.novice	...
Delo	20714	3219	4214	...
Dnevnik		15992	3642	...
Sl. novice			31997	...
.....



Slika 2: Slovenski dnevniki in revije.

Zaradi velikih razlik v številu bralcev posameznih revij je iz osnovnih podatkov težko sklepati na povezanost med revijami, zato je potrebno omrežje najprej normalizirati. Obstaja več smiselnih normalizacij. Na primer:

$$\text{Geo}_{ij} = \frac{A_{ij}}{\sqrt{A_{ii}A_{jj}}}$$

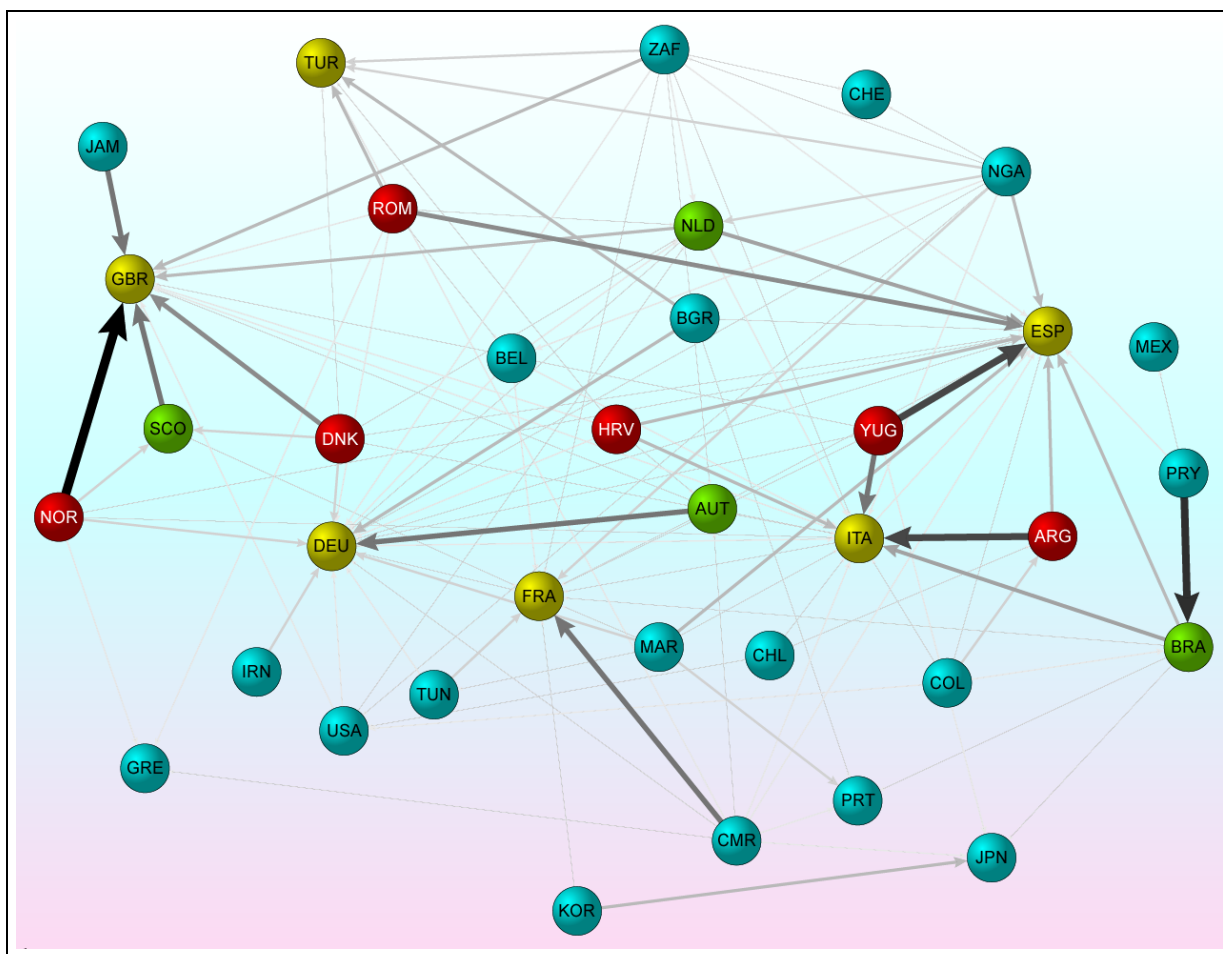
$$\text{MinDir}_{ij} = \frac{A_{ij}}{A_{ii}}, \text{ če je } A_{ii} \leq A_{jj}, \text{ in } 0 \text{ sicer}$$

Na sliki 2 so prikazani dnevniki in revije, med katerimi je povezanost Geo vsaj 0.2.

5. PRIMER: NOGOMETAŠI

V nadaljevanju bomo za analizo omrežja uporabili kazala in opise [3]:

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



Slika 3: Uvoz/izvoz nogometašev.

dobrih kazal.

Na svetovnem nogometnem prvenstvu v Parizu leta 1998 je sodelovalo 22 nogometnih reprezentanc. Ob tej priliki je Lothar Krempel zbral podatke o 'uvozu in izvozu' nogometašev [5]. Nogometaši pogosto igrajo v nogometnih ligah drugih držav – v našem primeru smo poiskali vse države iz katerih so nogometaši igrali v ligah teh 22 držav in vse države v katerih ligah so igrali nogometaši iz teh 22 držav. Dobimo 35 različnih držav. Relacijo *igrallec 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 (slika 3).

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) uvažajo igralce Španija (sledita Italija in Nemčija s po 16 državami).
- v največ držav izvažajo 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. Zato moramo 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 manjši delež svojih igralcev v najboljših ligah 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.

6. ZAKLJUČEK

Postopke prikazov omrežij v obliki SVG še razvijamo in sproti vključujemo v program Pajek. Tekoče stanje in povezave na vire podatkov ter druge zanimive informacije lahko najdete na predstavitveni strani programa Pajek [8].

Viri

- [1] BATAGELJ, V. (2000): SVG – Scalable Vector Graphics / Spletni opisi slik. *List SIO* 1.
<http://sio.edus.si/list/1/svg/>
- [2] BATAGELJ, V., MRVAR, A. (1998): Pajek – A Program for Large Network Analysis. *Connections* 21 (2), 47-57.
- [3] ZAVERŠNIK M., BATAGELJ V. (2000): Pomembni deli obsežnih omrežij. *Dnevi Slovenske informatike, Zbornik posvetovanja*, 149-156.
- [4] Cati Center – vhodna stran:
<http://www.cati.si>
- [5] Lothar Krempel, Max-Planck-Institut fuer Gesellschaftsforschung – vhodna stran:
<http://www.mpi-fg-koeln.mpg.de>
- [6] Plug-in za SVG:
<http://www.adobe.com/svg/viewer/install/>
- [7] Risalni program Mayura:
<http://www.mayura.com/>
- [8] Programski paket **Pajek**:
<http://vlado.fmf.uni-lj.si/pub/networks/pajek/>
- [9] Risalni program WebDraw:
<http://www.jasc.com/webdraw.asp>
- [10] Scalable Vector Graphics (SVG) Specification:
<http://www.w3.org/TR/SVG/>