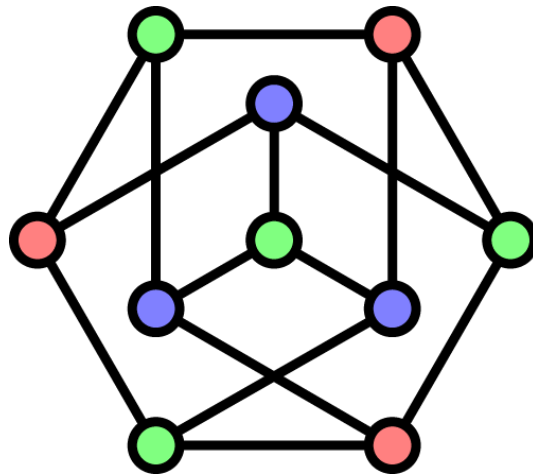


IZBRANA POGLAVJA IZ DISKRETNE MATEMATIKE

ZBORNİK SEMINARSKIH NALOG IZ DISKRETNE MATEMATIKE



Matjaž Krnc, Riste Škrekovski

Junij 2016 – Verzija 1.0

CIP – Kataložni zapis o publikaciji
Narodna in univerzitetna knjižnica, Ljubljana

519.1(082)(0.034.2)

IZBRANA poglavja iz diskretne matematike [Elektronski vir] :
zbornik seminarskih nalog iz diskretne matematike / avtorji Ur-
ban Marovt ... [et al.] ; [urednika] Matjaž Krnc, Riste Škrekovski.
- Verzija 1.0. - El. knjiga. - Ljubljana : samozal. M. Krnc, 2016.

Način dostopa (URL): [https://www.scribd.com/document/
321997378/UDM-2016](https://www.scribd.com/document/321997378/UDM-2016)

ISBN 978-961-92887-9-5 (pdf)

1. Marovt, Urban 2. Krnc, Matjaž, 1987- 286126336

Seminarske naloge iz tega zbornika so nastale na Fakulteti za matematiko in fiziko, v akademskem letu 2015-2016. Zbornik je pedagoške narave, namen je bralcu ponuditi nekatere zanimive teme iz diskretne matematike v slovenskem jeziku. Vsebina poglavij iz tega zbornika sledi priznanim objavljenim člankom ter knjigam v angleškem jeziku. Nekatere seminarske naloge so z namenom večje preglednosti urejene in/ali združene. Morebitne napake sporočite na matjaz.krnc@gmail.com.

Za izdelavo seminarskih nalog se zahvaljujemo študentom:
Urban Marovt, Anja Petković, Rok Fortuna, Vid Smrke, Larisa Gostenčnik, Vesna Iršič, Miha Eleršič, Rebeka Merhar, Gregor Rus, Sven Cerk, Bor Juros.

Izbrana poglavja iz diskretne matematike

Zbornik seminarskih nalog iz diskretne matematike

Urednika: Matjaž Krnc, Riste Škrekovski

Avtorji: Urban Marovt, Anja Petković,
Rok Fortuna, Vid Smrke, Larisa
Gostenčnik, Vesna Iršič, Miha
Eleršič, Rebeka Merhar, Gregor
Rus, Sven Cerk, Bor Juros

Samozaložba in oblikovanje: Matjaž Krnc

Fotografija na naslovnici: *Durer graph* by Koko90 - Own
work. Licensed under CC BY-
SA 3.0

ISBN: 978-961-92887-9-5

Ljubljana, Junij 2016

KAZALO

1	ZNAČILNOSTI POVEZANOSTI VOZLIŠČ GLEDE NA NJIHOVO STOPNJO	7
1.1	Uvod	7
1.2	Definicija matrike korelacije stopenj vozlišč	8
1.3	Definicija funkcije korelacije stopenj vozlišč	10
1.4	Definicija koeficienta korelacije stopenj vozlišč	12
1.5	Stukturne zgornje meje	13
1.6	Zaključek	16
2	KNESERJEVI GRAFI	21
2.1	Osnovne definicije	21
2.2	Osnovne lastnosti Kneserjevih grafov	22
2.3	Delno barvanje	23
2.4	Delna klika	25
2.5	Dualnost	26
2.6	Ciklični intervalni grafi	28
2.7	Erdős-Ko-Rado izrek	29
2.8	Homomorfizmi na Kneserjevih grafih	30
3	MERE CENTRALNOSTI	35
3.1	Uvod	35
3.2	Preproste mere centralnosti	35
3.3	Spektralna centralnost	36
3.4	Katzova centralnost	39
3.5	PageRank	40
3.6	Mere centralnosti v praksi	40
3.6.1	Omrežje sodelovanja igralcev	42
4	STRUKTURA SLUČAJNIH GRAFOV V ERDOS-RÉNYIJEVEM MODELU	47
4.1	Uvod	47
4.2	Erdős-Rényijev model	48
4.2.1	Procesi razvejanja	49
4.2.2	Povezava procesov razvejanja in slučajnih grafov	54
4.2.3	Zgled – epidemološki model	55
4.2.4	Zaključek	55
5	PROBLEM ŠTIRIH BARV	59
5.1	Uvod	59
5.1.1	Problem štirih barv	59
5.1.2	Prevedba na problem v teorijo grafov	60

5.2	Problem petih barv	60
5.3	Neizogibne konfiguracije	63
5.4	Reducibilnost	64
5.5	Problem štirih barv	66
5.6	Zaključek	67
6	HOMOMORFIZMI GRAFOV	71
6.1	Osnovni pojmi	71
6.2	Relacija \rightarrow	74
6.3	Retrakti	75
6.4	Povezava z barvanji grafov	76
6.5	Jedra	78
6.6	Zaključek	82
7	STISKANJE GRAFOV	85
7.1	Uvod	85
7.2	Stisnjena predstavitev grafa	86
7.3	Dolžina minimalne predstavitve	86
7.4	Aproksimacija predstavitve grafa	87
7.5	Algoritem	88
7.5.1	Ideja	88
7.5.2	Izračun najboljšega para za združevanje	88
7.5.3	Realizacija	89
7.5.4	Algoritem RANDOMIZED	89
7.6	Rezultati	91
7.6.1	Učinkovitost stiskanja	91
7.6.2	Hitrost stiskanja	91
7.6.3	Primerjava z drugimi algoritmi stiskanja	91
7.7	Zaključek	94
8	SNARKI	97
8.1	Osnovne definicije	98
8.2	Issacovi snarki	100
8.2.1	Točkast produkt	100
8.2.2	Zvezdast produkt	102
8.2.3	Cvetlični snarki	104
8.2.4	Dvojni zvezdast snark	105
8.3	k -snarki	105
8.4	Zaključek	105
9	HIPOHAMILTONOVI GRAFI	109
9.1	Osnovne definicije	109
9.2	Raziskovanje hipohamiltonovih grafov	114
9.3	Zaključek	116
10	LIMITE GRAFOV	119
10.1	Limite grafov	119
10.2	Gostota homomorfizmov in konvergenca grafov	120

10.3	Grafoni	122
10.4	Prostor grafonov	124
10.5	Preprost primer	125
10.6	Zaključek	126
11	JOHNSONOVİ IN J -GRAFI	129
11.1	Uvod	129
11.2	J -grafi	130
11.2.1	Lastnosti J -grafov	130
11.2.2	Izomorfizem J -grafov	131
11.2.3	Vozliščna tranzitivnost J -grafov	132
11.3	Johnsonovi grafi	133
11.3.1	Lastnosti Johnsonovih grafov	133
11.3.2	Premer Johnsonovih grafov	133
11.3.3	Hamiltonov cikel v Johnsonovih grafih	134
11.4	Zaključek	135

ZNAČILNOSTI POVEZANOSTI VOZLIŠČ GLEDE NA NJIHOVO STOPNJO

Urban Marovt

1.1 UVOD

Pri analizi omrežij poskušamo odkriti čim več karakteristik, ki bi čim bolj natančno opisale značilnosti posameznih grafov, saj želimo vsako omrežje predstaviti z določenim modelom oz. ga preslikati v model z določenimi značilnostmi, da bi lažje razumeli njegovo strukturo in obnašanje. Velikokrat nas zanima, v kakšnem odnosu so vozlišča, ki so med seboj povezana. Predvsem nas zanima, ali so si ta med seboj podobna ali različna. Pravimo, da je omrežje povezano podobno, če se povezujejo vozlišča, ki so si po določenih karakteristikah podobna (angl. assortative network). V nasprotnem primeru, če je za omrežje značilno, da se večinoma povezujejo različno karakterizirana vozlišča, pravimo, da je omrežje povezano različno (angl. disassortative network).

Karakteristiko povezovanja je prvi predstavil Mark Newman v iz leta 2002 [1], kjer je za podobnostno karakteristiko vzel stopnjo vozlišč, lahko pa vzamemo tudi vmestnost vozlišč, uteženost vozlišč, število vozlišč, ki jih lahko dosežemo s k koraki, itd.

Če za primer in motivacijo vzamemo omrežje poznanstev med ljudmi in pogledamo skupino zelo znanih ljudi, kamor spadajo slavni, voditelji podjetij, politiki, lahko predpostavimo, da bi se le-ti lahko družili s komerkoli v omrežju, če bi bilo omrežje socialnih povezav sestavljeno naključno. Vendar hitro opazimo, da je tendenca v socialnih omrežjih takšna, da se bolj znani ljudje velikokrat družijo med seboj in ne obratno. Ti ljudje predstavljajo v omrežju poznanih ljudi središča (angl. hubs) in lahko ugotovimo, da je za socialna omrežja značilno povezovanje vozlišč s podobno stopnjo [4].

Za razliko od socialnih omrežji pa imamo na drugi strani tudi takšna omrežja, kjer je tendenca ravno obratna in jim pravimo središčno izogibajoča omrežja (angl. hubs avoiding hubs). Takšna struktura je predvsem značilna za biološka in tehnološka vozlišča [6].

V nadaljevanju bomo predstavili tri možne mere za korelacijo stopenj vozlišč, ki so se razvile skozi čas in se med seboj razlikujejo predvsem po preglednosti in primerljivosti. Najprej bomo izpostavili matriko korelacije stopenj vozlišč, zatem bomo definirali funkcijo korelacije stopenj vozlišč, nato pa še koeficient korelacije stopenj vozlišč. Značilnosti vsake mere bomo razložili na primerih in poudarili njihove prednosti in slabosti. Na koncu bomo izpostavili še anomalijo strukturne zgornje meje, ki govori o omejitvi značilnosti podobnostnega povezovanja in predstavili omrežja, za katera je strukturna zgornja meja najbolj značilna.

1.2 DEFINICIJA MATRIKE KORELACIJE STOPENJ VOZLIŠČ

Za definicijo modela predpostavimo, da imamo neusmerjen enostaven graf G , ki je sestavljen iz N vozlišč in M povezav, s porazdelitvijo stopenj vozlišč p_k , kjer je p_k verjetnost, da bo naključno izbrano vozlišče stopnje k .

Recimo, da pridemo do posameznega vozlišča preko naključno izbrane povezave. Opazimo, da stopnja vozlišča ni porazdeljena tako kot p_k , saj je večja verjetnost, da končamo v vozlišču z višjo stopnjo, ker lahko do teh vozlišč pridemo po več različnih poteh, natančneje k . Od tod sledi, da je stopnja vozlišča, do katerega pridemo na ta način enaka porazdelitvi $k p_k$ in ne samo p_k . Na koncu to dobljeno porazdelitev normaliziramo in dobimo porazdelitev q_k , ki je enaka

$$q_k = \frac{k p_k}{\sum_j j p_j}. \quad (1)$$

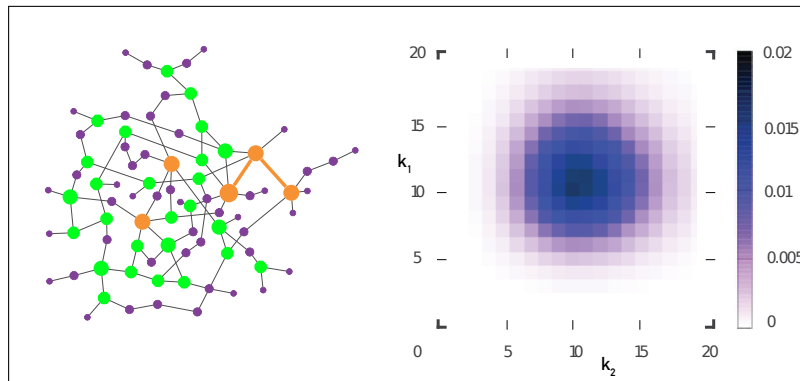
Sedaj definirajmo vrednosti matrike korelacije stopenj vozlišč e_{jk} , ki je naša prva mera za podobnostno karakteristiko omrežja. Vrednosti predstavljajo verjetnosti, da povežemo dve vozlišči s stopnjama j in k . Ta matrika je simetrična na neusmerjenih grafih in veljata naslednji dve enačbi

$$\sum_{jk} e_{jk} = 1, \quad (2)$$

$$\sum_j e_{jk} = q_k. \quad (3)$$

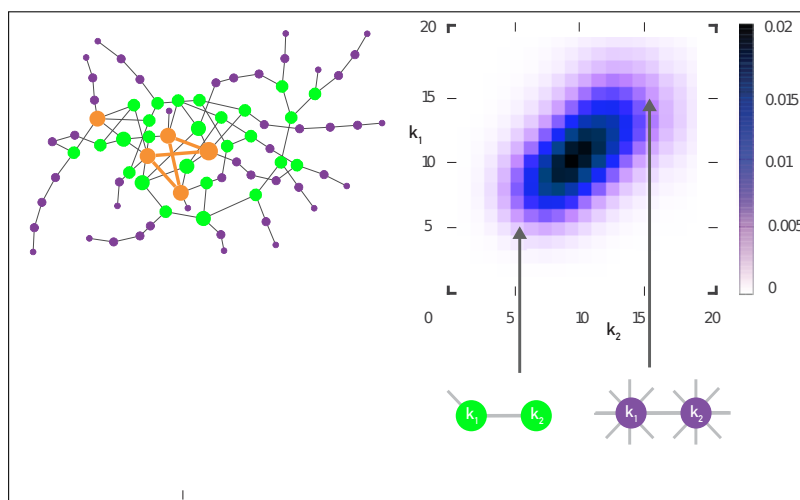
Na slikah 1, 2 in 3 imamo omrežja z enakim številom vozlišč in enako porazdelitvijo p_k , vendar različno karakteristiko med povezanimi vozlišči.

Če je omrežje nevtravno, za graf ni značilno niti, da se povezujejo podobna niti različna vozlišča in velja, da razporeditev povezovanja glede na stopnjo vozlišč e_{jk} zavzame vrednosti $q_j q_k$. Na sliki 1 vidimo primer nevtralnega omrežja in njegovo pripadajočo matriko stopenjske korelacije. Opazimo, da ni nobenega trenda med povezovanjem vozlišč visoke oz. nizke stopnje.



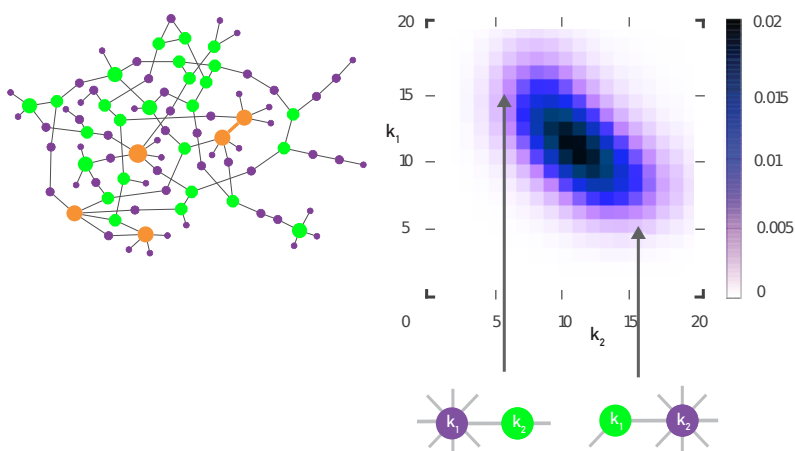
Slika 1: Nevtralno omrežje glede na povezovanje glede na stopnjo vozlišč

V nasprotnem primeru pa porazdelitev e_{jk} odstopa od prej definirane naključne porazdelitve in lahko opazimo določene omenjene karakteristike. Na sliki 2 imamo primer omrežja, ki ima karakteristiko podobnostnega povezovanja vozlišč. Opazimo lahko, da je porazdelitev zgoščena na vodilni diagonali matrike, kar predstavlja trend povezovanja vozlišč z podobno stopnjo.



Slika 2: Omrežje, ki ima karakteristiko podobnostnega povezovanja

Na sliki 3 imamo še primer omrežja, kjer je porazdelitev zgoščena po nasprotni diagonali, kar predstavlja povezovanje nasprotno karakteriziranih vozlišč glede na njihovo stopnjo.



Slika 3: Omrežje, ki ima karakteristiko različnostnega povezovanja

Mera, ki jo predstavlja korelacijska matrika, ni najboljša izbira, saj je samo odčitavanje nenatančno, težko primerljivo, matrika pa lahko vsebuje veliko vrednosti, natančneje $\frac{k_{\max}^2}{2}$.

1.3 DEFINICIJA FUNKCIJE KORELACIJE STOPENJ VOZLIŠČ

V primerjavi z zgoraj opisano matriko korelacije stopenj vozlišč 1.2, je potrebno poiskati bolj kompaktno in merljivo mero. En način, kako lahko izmerimo razmerje med povezanimi vozlišči je, da za vsako vozlišče izračunamo povprečno stopnjo vozlišč s katerim je povezan. To izračunamo z naslednjo enačbo

$$k_{nn}(v_i) = \frac{1}{k_i} \sum_{j=1}^N A_{ij} k_j, \quad (4)$$

kjer je k_i , stopnja specifičnega vozlišča v_i in A matrika sosednosti. Tako lahko to funkcijo razširimo na vsa vozlišča enake stopnje in dobimo funkcijo stopenjske korelacije

$$k_{nn}(k) = \sum_{k'} k' P(k'|k), \quad (5)$$

kjer je $P(k'|k)$ pogojna verjetnost, da lahko iz vozlišča s stopnjo k pridemo do vozlišča s stopnjo k' . Tako je $k_{nn}(k)$ povprečna stopnja sosedov vseh vozlišč stopnje k . Sedaj analiziramo podobnost povezanih vozlišč preko razmerja njihove stopnje in vrednosti funkcije stopenjske korelacije.

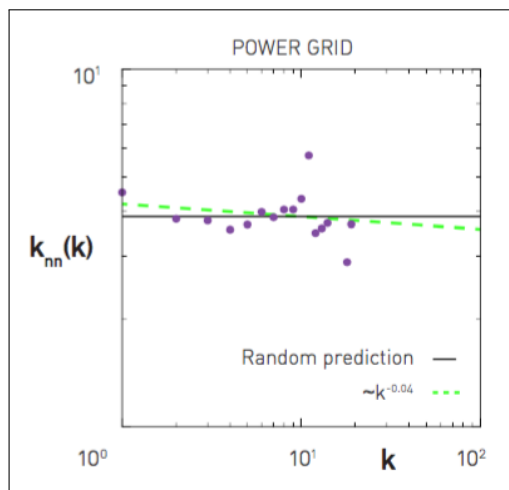
Najprej pogledamo primer, ko imamo nevtralno omrežje. Izračunamo vrednost prej definirane pogojne verjetnosti

$$P(k'|k) = \frac{e_{kk'}}{\sum_j e_{jk}} = \frac{e_{kk'}}{q_k} = \frac{q_{k'}q_k}{q_k} = q_{k'}. \quad (6)$$

Tako lahko izrazimo funkcijo $k_{nn}(k)$ kot

$$k_{nn}(k) = \sum_{k'} k' q_{k'} = \sum_{k'} k' \frac{k' p_{k'}}{\sum_i k_i p_i}. \quad (7)$$

Zgornji izraz je popolnoma neodvisen od parametra funkcije k , zato je korelacija med k in vrednostjo funkcije $k_{nn}(k)$ konstantna. Na sliki 4 imamo primer grafa funkcije $k_{nn}(k)$ neutralnega omrežja na logaritm-skih skalah. Opazimo, da so vrednosti funkcije $k_{nn}(k)$ nabrane okoli vodoravne črte.



Slika 4: Graf funkcije $k_{nn}(k)$ za nevtralno omrežje.

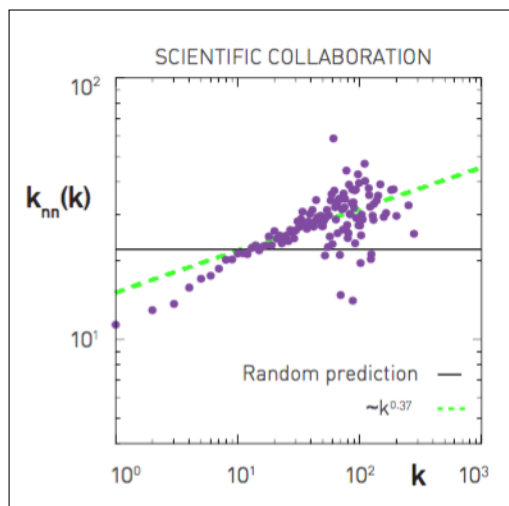
Za omrežja, ki imajo karakteristiko podobnostnega povezovanja, vemo, da z večanjem stopnje opazovanja vozlišča, narašča tudi vrednost funkcije $k_{nn}(k)$. Prikaz tega vidimo na sliki 5.

Nasprotno pa velja za omrežja, kjer se povezujejo vozlišča različnih stopenj. Za njih velja da vrednost funkcije $k_{nn}(k)$ pada, z naraščanjem vrednosti k . Rezultat tega vidimo na sliki 6.

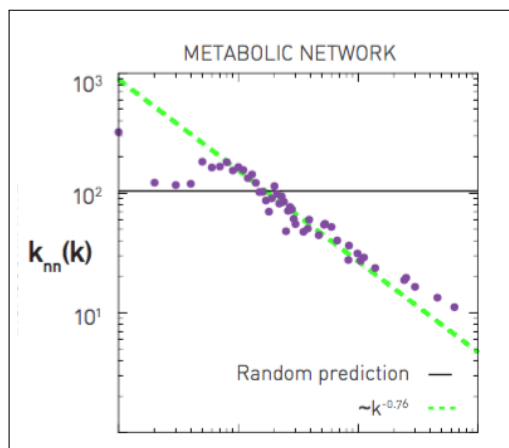
Iz slik 4, 5, 6 izluščimo, da lahko določimo aproksimacijsko vrednost funkcije $k_{nn}(k)$ s pomočjo naslednje enačbe

$$k_{nn}(k) = ak^\mu. \quad (8)$$

Na slikah 4, 5, 6 je graf te funkcije označen z zeleno črtkano črto. Če je parameter $\mu > 0$ potem ima omrežje značilnost podobnostnega povezovanja, če velja $\mu < 0$, potem ima omrežje značilnost različnega povezovanja. V primeru, da je $\mu = 0$ pa imamo nevtralno omrežje, glede na opazovano karakteristiko povezovanja.



Slika 5: Graf funkcije $k_{nn}(k)$ za omrežje, ki ima karakteristiko podobnostnega povezovanja



Slika 6: Graf funkcije $k_{nn}(k)$ za omrežje, ki ima karakteristiko različnega povezovanja

1.4 DEFINICIJA KOEFICIENTA KORELACIJE STOPENJ VOZLIŠČ

V nasprotju s predpostavko iz enačbe 8, je Mark Newmann v [1] predstavil koeficient korelacije stopenj vozlišč, ki ne predpostavi da sta vrednosti k in $k_{nn}(k)$ odvisni polinomske, ampak linearno.

Če je omrežje nevtravno velja enako kot prej, da za porazdelitev e_{jk} velja $e_{jk} = q_j q_k$. V nasprotnem, pa se porazdelitev razlikuje od omenjenega produkta in lahko izračunamo vrednost značilnosti povezovanja podobnih oz. različnih vozlišč s pomočjo funkcije $f(G) = \sum_{ij} ij(e_{ij} - q_i q_j)$. Ta funkcija zavzame vrednosti blizu 0 za omrežja brez opisanih značilnosti, pozitivne vrednosti za tista, kjer opazimo karakteristiko povezovanja podobnih vozlišč in negativne, če se v omrežju pogosteje povezujejo vozlišča različnih stopenj. Funkcija zavzame naj-

večje vrednosti za grafe, kjer imajo vsa vozlišča enako stopnjo (npr. polni grafi, regularni grafi). V teh primerih za porazdelitev e_{ij} velja $e_{ij} = q_j \delta_{ij}$, kjer je δ_{ij} Kroneckerjeva delta. Ta vrednost je enaka vrednosti variance porazdelitve q_k , ki je $\sigma^2 = \sum_k k^2 q_k - (\sum_k k q_k)^2$. Tako dobimo končno enačbo za naš koeficient r

$$r = \sum_{ij} \frac{ij(e_{ij} - q_i q_j)}{\sigma^2}. \quad (9)$$

Vrednost r Pearsonov korelacijski koeficient vozlišč na obeh straneh povezav in zavzame vrednosti med -1 in 1 .

V [2] je Mark Newman svojo definicijo podobnostnega koeficienta prenesel na realna omrežja in predstavil nove modele omrežij. Na sliki 7 vidimo nekaj primerov omrežji z njihovo velikostjo n in izračunanim koeficientom korelacije stopenj vozlišč r .

	network	n	r
real-world networks	physics coauthorship ^a	52 909	0.363
	biology coauthorship ^a	1 520 251	0.127
	mathematics coauthorship ^b	253 339	0.120
	film actor collaborations ^c	449 913	0.208
	company directors ^d	7 673	0.276
	Internet ^e	10 697	-0.189
	World-Wide Web ^f	269 504	-0.065
	protein interactions ^g	2 115	-0.156
	neural network ^h	307	-0.163
	food web ⁱ	92	-0.276
models	random graph ^u		0
	Callaway <i>et al.</i> ^v		$\delta/(1+2\delta)$
	Barabási and Albert ^w		0

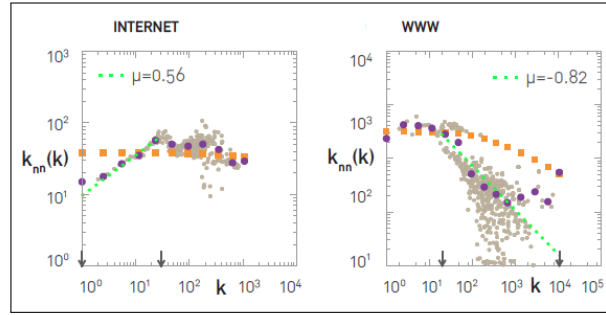
Slika 7: Tabela primerov omrežji z njihovo velikostjo n in koeficientom r .

1.5 STUKTURNE ZGORNJE MEJE

Čeprav je lahko intuitivno, da se v omrežjih kot sta Internet in svetovno medmrežje (angl. World Wide Web), povezujejo vozlišča podobnih stopenj, lahko v tabeli na sliki 7 najdemo vrednosti koeficienta r in opazimo, da v njih prevlada trend povezovanja vozlišč različnih stopenj.

V [3] so avtorji analizirali karakteristiko podobnostnega povezovanja v omrežjih scale-free. Predstavili so omejitev strukturne zgornje meje in dokazali njeno vrednost glede na število in povprečno stopnjo vozlišč.

Na sliki 8 vidimo grafa funkcije $k_{nn}(k)$ za oba omenjena primera in opazimo, da se vrednosti drugače obnašajo pri manjših kot pa pri velikih stopnjah.



Slika 8: Grafa funkcije $k_{nn}(k)$ za internetno omrežje in omrežje svetovnega medmrežja.

Do sedaj smo vselej opazovali enostavne grafe, kar pomeni, da ne vsebujejo večkratnih povezav med vozlišči in povezav samih vase. Čeprav v svetovnem medmrežju ena stran vsebuje več linkov do druge strani, imamo v opazovanem omrežju zgolj eno povezavo med vozliščema.

V enostavnih grafih obstaja konflikt med atributom scale-free in korelacijo stopenj vozlišč [5]. Za primer vzemimo graf G z $N = 300$ in $M = 450$, kjer imamo dve vozlišči največjih stopenj $k = 55$ in $k' = 46$. Iz porazdelitve $e_{kk'}$ dobimo pričakovano število povezav med opazovanima vozliščema

$$E_{kk'} = e_{kk'} N \sum_i p_i. \quad (10)$$

Predpostavimo, da je naše omrežje nevtralno glede na korelacijo stopenj vozlišč. Tako je pričakovano število povezav enako

$$E_{kk'} = \frac{k p_k k' p_{k'}}{\sum_i i p_i} N = \frac{55 \cdot 46}{300 \cdot 300} 300 = 2.8. \quad (11)$$

Vidimo, da bi med vozliščema pričakovali približno dve ali tri povezave, zaradi predpostavke o preprostem grafu, pa je lahko med njima zgolj ena povezava. Takšnemu pojavu pravimo omejitev strukturne zgornje meje.

V primeru ko imamo vozlišči z nizkima stopnjama k in k' nimamo težav, saj je tudi pričakovano število povezav manj od ena. Torej se težava pojavi zgolj takrat, ko opazujemo vozlišča katerih stopnja presega določeno mejo k_s , za katere bi pričakovali več kot eno povezavo.

Če želimo analizirati strukturne zgornje meje moramo najprej ugotoviti, v katerih grafih se sploh pojavi. Za primerjavo definirajmo strukturno zgornjo mejo k_s in naravno zgornjo mejo k_{max} , ki predstavlja pričakovano zgornjo mejo stopnje vozlišč v omrežju.

Trditev 1.1. *Strukturna zgornja meja k_s narašča kot $(\langle k \rangle N)^{\frac{1}{2}}$.*

Dokaz. Na začetku definirajmo $r_{kk'} = \frac{E_{kk'}}{m_{kk'}}$, kjer je $E_{kk'}$ število povezav med vozliščema stopnje k in k' za $k \neq k'$ in dvakrat število povezav med vozliščema za $k = k'$. V imenovalcu nastopa vrednost $m_{kk'} = \min\{kN_k, k'N_{k'}, N_k N_{k'}\}$, ki je največje možno število povezav med obema skupinama. Tako lahko naš koeficient $r_{kk'}$ malo razpišemo in dobimo

$$r_{kk'} = \frac{E_{kk'}}{m_{kk'}} = \frac{\langle k \rangle e_{kk'}}{\min\{kp_k, k'p_{k'}, Np_k p_{k'}\}}. \quad (12)$$

Ker je vrednost $m_{kk'}$ največja vrednost za $E_{kk'}$, potem je vrednost koeficienta $r_{kk'}$ vedno manjša ali enaka 1. Seveda pa to velja samo za enostavne grafe. Zato smo koeficient k_s definirali kot rešitev enačbe $r_{k_s k_s} = 1$.

Opazimo, da ko velja $k > Np_k$ in $k' > Np_{k'}$, dobi enačba za koeficient $r_{kk'}$ obliko

$$r_{kk'} = \frac{\langle k \rangle e_{kk'}}{Np_k p_{k'}}. \quad (13)$$

Za scale-free omrežja so ti pogoji zadoščeni, ko velja $k, k' > (aN)^{\frac{1}{\gamma+1}}$, kjer je a konstanta odvisna od p_k . Ta vrednost je manjša od naravne meje k_{\max} . Posledično ta vrednost predstavlja spodnjo vrednost za strukturno mejo, kar pomeni, da vedno ko stopnja vozlišč pade pod to vrednost, velja $r_{kk'} < 1$.

Za nevtralna omrežja vemo da velja

$$e_{jk} = q_j q_k = \frac{kk' p_k p_{k'}}{\langle k \rangle^2}. \quad (14)$$

Tako lahko koeficient $r_{kk'}$ zapišemo kot

$$r_{kk'} = \frac{kk'}{\langle k \rangle N}. \quad (15)$$

Od tod sledi, da ker mora strukturna meja zadostiti pogoju $r_{kk'}$, ima le-ta obliko

$$k_s(N) \sim (\langle k \rangle N)^{\frac{1}{2}}. \quad (16)$$

Opazimo, da je strukturna zgornja meja neodvisna od porazdelitve stopnje vozlišč, kar pomeni, da je tudi neodvisna od parametra γ v omrežjih scale-free. □

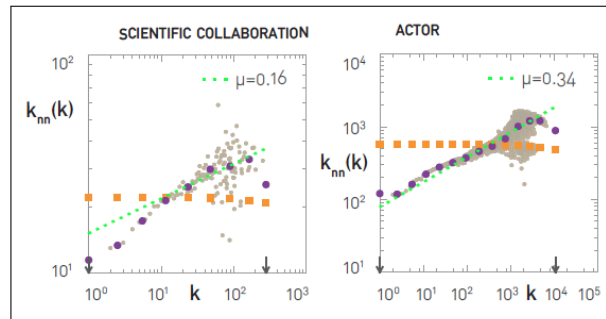
Za scale-free omrežja je ta naravna zgornja meja $k_{\max} = N^{\frac{1}{\gamma-1}}$. Tako lahko razlikujemo med dvema skupinama omrežij.

Prva skupina so omrežja, ki nimajo značilnosti scale-free omrežij ($\gamma > 3$), kar pomeni, da je ekponent v izrazu za k_{max} vedno manjši od $\frac{1}{2}$. Od tod pa sledi, da je tudi vedno manjši od k_s .

Druga skupina pa so omrežja z lastnostjo scale-free. Za te pa velja, da je $\gamma < 3$ in tako je pričakovana največja stopnja vozlišč k_{max} lahko več od vrednosti k_s . Tako lahko v omrežjih z lastnostjo scale-free pričakujemo vpliv strukturne zgornje meje.

Vendar značilnost povezovanja različnih vozlišč glede na stopnjo ni vedno posledica strukturne zgornje meje, zato je potrebno takšno obnašanje zaznati. To lahko storimo z naključnostnim vzdrževanjem stopenj vozlišč s pomočjo enostavnih povezav (angl. Degree Preserving Randomization with Simple Links). Algoritem deluje tako, da izvajamo naključnostno vzdrževanje stopenj vozlišč (vzamemo določeno število povezav, jih premešamo tako, da ohranimo stopnje vozlišč), po vsakem koraku pa odstranimo morebitne večkratne povezave. Če se po izvajanju algoritma vrednosti funkcij $k_{nn}(k)$ in $k_{nn}^{R-S}(k)$ ne razlikujeta, potem smo zaznali vpliv strukturne zgornje meje. V nasprotnem primeru, pa lastnost različnostnega povezovanja izhaja iz drugega neznanega procesa.

Na sliki 9 vidimo dve omrežji, za kateri velja, da sta podobnostno povezani in nanju ne vpliva strukturna zgornja meja, saj le-ta vpliva samo na različno povezana omrežja.



Slika 9: Grafa funkcije $k_{nn}(k)$ omrežje sodelovanja znanstvenikov in za omrežje filmskih igralcev .

1.6 ZAKLJUČEK

Pri analizi omrežji želimo z določenimi splošnimi karakteristikami vseh omrežij, čim natančneje opisati posamezne splošne modele. Opisali smo mero podobnostnega povezovanja, kjer nas zanima ali se med seboj povezujejo karakteristično podobna ali različna vozlišča. Mark Newman je leta 2002 v [1] prvi predstavil mero podobnostnega

povezovanja in se osredotočil na povezovanje vozlišč, glede na njihovo stopnjo.

V besedilu smo pregledali značilnosti treh mer karakteristike podobnostnega povezovanj, poudarili njihove prednosti in slabosti ter jih med seboj primerjali. V zaključku smo predstavili še pojav strukturne zgornje meje in dokazali spodnjo mejo glede na število in povprečno stopnjo vozlišč v omrežju.

Karakteristika podobnostnega povezovanja je veliko uporabljena pri analizi bioloških in socialnih omrežji, predvsem pri prepoznavanju potencialnih novih povezav. Vsaka dodatna karakteristika na področju analize omrežji prinese možnost natančnejše definicije modelov in razumevanje različnih pojavov v omrežjih.

LITERATURA

- [1] M. E. J. Newman, Assortative mixing in networks, *Physical review letters*, št. 89, zv. 20, str. 208701, 2002.
- [2] M. E. J. Newman, Mixing patterns in networks, *Physical Review E*, št. 67, zv. 2, str. 026126, 2003.
- [3] M. Boguná, R. Pastor-Satorras, A. Vespignani, The European Physical Journal B-Condensed Matter and Complex Systems, *The European Physical Journal B-Condensed Matter and Complex Systems*, št. 38, zv. 2, str. 205–209, 2004.
- [4] A. L. Barabaasi, *Network Science*, Citeseer, 2016.
- [5] S. Maslov, K. Sneppen in A. Zaliznyak, Detection of topological pattern in complex networks: Correlation profile of the Internet, *Physica A: Statistical Mechanics and its Applications*, št. 333, str. 529-540, 2004.
- [6] E. Estrada, *The structure of complex networks: theory and applications*, OUP Oxford, 2011.

Kneserjevi grafi so poimenovani po Martinu Kneserju, ki jih je prvi preučeval leta 1955. Uvedel jih je Lovász, da je z njimi leta 1978 dokazal Kneserjevo domnevo [5]. Pogledali si bomo nekaj zanimivih lastnosti Kneserjevih grafov, osredotočili pa se bomo na njihovo kromatično število in si ogledali nekaj korakov v dokazu le-tega. Večina informacij je vzeta iz [4], kjer se nahajajo tudi podrobnejši dokazi trditev in izrekov.

2.1 OSNOVNE DEFINICIJE

Definicija 2.1. *Kneserjev graf* je graf, za katerega velja:

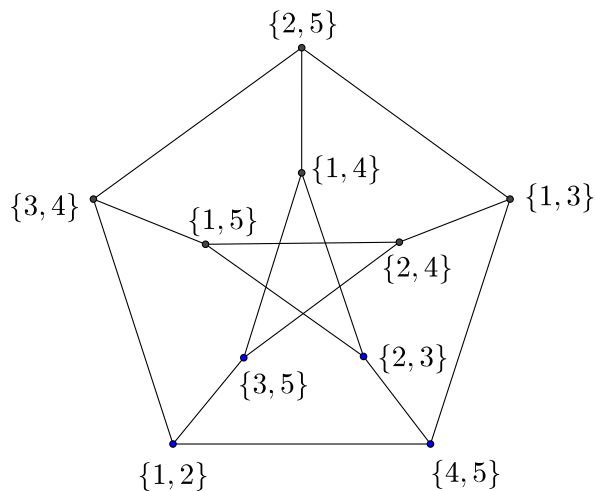
- vozlišča so r -podmnožice množice $\{1, 2, \dots, d\}$,
- vozlišči sta povezani, če sta pripadajoči r -podmnožici disjunktni.

Označimo ga $K(d, r)$, $K_{d:r}$ ali $KG_{d,r}$.

Oglejmo si nekaj primerov Kneserjevih grafov.

Zgled 2.2. Nekaj standardnih primerov Kneserjevih grafov:

- $K(d, 1)$ je izomorfen polnemu grafu na d vozliščih.
- $K(5, 2)$ je izomorfen Petersenovemu grafu. Primer izomorfizma je na diagramu na sliki 10.
- $K(d, 2)$ je komplement povezavnega grafa od polnega grafa K_d . Razmislimo: V komplementu povezavnega grafa sta dve vozlišči povezani natanko tedaj, ko v povezavnem grafu $L(K_d)$ nista povezani. Označimo vozlišča grafa K_d z naravnimi števili $V(K_d) = \{1, 2, \dots, d\}$ in vozlišča povezavnega grafa $V(L(K_d)) =$



Slika 10: Izomorfizem med $K(5,2)$ in Petersenovim grafom.

$\{\{1,2\}, \{1,3\}, \dots, \{d-1, d\}\}$, ki predstavljajo neusmerjene povezave polnega grafa K_d . Vozlišči v povezavnem grafu sta povezani, ko imata pripadajoči povezavi v K_d skupno krajišče, torej ko imata množici, ki ju označujeta, neprazen presek. Torej vozlišči v $L(K_d)$ nista povezani natanko tedaj, ko sta disjunktni, kar ustreza definiciji Kneserjevega grafa.

- $K(d, d-1)$ je izomorfen grafu brez povezav, $\overline{K_d}$, saj nobeni dve $(d-1)$ -množici nista disjunktni. V splošnem, če velja $d < 2r$, je $K(d, r)$ graf brez povezav, zato ponavadi privzamemo, da velja $d \geq 2r$.
- Poseben primer Kneserjevih grafov so Schrijverjevi grafi, ki so definirani na naslednji način: Vozlišča so r -podmnožice množice $\{1, 2, \dots, d\}$, pri čemer velja, da r -podmnožica ne vsebuje nobenih dveh zaporednih elementov (gledano ciklično, torej tudi ne vsebuje hkrati d in 1). Dve vozlišči sta povezani, če sta disjunktni.

2.2 OSNOVNE LASTNOSTI KNESERJEVIH GRAFOV

Oglejmo si nekaj osnovnih lastnosti Kneserjevih grafov $K(d, r)$:

- Kneserjevi grafi imajo $\binom{d}{r}$ vozlišč. Vsako vozlišče je krajišče $\binom{d-r}{r}$ povezav (število načinov za izbiro r -podmnožice, ki je disjunktna s prvo). Graf je torej $\binom{d-r}{r}$ -regularen. Od tod zlahka izračunamo, da ima $\frac{1}{2}\binom{d}{r}\binom{d-r}{r}$ povezav.
- Je vozliščno in povezavno tranzitiven graf.
- Če je $d < 3r$ graf nima trikotnikov. To je očitno, saj ne obstajajo tri paroma disjunktna r -podmnožice. Podobno, če je $d \geq 2r + 2$, graf vsebuje cikel dolžine 4 [3].
- Če je $d \geq 3r$, ima Kneserjev graf Hamiltonski cikel [2]. Tega ne bomo dokazovali.
- Kromatično število je $\chi(K(d, r)) = d - 2r + 2$. To dejstvo je prvi dokazal Lovász [5].
- Velikost največje neodvisne množice je $\alpha(K(d, r)) = \binom{d-1}{r-1}$. To je posledica Erdős-Ko-Rado izreka.

Zadnjima dvema lastnostma se bomo še posebej posvetili. Ker dokaz kromatičnega števila ([4]) uporablja topološke rezultate, si bomo pogledali le, kako pokažemo zgornjo mejo, torej da velja $\chi(K(d, r)) \leq d - 2r + 2$. Zajeli bomo tudi bistvo dokaza Erdős-Ko-Rado izreka, nekaj zoprnih podrobnosti pa bomo izpustili, saj za razumevanje ideje niso bistvene. Za razumevanje dokazov bomo potrebovali dva nova pojma, *delno barvanje (fractional colouring)* in *delna klika (fractional clique)*, ki ju zaporedoma predstavimo v naslednjih dveh razdelkih.

2.3 DELNO BARVANJE

Označimo z $I(G)$ množico vseh neodvisnih množic grafa G , z $I(G, u)$ pa množico vseh neodvisnih množic, ki vsebujejo vozlišče u .

Definicija 2.3. *Delno barvanje grafa G je nenegativna realna funkcija $f : I(G) \rightarrow \mathbb{R}$, da za vsako vozlišče $u \in V(G)$ velja*

$$\sum_{S \in I(G, u)} f(S) \geq 1.$$

Teža delnega barvanja je vsota vseh vrednosti. Delno barvanje je regularno, če za vsako vozlišče $u \in V(G)$ velja

$$\sum_{S \in I(G, u)} f(S) = 1.$$

Podobno kot pri (običajnem) barvanju definiramo *delno kromatično število* $\chi^*(G)$ grafa G kot najmanjša možna teža delnega barvanja na grafu G . Izkaže se, da je to dobro definirano (dokaz je preko linearnega programiranja) in da je vedno racionalno število.

Oglejmo si primer delnega barvanja.

Zgled 2.4. Definirajmo delno barvanje na ciklu C_5 . Graf ima natanko 5 neodvisnih množic velikosti 2 in vsako vozlišče leži v natanko dveh takih neodvisnih množicah. Definirajmo naše delno barvanje f na sledeči način:

$$f(S) = \begin{cases} \frac{1}{2}; & |S| = 2 \\ 0; & \text{sicer.} \end{cases}$$

Teža tega barvanja je $\frac{5}{2}$, vemo pa, da je $\chi(C_5) = 3$. Opazimo torej, da je $\chi^*(C_5) < \chi(C_5)$. Ali pa to velja v splošnem?

Porodi se torej vprašanje, kakšna je zveza med kromatičnim številom in delnim kromatičnim številom. Denimo, da imamo k -barvanje grafa G . Vsak barvni razred razdeli množico vozlišč na k paroma disjunktnih množic V_1, V_2, \dots, V_k , njihova unija pa je $V(G)$. Definiramo sedaj delno barvanje f takole:

$$f(S) = \begin{cases} 1; & S = V_i \text{ za nek } i \\ 0; & \text{sicer.} \end{cases}$$

To delno barvanje ima težo k , torej takoj velja

$$\chi^*(G) \leq \chi(G).$$

Na prejšnjem zgledu s ciklom dolžine 5 smo videli, da je delno kromatično število lahko strogo manjše od kromatičnega števila. Sedaj pa predpostavimo, da imamo delno barvanje grafa G z vrednostmi 0 in 1, ki ima težo k . Tedaj nosilec delnega barvanja tvori k paroma disjunktnih neodvisnih množic V_1, V_2, \dots, V_k , katerih unija je $V(G)$. Vsako vozlišče u pobarvamo s takšnim i , da je V_i množica z najmanjšim indeksom, ki vsebuje u . Dobimo pravilno barvanje grafa G , torej je kromatično število najmanjša teža delnega barvanja z vrednostmi 0 in 1.

Sedaj si oglejmo, kaj vemo o delnem barvanju na Kneserjevih grafih. Oglejmo si graf $K(d, r)$. Vse r -množice, ki vsebujejo nek izbran element $i \in \{1, 2, \dots, d\}$ tvorijo neodvisno množico velikosti $\binom{d-1}{r-1}$ in vsako vozlišče leži v natanko r takih neodvisnih množicah. Definiramo delno barvanje f tako, da ima vrednost $\frac{1}{r}$ na teh neodvisnih množicah in vrednost 0 na preostalih. Ker je takšnih neodvisnih množic d (za vsak element ena), je teža tega barvanja $\frac{d}{r}$, torej velja

$$\chi^*(K(d, r)) \leq \frac{d}{r}.$$

Ta rezultat bo pomemben v nadaljevanju.

Pri delnem barvanju moramo upoštevati še prazno množico, ki v principu je neodvisna množica, torej je za delno barvanje f definirano tudi $f(\emptyset)$. Vendar, če $f(\emptyset) \neq 0$, lahko definiramo, da je $f(\emptyset) = 0$ in dobimo delno barvanje z manjšo težo, torej lahko brez škode privzamemo, da je $f(\emptyset) = 0$.

2.4 DELNA KLIKA

Za razliko od običajne klike, delna klika ni podmnožica vozlišč, ampak je funkcija.

Definicija 2.5. *Delna klika (fractional clique)* grafa G je nenegativna realna funkcija $f : V(G) \rightarrow \mathbb{R}$, tako da za vsako neodvisno množico $S \subset V(G)$ velja

$$\sum_{v \in S} f(v) \leq 1.$$

Teža delne klike je vsota vrednosti funkcije.

Podobno lahko definiramo *delno klikno število (fractional clique number)* grafa G , ki je največja možna teža delne klike in ga označimo $\omega^*(G)$. Če imamo kliko velikosti k , potem je karakteristična funkcija te klike tudi delna klika s težo k , torej velja

$$\omega(G) \leq \omega^*(G),$$

kjer $\omega(G)$ označuje velikost največje klike v grafu G . Neenakost je lahko tudi stroga, kar nam pokaže naslednji zgled.

Zgled 2.6. Za graf vzemimo cikel C_5 in na njem definiramo delno kliko $f(v) = \frac{1}{2}$. Teža te delne klike je $\frac{5}{2}$, velikost največje klike v C_5 pa je 2, torej je neenakost res lahko stroga.

Označimo z $\alpha(G)$ največjo neodvisno množico v grafu G , torej $\alpha(G) = \omega(\overline{G})$. Velja naslednja lema:

Lema 2.7. *Za vsak graf G velja*

$$\omega^*(G) \geq \frac{|V(G)|}{\alpha(G)}.$$

Dokaz. Naj bo g delna klika s predpisom $g(v) := \frac{1}{\alpha(G)}$. Naj bo S poljubna neodvisna množica. Velja, da je $|S| \leq \alpha(G)$, zato je

$$\sum_{v \in S} g(v) = \frac{|S|}{\alpha(G)} \leq 1,$$

in torej je g res delna klika. Lema sledi iz dejstva, da je teža te delne klike $\frac{|V(G)|}{\alpha(G)}$. \square

Ker vemo, da so Kneserjevi grafi vozliščno tranzitivni, nas zanima tudi naslednja lastnost.

Trditev 2.8. Če je G vozliščno tranzitiven graf, je

$$\omega^*(G) = \frac{|V(G)|}{\alpha(G)}$$

in $f(v) = \frac{1}{\alpha(G)}$ je delna klika s to težo.

Dokaz. Naj bo g neničelna delna klika grafa G in naj bo $\gamma \in \text{Aut}(G)$. Definirajmo funkcijo g^γ s predpisom

$$g^\gamma(v) = g(\gamma(v)).$$

Tudi g^γ je delna klika z enako težo kot g (saj je γ avtomorfizem). Tedaj je tudi

$$\hat{g} := \frac{1}{|\text{Aut}(G)|} \sum_{\gamma \in \text{Aut}(G)} g^\gamma$$

delna klika z enako težo kot g . Ker je G vozliščno tranzitiven, je \hat{g} konstantna na vseh vozliščih iz G . Če želimo, da je konstantna preslikava c delna klika, mora (zaradi velikosti neodvisnih množic in definicije klike) veljati $c \leq \frac{1}{\alpha(G)}$, torej $\hat{g} \leq \frac{1}{\alpha(G)}$. Po prejšnji lemi rezultat sledi. \square

Velja tudi naslednja lema, ki je ne bomo dokazovali, je pa njen rezultat pomemben v nadaljevanju:

Trditev 2.9. Če je G vozliščno tranzitiven graf, velja

$$\chi^*(G) \leq \frac{|V(G)|}{\alpha(G)}.$$

2.5 DUALNOST

Delno barvanje in delno kliko si lahko predstavljamo tudi z matriko. Naj bo B 0/1-matrika z vrsticami indeksiranimi z vozlišči grafa G in stolpci karakterističnimi vektorji neodvisnih množic. Oglejmo si primer.

Zgled 2.10. Za graf vzemimo cikel na štirih vozliščih, C_4 . Označimo njegova vozlišča $\{1, 2, 3, 4\}$.

Neodvisne množice so $(\emptyset, \{1\}, \{2\}, \{3\}, \{4\}, \{1, 4\}, \{2, 3\})$, stolpci matrike si sledijo v enakem vrstnem redu. Pripadajoča matrika je

$$B = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 \end{bmatrix}.$$

Dimenzije matrice so 4×7 , vrstice namreč ustrezajo vozliščem, stolpci pa neodvisnim množicam.

Označimo z $\mathbb{1}$ vektor samih enic. Nenegativen vektor f , za katerega velja $Bf \geq \mathbb{1}$, to pomeni, da je vsaka koordinata vektorja Bf vsaj 1, predstavlja delno barvanje. V tem primeru i -ta koordinata vektorja f predstavlja vrednost delnega barvanja na i -ti neodvisni množici. Ker delamo s končnimi grafi, je to res stolpec dolžine enake številu neodvisnih množic v grafu G . Pogoja sta res ekvivalentna, i -ta komponenta vektorja Bf predstavlja skalarni produkt med barvanjem f in i -to vrstico matrice B , torej vrstico, ki pripada i -temu vozlišču (v_i). Ta vrstica ima enice na mestih, kjer v_i pripada neodvisni množici in ničle drugod. Skalarni produkt nam da ravno pogoj za i -to vozlišče v definiciji delnega barvanja.

Podobno je nenegativen vektor g , za katerega velja $g^T B \leq \mathbb{1}^T$, delna klika grafa G . Zopet i -ta koordinata vektorja g predstavlja vrednost delne klike v i -tem vozlišču. Produkt $g^T B \leq \mathbb{1}$ je sedaj vrstica, na vsaki komponenti dobimo pogoj za eno neodvisno množico.

S to interpretacijo lahko iskanje $\chi^*(G)$ in $\omega^*(G)$ formuliramo kot minimizacijski oziroma maksimizacijski problem, ki ju rešujemo z linearnim programiranjem.

Problem 2.11. *Problem iskanja delnega kromatičnega števila $\chi^*(G)$:*

$$\begin{aligned} \min \quad & \mathbb{1}^T f \\ & Bf \geq \mathbb{1} \\ & f \geq 0. \end{aligned}$$

Problem 2.12. *Problem iskanja delnega klinkega števila $\omega^*(G)$:*

$$\begin{aligned} \max \quad & g^T \mathbb{1} \\ & g^T B \leq \mathbb{1} \\ & g \geq 0. \end{aligned}$$

Oba problema sta problema iz linearnega programiranja, še več, problema sta dualna. Če uporabimo znan šibki izrek o dualnosti za linearne programe, dobimo pomemben rezultat:

Trditev 2.13. *Za poljuben graf G velja*

$$\omega^*(G) \leq \chi^*(G).$$

Če uporabimo krepki izrek o dualnosti iz linearnega programiranja, velja celo enakost

$$\omega^*(G) = \chi^*(G).$$

Za vozliščno tranzitivne grafe lahko o tej enakosti povemo še nekaj več:

Posledica 2.14. Če je G vozliščno tranzitiven graf, je

$$\omega^*(G) = \chi^*(G) = \frac{|V(G)|}{\alpha(G)}.$$

Dokaz. Če uporabimo šibko dualnost, torej trditve 2.13, ter lemi 2.7 in 2.9, dobimo

$$\frac{|V(G)|}{\alpha(G)} \leq \omega^*(G) \leq \chi^*(G) \leq \frac{|V(G)|}{\alpha(G)}.$$

Trditve sledi. □

Iz leme 2.7 in iz šibke dualnosti, torej trditve 2.13, sledi tudi, da za vsak graf G velja

$$\chi^*(G) \leq \frac{|V(G)|}{\alpha(G)}.$$

2.6 CIKLIČNI INTERVALNI GRAFI

Z namenom, da kaj povemo o delnem kromatičnem številu Kneserjevih grafov, si ogledamo ciklične intervalne grafe.

Definicija 2.15. *Ciklični intervalni graf* $C(d, r)$ je graf dobljen na sledeč način:

- vozlišča so cilkični premiki podmnožice $\{1, 2, \dots, r\}$ v množici $\Omega = \{1, 2, \dots, d\}$ modulo d ,
- dve vozlišči sta povezani, če sta pripadajoči množici disjunktni.

Že iz same definicije sledi, da je ciklični intervalni graf $C(d, r)$ induciran podgraf v Kneserjevem grafu $K(d, r)$ in da je $C(d, r)$ cikličen, torej tudi vozliščno tranzitiven. Če je $d < 2r$ je $C(d, r)$ brez povezav, zato nas zanimajo le takšni, za katere velja $d \geq 2r$. V njih lahko karakteriziramo maksimalno velikost neodvisne množice.

Trditev 2.16. Če je $d \geq 2r$, je neodvisna množica v $C(d, r)$ velikosti največ r . Še več, neodvisna množica velikosti r vsebuje vozlišča, ki vsa kot r -množice vsebujejo isti element $j \in \{1, 2, \dots, d\}$.

Dokaz. Naj bo S neodvisna množica v $C(d, r)$. Ker je $C(d, r)$ vozliščno tranzitiven, lahko predpostavimo, da S vsebuje r -množico $\beta = \{1, 2, \dots, r\}$. Naj bo S_1 množica tistih r -množic v S , ki vsebujejo 1 in S_r množica tistih r -množic v S , ki vsebujejo r .

Naj bo j najmanjši element, ki leži v vseh r -množicah v S_r (tak j obstaja, ker vemo, da r leži v vseh množicah iz S_r). Najmanjši element v vsaki množici iz S_r je torej največ j in ker so najmanjši elementi za

vse množice iz S_r različni (imamo le ciklične premike), sledi, da je $|S_r| \leq j$. Nadalje ima vsaka r -množica v S_1 skupen element z vsako r -množico iz S_r in ker so r -množice dobljene ciklično, velja, da je j vsebovan v vseh r -množicah iz S_1 in posledično velja $|S_1| \leq r - j + 1$.

Ker je $d \geq 2r$, se lahko zgodi, da je $S_1 \cap S_r = \{\beta\}$ (preseka je v splošnem lahko tudi kaj več, zanima nas najslabši primer). Ker imamo le ciklične premike, je očitno $S = S_1 \cup S_r$, torej velja:

$$|S| \leq |S_1| + |S_r| - 1 \leq (r - j + 1) + j - 1 = r.$$

Če velja enakost, potem S sestavljajo vozlišča v $C(d, r)$, ki vsebujejo element j . \square

Posledica 2.17. Če je $d \geq 2r$, potem velja $\chi^*(C(d, r)) = \frac{d}{r}$.

Dokaz. Ker je $C(d, r)$ vozliščno tranzitiven, po trditvi 2.14 velja $\chi^*(C(d, r)) = \frac{|V(C(d, r))|}{\alpha(C(d, r))}$. Po prejšnji trditvi 2.16 je velikost največje neodvisne množice enaka r . Velja $|V(C(d, r))| = d$, saj so to ravno vsi ciklični premiki. Ko združimo enakosti, posledica sledi. \square

Posledica 2.18. Če je $d \geq 2r$, potem velja $\chi^*(K(d, r)) = \frac{d}{r}$.

Dokaz. Videli smo že, da je $\chi^*(K(d, r)) \leq \frac{d}{r}$. Ker je $C(d, r)$ podgraf v Kneserjevem grafu, velja:

$$\frac{d}{r} = \chi^*(C(d, r)) \leq \chi^*(K(d, r)) \leq \frac{d}{r}.$$

Torej je $\chi^*(K(d, r)) = \frac{d}{r}$.

Da velja neenakost $\chi^*(C(d, r)) \leq \chi^*(K(d, r))$ ni povsem očitno. V splošnem velja, da če je $\varphi : G \rightarrow H$ homomorfizem, potem je $\chi^*(G) \leq \chi^*(H)$, a tega ne bomo dokazovali. \square

2.7 ERDÖS-KO-RADO IZREK

Ko uporabimo izpeljane rezultate na Kneserjevih grafih, dobimo znan izrek.

Izrek 2.19 (Erdős-Ko-Rado izrek). Naj bo $d > 2r$. Tedaj velja

$$\alpha(K(d, r)) = \binom{d-1}{r-1}.$$

Neodvisna množica velikosti $\binom{d-1}{r-1}$ vsebuje vse r -podmnožice, ki vsebujejo izbran element iz $\Omega = \{1, 2, \dots, d\}$.

Ideja dokaza. Iz posledice 2.18 sledi $\chi^*(K(d, r)) = \frac{d}{r}$ in iz trditve 2.14 dobimo $\chi^*(K(d, r)) = \frac{|V(K(d, r))|}{\alpha(K(d, r))}$. Preoblikujemo neenakosti in dobimo:

$$\alpha(K(d, r)) = \frac{1}{\chi^*(K(d, r))} |V(K(d, r))| = \frac{r}{d} \binom{d}{r}.$$

In ko razpišemo po definiciji binomskega simbola:

$$\frac{r}{d} \binom{d}{r} = \frac{r \cdot d!}{d \cdot r! (d-r)!} = \frac{r \cdot d \cdot (d-1)!}{d \cdot r \cdot (r-1)! (d-1-(r-1))!} = \binom{d-1}{r-1}.$$

S preštevilčenjem nato pokažemo, da so te neodvisne množice ravno tiste, ki vsebujejo določen element iz Ω . \square

Opazimo, da je v izreku predpostavljena stroga neenakost $d > 2r$. V primeru enakosti $d = 2r$, je meja še vedno veljavna, maksimalna velikost neodvisne množice je $\binom{2r-1}{r-1} = \frac{1}{2} \binom{2r}{r}$. Vendar pa je $K(2r, r)$ izomorfen $\binom{2r-1}{r-1}$ vozliščno disjunktnim kopijam K_2 , zato ima $2^{\binom{2r-1}{r-1}}$ maksimalnih neodvisnih množic in ne le $2r$ tistih opisanih v izreku.

2.8 HOMOMORFIZMI NA KNESERJEVIH GRAFIH

S pomočjo homomorfizmov bomo izpeljali zgornjo mejo za kromatično število Kneserjevih grafov. Zanimivo je tudi dejstvo, da so Kneserjevi grafi jedra.

Trditev 2.20. Če je $d > 2r$, je Kneserjev graf $K(d, r)$ jedro.

Ideja dokaza. Za poljuben homomorfizem pokažemo, da je surjekcija: gledamo presek maksimalnih neodvisnih množic, ki vsebujejo neko vozlišče in si pomagamo z Erdős-Ko-Rado izrekom. \square

Oglejmo si izrek, ki nam daje poglavitni homomorfizem med Kneserjevimi grafi.

Izrek 2.21. Če je $d \geq 2r$ in $r \geq 2$, obstaja homomorfizem grafov iz grafa $K(d, r)$ v graf $K(d-2, r-1)$.

Dokaz. Če je $d = 2r$, je $K(d, r) \cong \binom{2r-1}{r-1} K_2$, ki dopušča homomorfizem v kateri koli graf s povezavo (v posebnem: je dvodelen graf, dopušča homomorfizem v K_2). Zato lahko sedaj predpostavimo, da je $d > 2r$ in da gledamo r -množice nad množico $\Omega = \{1, 2, \dots, d\}$.

Zlahka najdemo homomorfizem φ iz $K(d-1, r)$ v $K(d-2, r-1)$: vsako r -množico preslikamo v $(r-1)$ -podmnožico tako, da izbrišemo največji element. Sedaj identificiramo $K(d-1, r)$ s podgrafom $K(d, r)$ induciranim s tistimi vozlišči, ki ne vsebujejo elementa d , in skušamo razširiti φ na celoten graf $K(d, r)$.

Opazimo, da vozlišča, na katerih φ še ni definiran (torej tista, ki so v $K(d, r)$, ne pa v $K(d - 1, r)$), vsa vsebujejo element d , torej niso disjunktna in tvorijo neodvisno množico, označimo jo z S . Naj bo S_i množica vozlišč, ki vsebujejo elemente $d, d - 1, \dots, d - i + 1$, ne vsebujejo pa elementa $d - i$. Množice S_1, \dots, S_r tvorijo particijo (razdelitev) množice S .

Sedaj postopoma definiramo $\varphi : K(d, r) \rightarrow K(d - 2, r - 1)$. Če je $\alpha \in S_1$, postavimo $\varphi(\alpha) = \alpha \setminus \{d\}$. Če je $i > 1$ in $\alpha \in S_i$, potem velja $d - i \notin \alpha$. Potem $\varphi(\alpha)$ dobimo tako, da iz α izbrišemo d in zamenjamo $d - 1$ z $d - i$.

Preverimo, ali je funkcija smiselno definirana. Če je $\alpha \in S_1$, vemo, da α ne vsebuje elementa $d - 1$. Sledi $\varphi(\alpha) = \alpha \setminus \{d\}$ je množica moči $r - 1$ in ne vsebuje niti d niti $d - 1$, torej je $(r - 1)$ -podmnožica množice $\{1, 2, \dots, d - 2\}$. Če je $i > 1$ in $\alpha \in S_i$, potem smo zbrisali ven tako d kot $d - 1$ in dodali en nov element, spet dobimo $(r - 1)$ -podmnožico množice $\{1, 2, \dots, d - 2\}$.

Preverimo še, če se povezave slikajo v povezave. Denimo, da sta vozlišči x in y v grafu $K(d, r)$ povezani, torej sta pripadajoči množici disjunktni. Ker sta povezani, ne moreta obe vsebovati elementa d . Če nobeno od vozlišč x in y ne vsebuje d , potem φ le zbrise največji element iz vsake množice, vozlišči ostaneta disjunktni in torej sta sliki $\varphi(x)$ in $\varphi(y)$ povezani. Sedaj predpostavimo, da eno od vozlišč (brez škode je to x) vsebuje element d . Vozlišču y smo zgolj zbrisali največji element in ostane disjunkten z vozliščem x . V primeru, da je $x \in S_1$, smo le zbrisali d in x ter y ostaneta disjunktni množici, torej sliki povezani. Če pa je $x \in S_i$ za $i > 1$, potem smo poleg brisanja d (ki očitno ne pokvari disjunktnosti) dodali še element $d - i$. Problem bi nastal, če y vsebuje $d - i$. Vendar pa vemo, da je prej x vseboval vse elemente $d, d - 1, \dots, d - i + 1$ in ker sta x in y disjunktni množici, je največji možen element, ki ga lahko y vsebuje, enak $d - i$. Torej, če je y vseboval $d - i$, je bil to gotovo največji element in smo ga s φ izbrisali, torej $\varphi(y)$ gotovo ne vsebuje $d - i$, kar pa pomeni, da sta sliki $\varphi(x)$ in $\varphi(y)$ disjunktni kot množici in torej vozlišči povezani. \square

Posledica 2.22. Za Kneserjev graf $K(d, r)$ velja

$$\chi(K(d, r)) \leq d - 2r + 2.$$

Dokaz. Ko zaporedoma $(r - 1)$ -krat uporabimo zgornji izrek 2.21, dobimo kompozitum homomorfizmov $K(d, r) \rightarrow K(d - 2r + 2, 1)$. Ker je $K(d - 2r + 2, 1)$ poln graf na $d - 2r + 2$ vozliščih, dobimo

$$\chi(K(d, r)) \leq d - 2r + 2.$$

\square

Tako smo pokazali zgornjo mejo za kromatično število za Kneserjeve grafe. Da pokažemo tudi spodnjo mejo, si pomagamo s Borsukovim izrekom iz topologije. Natančen dokaz je v [4], ki se naslanja na Bárányjev dokaz Kneserjeve domneve [1].

LITERATURA

- [1] I. Bárány, *A short proof of Kneser's conjecture*, Journal of Combinatorial Theory Series A, 25 (1987), 325-326.
- [2] Y. Chen (2000), *Kneser graphs are Hamiltonian for $n \geq 3k$* , Journal of Combinatorial Theory, Series B 80 (1), 69-79.
- [3] T. Denley (1997), *The odd girth of the generalised Kneser graph*, European Journal of Combinatorics 18 (6), 607-611.
- [4] C. Godsil, G. Royle, *Algebraic Graph Theory*, Graduate Texts in Mathematics 207, Springer, 2001, 9-10, 135-161.
- [5] L. Lovász, *Kneser's conjecture, chromatic number, and homotopy*, Journal of Combinatorial Theory Series A, 25 (1978), 319-324.
- [6] *Kneser Graph*, v: Wikipedia: The Free Encyclopedia, [ogled 10. 4. 2016], dostopno na https://en.wikipedia.org/wiki/Kneser_graph
- [7] *Kneser Graph*, v: Wolfram Mathworld, [ogled 20. 5. 2016], dostopno na <http://mathworld.wolfram.com/KneserGraph.html>

MERE CENTRALNOSTI

Rok Fortuna

3.1 UVOD

Če poznamo strukturo omrežja lahko iz nje izračunamo marsikatere uporabne mere, ki kvantificirajo različne lastnosti omrežja. Dostikrat nas zanimajo pomembna vozlišča v omrežju. Pri tem nam pomagajo mere centralnosti (angl. centrality measures). Veliko idej kako prepoznati pomembna vozlišča v omrežju izvira iz področja analize socialnih omrežij, uporabljajo pa se tudi v računalništvu, biologiji, fiziki itd.

Ideje, ki jih bomo matematično opisali v nadaljevanju temeljijo na sledečih lastnostih vozlišč:

- zmožnost komuniciranja vozlišča z drugimi,
- razdalja vozlišča od večine drugih vozlišč,
- med koliko vozlišči se neko vozlišče nahaja.

3.2 PREPROSTE MERE CENTRALNOSTI

Najbolj preprosta mera centralnosti je stopnja vozlišča. Dostikrat je to zelo ustrezna in uporabna mera. V socialnih omrežjih imajo posamezniki z veliko povezavami načeloma velik vpliv oz. velik dostop do informacij.

V mnogih primerih je bolje uporabiti kakšno drugo mero centralnosti. Če imamo omrežje v katerem vozlišča med seboj komunicirajo (npr. prometno omrežje) in nas zanimajo vozlišča prek katerih poteka

največ komunikacije, potem lahko hitro vidimo, da je stopnja vozlišča neustrezna mera centralnosti. Ustrezna izbira za mero centralnosti je npr. vmesnostna centralnost (angl. betweenness centrality). Za posamezno vozlišče se izračuna kot število najkrajših poti (med vsemi ostalimi pari vozlišč), na katerih to vozlišče leži. Večina algoritmov, ki računajo vmesnostno centralnost ima časovno zahtevnost reda $O(n^3)$. Leta 2001 je bil odkrit algoritem s časovno zahtevnostjo $O(nm)$ [6], kar je v realnosti (ker so omrežja redki grafi) bistveno boljše.

Pogosto se uporablja tudi bližinska centralnost, pri kateri so najbolj pomembna vozlišča tista, ki so najbližja ostalim. Za posamezno vozlišče se izračuna kot obratna vrednost vsote razdalj do vseh ostalih vozlišč.

Preprostost teh treh pristopov pa za mnoge aplikacije omrežne znanosti ni dovolj. Zlasti se veliko uporabljajo spektralne metode, med katerimi je najbolj znan PageRank [9]. V nadaljevanju bomo natančneje preučili spektralne metode za izračun centralnosti [2].

3.3 SPEKTRALNA CENTRALNOST

Spektralna centralnost temelji na ideji, da je vozlišče pomembno, če je povezano z drugimi pomembnimi vozlišči. Koncept spektralnih metod si lahko predstavljamo na sledeč način. Namesto, da vozlišču damo točke za vsakega sosedu, ki ga ima (kot pri centralnosti glede na stopnje), mu damo točke glede na to, koliko točk imajo njegovi sosedi [5].

Za nadaljne izpeljave potrebujemo sledeč Perron-Frobeniusev za nereducibilne matrike. To so matrike, ki jih s permutiranjem vrstic in stolpcev ne moremo transformirati v bločno zgornje trikotno matriko. V primeru matrike sosednosti nereducibilnost pomeni, da imamo bodisi neusmerjen graf, ali pa da je usmerjen graf krepko povezan.

Izrek 3.1 (Perron-Frobenius). *Naj bo A nereducibilna nenegativna kvadratna matrika $\rho(A) = r$ njen spektralni radij (največja lastna vrednost). Potem držijo naslednje trditve:*

- Število r je pozitivno realno število.
- Kratnost r -ja kot ničle karakterističnega polinoma je 1.
- Lastni vektor pripadajoč lastni vrednosti r je pozitiven.

Nadaljujmo s spektralno centralnostjo. Naj bo A matrika sosednosti krepko povezanega grafa in naj bo $x(0)$ vektor začetne razdelitve centralnosti. Nato izračunamo:

$$x(1)_i = \sum_j A_{ij}x(0)_j,$$

kjer je A_{ij} element matrike sosednosti. V primeru usmerjenega grafa je $A_{ij} = 1$, če obstaja usmerjena povezava od j do i , sicer je 0. Če ta postopek ponovimo t -krat, dobimo

$$x(t) = A^t x(0).$$

Vsako množenje vektorja z matriko sosednosti vsakemu vozlišču dodeli vsoto vrednosti sosedov. Predpostavimo, da se $x(0)$ da zapisati kot linearna kombinacija lastnih vektorjev matrike A

$$x(0) = \sum_i c_i v_i,$$

kjer je v_i i -ti lastni vektor in c_i neka konstanta. Naj bo λ_1 največja lastna vrednost matrike A (iz Perron-Frobeniusevega izreka vemo, da je ena sama). Potem

$$x(t) = A^t \sum_i c_i v_i = \sum_i \lambda_i^t c_i v_i = \lambda_1^t \sum_i c_i \left(\frac{\lambda_i}{\lambda_1}\right)^t. \quad (17)$$

Ker je $\lambda_i/\lambda_1 < 1$ za vsak $i \neq 1$, gredo vsi členi zgornje vsote proti 0, ko $t \rightarrow \infty$ in gre $x(t) \rightarrow c_1 \lambda_1^t v_1$. Kar pomeni, da je vektor, ki ga dobimo soracionalen vodilnemu lastnemu vektorju.

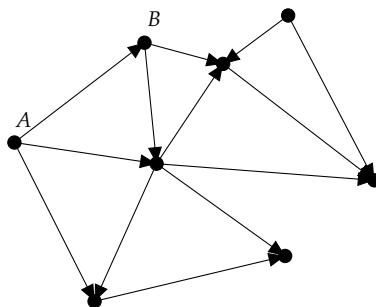
To je ravno postopek iteracije moči (angl. power iteration), s katerim iščemo lastni vektor matrike sosednosti, ki pripada po absolutni vrednosti največji lastni vrednosti [5]. Pogoj postopka iteracije moči za konvergenco je, da začetni vektor in vektor ki ga iščemo nista ortogonalna. Iz Perron-Frobeniusevega izreka vemo, da bo vektor, ki ga iščemo pozitiven, zato lahko za začetni vektor izberemo katerikoli pozitiven vektor npr. $x(0)_i = 1$ za vse i . Vektor, ki ga dobimo torej ustreza enačbi:

$$Ax = \lambda x,$$

kjer je i -ta komponenta x ravno centralnost i -tega vozlišča. Rešitev bi lahko dobili tudi z računanjem inverza matrike, kar pa je v praksi računsko prezahtevno.

V teoriji se lahko spektralna centralnost uporablja za usmerjena in neusmerjena omrežja [3]. Kljub temu, pa se najboljše obnese pri neusmerjenih omrežjih. Problem pri usmerjenih omrežjih nastane že pri matriki A , ki je nesimetrična. Zaradi nesimetričnosti Perron-Frobeniusev ne velja in posledično nimamo zagotovljene konvergenca pri postopku iteracije moči [1].

Poglejmo si primer, ko spektralna centralnost ne deluje, kot bi si želeli. Iz slike 11, vidimo, da ima vozlišče A po spektralni centralnosti vrednost 0. S tem ni nič narobe, saj je smiselno, da imajo vozlišča, na katere noben ne kaže nobeno drugo vozlišče centralnost 0. Če



Slika 11: Del usmerjenega omrežja.

pogledamo drugo vozlišče B, vidimo, da nanj kaže le vozlišče A, ki pa ima centralnost 0, zato ima tudi B centralnost 0. Ta pojav bi se lahko razširjal skozi generacije vozlišč in bi na koncu imeli veliko vozlišč, ki bi imelo neupravičeno centralnost 0. V naslednjem razdelku si bomo ogledali mero centralnosti, ki opisane težave odpravlja.

3.4 KATZOVA CENTRALNOST

Katzova centralnost naslavlja probleme spektralne centralnosti in jih rešuje na sledeč način. Vsakemu vozlišču da nekaj centralnosti za stonj, ne glede na strukturo omrežja. Katzova centralnost je definirana kot

$$x_i = \alpha \sum_j A_{ij} x_j + \beta,$$

kjer je A_{ij} element matrike sosednosti, α parameter, ki kontrolira prvi člen vsote in β konstanta (centralnost, ki jo vsako vozlišče dobi za stonj).

V vektorski obliki enačbo zapišemo kot

$$x = \alpha Ax + \beta \mathbf{1},$$

kjer je $\mathbf{1}$ vektor enic. Če izrazimo x dobimo

$$x = \beta(I - \alpha A)^{-1} \cdot \mathbf{1} = \beta \sum_{i=0}^{\infty} (\alpha A)^i.$$

Iz zgornje enačbe je razvidno, da je Katzova centralnost vozlišča ekvivalentna številu poti v grafu, ki dosežejo to vozlišče.

Ko je $\alpha = 0$, potem del s spektralno centralnostjo odpade in dobijo vsa vozlišča centralnost β . Ko α večamo, se manjša učinek β in daljše poti dobivajo večji vpliv na centralnost. Ko je $\det(I - \alpha A) = 0$, potem matrika $(I - \alpha A)$ postane neobrnljiva. To povzroči divergenco centralnosti. Matrika $I - \alpha A$ prvič postane neobrnljiva, ko je $\alpha = 1/\lambda$, kjer je λ največja lastna vrednost matrike A . To lahko vidimo, če enačbo $\det(I - \alpha A) = 0$ preoblikujemo v karakteristično enačbo za A : $\det(A - \alpha^{-1}I) = 0$. Če enačimo $\alpha^{-1} = \lambda$, vidimo da bo najmanjša vrednost α , da zadanemo lastno vrednost, ravno $1/\lambda$, kjer je λ največja lastna vrednost. V praksi tako izberemo parameter $\alpha < 1/\lambda$.

Podobno kot spektralna centralnost, pa ima tudi Katzova centralnost pomankljivosti pri usmerjenih grafih. Problem se pojavi pri vozliščih, ki imajo visoko vrednost Katzove centralnosti. Vozlišča na katera taka vozlišča kažejo, dobijo celotno njihovo centralnost. To v praksi pomeni, da vsa vozlišča na katera kaže neko pomembno vozlišče postanejo prav tako zelo pomembna, čeprav so mogoče le ena izmed mnogih. Želeli bi si, da vsa vozlišča na katere kaže pomembno vozlišče dobijo le delež te pomembnosti. Te težave odpravlja Page-Rank, algoritem, ki ga podjetje Google uporablja v svojem iskalniku.

3.5 PAGERANK

PageRank naslavlja problem Katzove centralnosti, poleg tega pa imajo vrednosti centralnosti lep verjetnostni pomen. PageRank uporablja stolpično stohastično matriko S , kjer je $S_{ij} = A_{ij}/d_j$, kjer je d_j stopnja vozlišča j (v primeru usmerjenega grafa izhodna stopnja). S tem rešuje problem pri Katzovi centralnosti, da se je iz vozlišča v vse njegove sosedne prenesla celotna centralnost [4].

PageRank uporablja t.i. "Googlovo matriko":

$$G = \alpha S + (1 - \alpha) \frac{1}{n} E,$$

kjer je S stohastična matrika definirana zgoraj, $0 < \alpha < 1$ poljuben parameter in E matrika samih enic.

Preprost premislek nam pokaže, da ker je S stolpično stohastična, je tudi G stolpično stohastična [1]. S tem ko smo prišteli polno matriko E , smo dobili pozitivno matriko G (nima ničelnih vrednosti). Za pozitivne stolpično stohastične matrike nam pride prav sledeča različica Perron-Frobeniusovega izreka.

Izrek 3.2 (Perron-Frobenius). *Naj bo A pozitivna stolpično stohastična kvadratna matrika. Potem držijo naslednje trditve:*

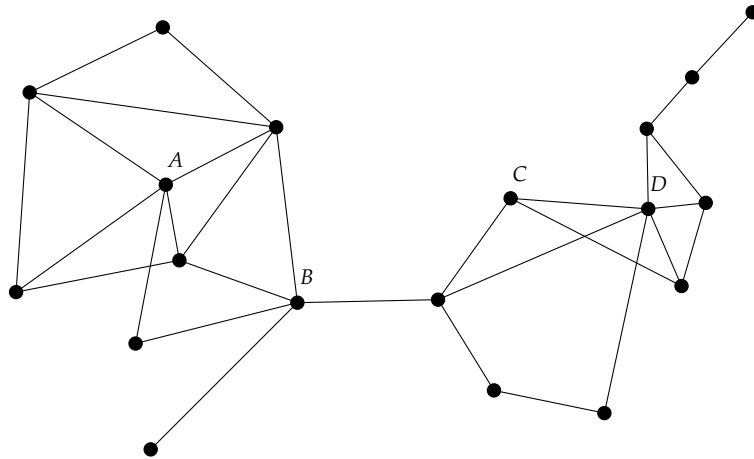
- Število 1 je lastna vrednost matrika A kratnosti ena.
- Število 1 je največja lastna vrednost, vse ostale so po absolutni vrednosti manjše od 1.
- Obstaja natanko en lastni vektor pripadajoč lastni vrednosti 1 tak, da se njegove komponente seštejejo v 1.

Zgornje lastnosti nam ponovno omogočajo, da izračunamo željen lastni vektor s pomočjo metode power iteration.

Pomen vrednosti centralnosti PageRank je povezan z naključnimi sprehodi po omrežju. Če se nek sprehajalec sprehaja po omrežju dovolj dolgo časa, nam vrednost x_i v dobljenem lastnem vektorju nam pove verjetnost, da bo v tem trenutku na strani i . Parameter α v enačbi matrike G ima pomen verjetnosti, da bo sprehajalec v nekem trenutku, ko je na neki strani dejansko sledil eni izmed povezav na tej strani, v nasprotnem primeru bo (z verjetnostjo $1 - \alpha$) skočil na poljubno stran v omrežju. Parameter α je ponavadi izbran kot 0.85 [1].

3.6 MERE CENTRALNOSTI V PRAKSI

V realnosti za različne probleme izbiramo različne mere centralnosti.



Slika 12: Primer obnašanja mer centralnosti.

vozlišče	lastnost
A	najvišja spektralna centralnost
B	najvišja vmesnostna centralnost
C	najvišja bližinska centralnost
D	najvišja centralnost glede na stopnjo vozlišča

Tabela 13: Vozlišča iz grafa 12 in njihove lastnosti.

Slika 12 prikazuje majhno umetno omrežje prikazuje in najbolj pomembna vozlišča v njem glede na različne mere centralnosti. Iz tabele 13, ki pripada grafu na sliki 12, lahko vidimo, da vmesnostna centralnost favorizira vozlišča na sredini grafa zato je najbolj pomembno prerezno vozlišče. Bližinska centralnost favorizira vozlišča, ki so najbližje večini vozlišč in je zato podobna vmesnostni centralnosti, le da ne poudarja toliko prereznih vozlišč.

3.6.1 Omrežje sodelovanja igralcev

V omrežju sodelovanja igralcev (vir [8]) sta dva igralca povezana, če sta skupaj igrala v vsaj enem filmu. Z merami centralnosti lahko ugotovimo, kateri igralci so glede na te mere najbolj pomembni. Sledijo rezultati mer centralnosti, ki smo jih pridobili z uporabo programskega orodja SnapPy [7].

igralec	centralnost glede na stopnje
Davis, Mark	0.044918
Sanders, Alex	0.034949
North, Peter	0.034319
Marcus, Mr.	0.033459
Tedeschi, Tony	0.032142
Dough, Jon	0.031798
Stone, Lee	0.031225
Voyeur, Vince	0.030537
Lawrence, Joel	0.028647
Steele, Lexington	0.028646

Tabela 14: 10 najbolj pomembnih igralcev glede na stopnje vozlišč.

igralec	vmesnostna centralnost
Jeremy, Ron	0.000665
Jackie, Chan	0.000321
Cruz, Penelope	0.000295
Shahlavi, Darren	0.000293
Del Rosario, Monsour	0.000291
Depardieu, Gerrard	0.000275
Bachchan, Amitabh	0.000175
Jackson, Samuel L.	0.000173
Soulalem, Zinedine	0.000162
Del Rio, Olivia	0.000158

Tabela 15: 10 najbolj pomembnih igralcev glede na vmesnostno centralnost.

Iz tabele 14 lahko vidimo, da nam stopnja vozlišča kot mera centralnosti ne da rezultatov, kot bi jih želeli. Želeli bi, da so najbolj pomembni igralci iz Hollywoodske filmske industrije, kar je v tem primeru delno res, vendar je kljub temu med najbolj pomembnimi igralci dosti neznanih imen.

Vmesnostna centralnost naredi bolj pomembna vozlišča, ki so na več najkrajših poteh med ostalimi pari vozlišč. Iz seznama 15 lahko

igralec	PageRank centralnost
Hanks, Tom	0.000665
Jackson, Samuel L.	0.000647
Goldberg, Whoopi	0.000611
Stiller, Ben	0.000565
Davis, Mark	0.000551
Lopez, Jennifer	0.000546
Berland, Francois	0.000544
Berry, Halle	0.000540
Diaz, Cameron	0.000518
Travolta, John	0.000490

Tabela 16: 10 najbolj pomembnih igralcev glede na PageRank.

razberemo, da je med najbolj pomembnimi igralci veliko igralcev, ki povezujejo različne svetove. Npr. Jackie Chan povezuje kitajsko filmsko industrijo in Hollywood, Gerard Depardieu Evropo s Hollywoodom itd. Tako vidimo, da je vmesnostna centralnost za določene primere zelo smiselna mera.

PageRank nam po pričakovanjih da najbolj smiselne rezultate. Večina najbolj pomembnih igralcev po meri PageRank je znanih Hollywoodskih igralcev.

Omrežje sodelovanja igralcev je neusmerjen graf, zato se prednosti PageRanka ne pokažejo popolnoma. V primeru spleta bi dobili še boljše rezultate.

LITERATURA

- [1] M. E. J. Newman, Newman, Mark. Networks: an introduction, OUP Oxford, 2010.
- [2] A. L. Barabaasi, Network Science, Press, 2015.
- [3] R. Zafarani, M. Abbasi, H. Liu, Social Media Mining: An Introduction, Cambridge University Press, 2014.
- [4] E. Estrada, The structure of complex networks: theory and applications, OUP Oxford, 2011.
- [5] P. Bonacich, Power and centrality: A family of measures, Am. J. Sociology, 1987.
- [6] U. Brandes, A faster algorithm for betweenness centrality, Journal of mathematical sociology, 2001.
- [7] M. Culler, N. M. Dunfield, M. Goerner, and J. R. Weeks SnapPy, a computer program for studying the geometry and topology of 3-manifolds, Dostopno na <http://snappy.computop.org>, 15.5.2016.
- [8] Actor collaborations network dataset - KONECT, Dostopno na <http://konect.uni-koblenz.de/networks/actor-collaboration>, 25.5.2016.
- [9] L. Page, S. Brin, R. Motwani, T. Winograd, The PageRank citation ranking: bringing order to the web, Stanford InfoLab, 1999.

STRUKTURA SLUČAJNIH GRAFOV V ERDOS-RÉNYIJEVEM MODELU

Vid Smrke

4.1 UVOD

Večino omrežji v sodobnem svetu sestavlja ogromna količina podatkov. Klasični primeri zajemajo omrežja, ki jih tvorijo spletne strani, omrežja telefonskih klicev, socialna omrežja, omrežja letalskih linij, modeli epidemij in podobno. Veliko lastnosti teh velikih grafov, ki so plod primerov iz resničnega življenja, se da razumeti s pomočjo študije slučajnih grafov. Za velike grafe porojene naprimer iz zgornjih primerov so značilne naslednje lastnosti.

- Stopnja večine vozlišč je zelo majhna v primerjavi z velikostjo celotnega grafa; naprimer na socialnih omrežjih (facebook) – večina oseb ima majhno število prijateljev v primerjavi s številom vseh profilov na facebooku.
- V grafu je nekaj vozlišč z zelo veliko stopnjo; nekaj oseb – naprimer kakšen zelo priljubljen športnik ali glasbenik – ima zelo veliko prijateljev.
- Majhen svet (small world). Vozlišča so si zelo blizu (razdaljo med vozliščema definiramo z najmanjšim številom povezav, ki loči dani vozlišči). Če bi recimo iz celotnega omrežja facebooka izluščili vse slovenske uporabnike, bi bila ta lastnost verjetno dobro vidna – v Sloveniji praktično vsakdo pozna vsakogar ali pa vsaj pozna nekoga, ki dano osebo pozna.
- Tranzitivnost – prijatelj mojega prijatelja je moj prijatelj. To seveda implicira cikle.

Zgornje lastnosti so za grafe, ki jih porodijo nekateri primeri iz prakse, intuitivno pričakovane. Ker želimo velike grafe študirati s pomočjo slučajnih grafov, bi bilo ugodno, če bi imeli le ti tudi zgornje lastnosti, ali če bi lahko vsaj z neko verjetnostjo rekli, kdaj jih imajo in kdaj ne. Seveda je tudi od primera do primera odvisno kakšne lastnosti pričakujemo od našega grafa. Če naprimer študiramo epidemološki model, v grafu ne želimo imeti ciklov (nekdo, ki je bolezen prebolel, je bodisi umrl, bodisi je postal na bolezen imun, torej ne more še enkrat zboleti). Pogosto je torej smiselno privzeti neke omejitve. Obstajajo številni modeli slučajnih grafov. Mi se bomo posvetili najbolj osnovnemu in kronološko gledano prvemu modelu. Ogleдали si bomo nekaj njegovih osnovnih lastnosti in premislili kakšno zgradbo grafa lahko pričakujemo v odvisnosti od določenih parametrov.

4.2 ERDŐS-RÉNYIJEV MODEL

Na začetku razvoja teorije slučajnih grafov (v sredini dvajsetega stoletja) sta madžarska matematika Paul Erdős in Alfréd Rényi [3] razvila dva modela. Prvi model je slučajni graf na danem številu vozlišč in z danim številom povezav definiral kot naključno izbran element množice $G(n, m)$ vseh grafov na n vozliščih z m povezavami. Drugi model, ki je doživel precej večji razvoj in kateremu se bomo posvetili tudi mi, je vpeljan na sledeč način.

Naj bo $n \in \mathbb{N}$ in $p \in [0, 1]$. Z $G(n, p)$ označimo verjetnostni prostor, katerega dogodki so neusmerjeni grafi na n vozliščih, v katerih sta vsaki dve vozlišči povezani z verjetnostjo p neodvisno od ostalih vozlišč. Kot vemo ima graf na n vozliščih največ $N = \binom{n}{2}$ povezav. Moč množice dogodkov je torej enaka 2^N . Z $G_{n,p}$ označimo konkreten dogodek iz te množice.

Zgled 4.1. V $G(n, p)$ ima poln graf verjetnost p^N – vsak par vozlišč je povezan z verjetnostjo p , parov vozlišč je N in dogodki »vozlišči i in j sta povezani« so za $1 \leq i < j \leq N$ neodvisni. Podobno ima prazen graf verjetnost $(1 - p)^N$. Graf z m povezavami ima verjetnost $\binom{N}{m} p^m (1 - p)^{N-m}$ – izmed vseh parov vozlišč jih izberemo m , ki bodo povezana, ostala so nepovezana.

Opomba 4.2. Nekatero lastnosti slučajnih grafov lahko zaradi njihove konstrukcije obravnavamo kot slučajne spremenljivke. Stopnja posameznega vozlišča v danem modelu ima naprimer binomsko porazdelitev s parametroma $n - 1$ in p . Je namreč vsota $n - 1$ neodvisnih Bernoullijevih slučajnih spremenljivk. Vsako vozlišče ima potemtakem pričakovano stopnjo $(n - 1)p$ z varianco $(n - 1)p(1 - p)$.

Kot omenjeno že v prvem poglavju, bi želeli, da imajo slučajni grafi neke tipične lastnosti velikih grafov in morda neke specifične lastnosti, ki jih zahteva konkreten primer. Glede na konstrukcijo grafov v obravnavanem modelu se ne zdi, da bi bilo lahko napovedati lastnosti slučajnega grafa, preden bi dejansko imeli pred seboj nek konkreten graf.

4.2.1 Procesi razvejanja

Ogledali si bomo, kaj lahko povemo o nekaterih lastnostih grafa (predvsem kar se tiče njegove strukture), ko število vozlišč pošljemo proti neskončnosti. Pri tem ohranimo verjetnost p fiksno.

Razvoj rodovnih funkcij se je začel s Sirom Francsom Galtonom leta 1871 [4]. Zanimala ga je verjetnost izumrtja rodbine pri naslednjih predpostavkah. Začnemo z eno osebo. Ta oseba ima nič, enega, dva ... potomca z verjetnostmi $p_0, p_1, p_2 \dots$. Njegovi potomci bodo imeli spet slučajno število potomcev in tako naprej. Označimo z $X_0 = 1$ začetno stanje rodbine in naj bo X_k število potomcev v k -ti generaciji. Naj bo $A = \{\text{dogodek da rodbina izumre}\}$; to pomeni, da obstaja tak $k < \infty$, da je $X_k = 0$.

Zgled 4.3. V spodnjem primeru 17 je $X_0=1$, $X_1 = 2$ in $X_2 = 3$.

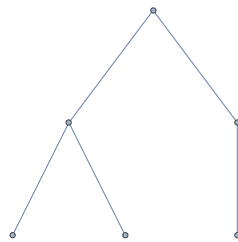


Tabela 17: Proces razvejanja

Opomba 4.4. Takemu procesu rečemo proces razvejanja.

Za to, da rodbina pri predpostavki, da ima koren k potomcev, izumre, mora izumreti vseh k podrodbin (to je poddreves). Velja torej

$$P(A|X_1 = k) = P(A)^k.$$

Vsako izmed k -tih poddreves izumre z enako verjetnostjo. Pri tem predpostavljamo neodvisnost med poddrevesi (to je disjunktnimi poddrevesi). Če v zgornji enačbi obe strani pomnožimo s $P(X_1 = k)$ in ju seštejemo po k , dobimo naslednjo enakost

$$\sum_{k=0}^{\infty} P(A|X_1 = k) \cdot P(X_1 = k) = \sum_{k=0}^{\infty} P(A)^k \cdot P(X_1 = k).$$

Po formuli za popolno verjetnost, je leva stran enaka $P(A)$, torej je

$$P(A) = \sum_{k=0}^{\infty} P(A)^k \cdot P(X_1 = k).$$

Definicija 4.5. Naj bo X nenegativna celoštevilka slučajna spremenljivka. Rodovno funkcijo $g_X(s)$ definiramo kot potenčno vrsto:

$$g_X(s) = \sum_{k=0}^{\infty} s^k P(X = k).$$

Če z η označimo $P(A)$, je verjetnost izumrtja rodbine fiksna točka ustrezne rodovne funkcije.

$$\eta = g_{X_1}(\eta), \text{ kjer je } g_{X_1}(s) = \sum_{k=0}^{\infty} p_k \cdot s^k$$

Za nadaljne sklepe moramo povedati nekaj osnovnih lastnosti rodovnih funkcij. Hitro lahko izpeljemo naslednje enakosti:

- Definirajmo slučajno spremenljivko s^X kot funkcijo slučajne spremenljivke X : $f(X) = s^X$. Rodovna funkcija $g_X(s)$ je po definiciji enaka pričakovani vrednosti slučajne spremenljivke s^X : $g_X(s) = E(s^X)$;
- $g_X(0) = 0^0 \cdot P(X = 0) = P(X = 0)$;
- $g_X(1) = E(1^X) = E(1) = 1$;
- $g'_X(1) = E(X)$, saj je $g'_X(s) = E(X \cdot s^{X-1})$;
- Če imata dve slučajni spremenljivki enako porazdiletev, imata tudi enako rodovno funkcijo.

Trditev 4.6. Naj bosta X in Y nenegativni celoštevilski slučajni spremenljivki. Če sta neodvisni, je

$$g_{X+Y}(s) = g_X(s) \cdot g_Y(s).$$

Dokaz. Dokaz temelji na enakosti $g_X(s) = E(s^X)$. Upoštevamo še predpostavko o neodvisnosti.

$$\begin{aligned} g_{X+Y}(s) &= E(s^{X+Y}) = E(s^X \cdot s^Y) = \\ &= E(s^X) \cdot E(s^Y) = g_X(s) \cdot g_Y(s). \end{aligned}$$

□

Izrek 4.7. Naj bo T nenegativna celoštevilska slučajna spremenljivka, ki je neodvisna od zaporedja neodvisnih enako porazdeljenih nenegativnih celoštevilskih slučajnih spremenljivok Z_1, Z_2, \dots . Označimo $X = \sum_{i=0}^T Z_i$ in naj bodo g_Z, g_T in g_X rodovne funkcije slučajnih spremenljivok Z_i, T in X . Potem velja

$$g_X(s) = (g_T \circ g_Z)(s) = g_T(g_Z(s)).$$

Dokaz.

$$g_X(s) = E(s^X)$$

Uporabimo formulo za popolno pričakovano vrednost.

$$\begin{aligned} E(s^X) &= \sum_{n=0}^{\infty} E(s^X | T = n) P(T = n) \\ &= \sum_{n=0}^{\infty} E(s^{Z_1 + \dots + Z_n} | T = n) P(T = n). \end{aligned}$$

Upoštevamo, da so Z_1, Z_2, \dots neodvisne od T . Tako je seveda od T neodvisna tudi njihova vsota.

$$\sum_{n=0}^{\infty} E(s^{Z_1 + \dots + Z_n} | T = n) P(T = n) = \sum_{n=0}^{\infty} E(s^{Z_1 + \dots + Z_n}) P(T = n).$$

Uporabimo prejšnjo trditev in definicijo rodovne funkcije.

$$\sum_{n=0}^{\infty} E(s^{Z_1 + \dots + Z_n}) P(T = n) = \sum_{n=0}^{\infty} [g_Z(s)]^n \cdot P(T = n) = g_T(g_Z(s)).$$

□

Zdaj, ko imamo osnovno orodje za delo z rodovnimi funkcijami, se lahko vrnemo k verjetnosti izumrtja procesa razvejanja (oziroma rod-bine). Definirajmo $Z_{k,l}$ kot število otrok l -tega vozlišča v k -ti generaciji. Tako kot za začetno vozlišče predpostavimo:

$$P(Z_{k,l} = i) = p_i \quad \text{za } i = 0, 1, 2, \dots$$

Velja torej

$$X_{n+1} = \sum_{l=1}^{X_n} Z_{n,l}.$$

Označimo $g_{X_n}(s) = g_n(s)$. Ker smo predpostavili, da je $X_0 = 1$, velja $g_0(s) = s$. Predpostavimo lahko, da je število vozlišč v n -ti generaciji neodvisno od tega, koliko imajo vozlišča v tej generaciji otrok – to je X_n je neodvisna od $Z_{n,1}, Z_{n,2}, \dots$. Iz zgornjega izreka in zveze med X_{n+1} in $Z_{k,l}$ -ji torej neposredno sledi spodnja enakost.

$$g_{n+1}(s) = g_n(g_{Z_{k,l}}(s)).$$

$Z_{k,l}$ je za vsak k in l enako porazdeljen in je porazdeljen enako kot X_1 . Označimo rodovno funkcijo X_1 kar z $g(s) = \sum_{k=0}^{\infty} p_k \cdot s^k$. Zgornjo enakost lahko prepišemo v

$$g_{n+1}(s) = g_n(g(s)).$$

Lema 4.8. *Pri zgoraj vpeljanih oznakah, velja enakost $g_{n+1}(s) = g(g_n)(s)$.*

Dokaz. Z večkratno uporabo zgornje enakosti in še nekaj poznanih enakosti dobimo

$$\begin{aligned} g_{n+1}(s) &= g_n \circ g(s) = \dots = g_0(g \circ \dots \circ g)(s) = \\ &= g_0(g^{n+1})(s) = g^{n+1}(s) = g(g_0(g^n))(s) = g(g_n)(s). \end{aligned}$$

□

Izrek 4.9 (Generating functions). $\eta = P(\text{rodbina izumre})$ je najmanjša nenegativna fiksna točka funkcije $g(s)$.

Dokaz. Naj bo $A_n = \{\text{rodbina je izumrla do } n\text{-te generacije}\} = \{X_n = 0\}$. Če je rodbina izumrla do i -te generacije, je seveda izumrla tudi do j -te generacije za vsak $j \geq i$, torej $A_1 \subset A_2 \subset \dots$ in $A = \cup_{k=1}^{\infty} A_k$. Potemtakem je

$$P(A) = \lim_{k \rightarrow \infty} P(A_k).$$

Iz ene od osnovnih lastnosti rodovnih funkcij dobimo $P(A_k) = P(X_k = 0) = g_k(0)$. Če na obeh straneh te enakosti pošljemo k v neskončnost, dobimo

$$\eta = \lim_{k \rightarrow \infty} g_k(0).$$

Upoštevamo lemo.

$$\eta = \lim_{k \rightarrow \infty} g_{k+1}(0) = \lim_{k \rightarrow \infty} g(g_k(0)).$$

Ker je g zvezna v okolici ničle, lahko limito nesemo v funkcijo.

$$\lim_{k \rightarrow \infty} g(g_k(0)) = g(\lim_{k \rightarrow \infty} g_k(0)) = g(\eta).$$

Pokažimo še, da je η najmanjša nenegativna fiksna točka funkcije g . Ker je η verjetnost, je gotovo nenegativna. Hkrati lahko hitro opazimo, da je g na nenegativnih številih gotovo nepadajoča, torej $g(0) \leq g(\eta)$. Naj bo η' neka druga nenegativna fiksna točka funkcije g .

$$g(0) \leq g(\eta') = \eta'.$$

$$g(g(0)) \leq g(\eta') = \eta'.$$

Če nadaljujemo s komponiranjem na obeh straneh, v limiti dobimo

$$(g \circ \dots \circ g \circ \dots)(0) = \lim_{k \rightarrow \infty} g_k(0) = \eta \leq \eta'.$$

□

Opomba 4.10. To, da je verjetnost izumrtja fiksna točka ustrezne rodovne funkcije, smo z osnovnimi prijemi verjetnosti pokazali že na začetku razdelka. Zdaj smo s pomočjo znanja o rodovnih funkcijah pokazali še, da je to najmanjša nenegativna fiksna točka.

Opazimo lahko, da nenegativna fiksna točka funkcije g vedno obstaja, saj je g rodovna funkcija in zanjo potemtakem velja $g(1) = 1$. Če g nima manjše fiksne točke od 1, bo torej rodbina zagotovo izumrla.

Za dokaz izreka 4.12, bomo potrebovali Taylorjev izrek [5], ki je bolj splošna oblika dobro znanega Lagrangevega izreka [6]. Povejmo ga brez dokaza.

Izrek 4.11 (Taylor). Naj bo $k \geq 1$ celo število in naj bo funkcija $f : \mathbb{R} \rightarrow \mathbb{R}$ k -krat odvedljiva v neki točki $a \in \mathbb{R}$. Potem obstaja funkcija $h_k : \mathbb{R} \rightarrow \mathbb{R}$, da velja

$$f(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \dots \\ \dots + \frac{f^{(k)}(a)}{k!}(x-a)^k + h_k(x)(x-a)^k.$$

Pri tem je $\lim_{x \rightarrow a} h_k(x) = 0$.

Definirajmo pričakovano število potomcev vsakega posameznika v drevesu.

$$\mu := \sum_{k=0}^{\infty} k \cdot p_k$$

Izrek 4.12 (Galton-Watson main theorems). Če porazdelitev slučajnih spremenljivok $Z_{k,l}$ ni oblike $p_1 = 1$ in $p_i = 0$ za vsak $i \neq 1$, ima enačba $g(s) = s$ eno ali dve rešitvi. Če je $\mu > 1$, obstaja enolična rešitev $s = \eta < 1$. Če je $\mu \leq 1$, je edina rešitev $s = 1$.

Opomba 4.13. Upoštevajoč prejšnji izrek, to pomeni, da je verjetnost izumrtja gotova (to je, je enaka 1) natanko tedaj, ko je $\mu \leq 1$ oziroma, da je verjetnost izumrtja manjša od 1 natanko tedaj, ko je $\mu > 1$.

Dokaz. Osredotočimo se na interval $[0, 1]$. Z izračunom drugega odvoda funkcije g lahko hitro preverimo, da je g konveksna na $[0, 1]$

$$g''(s) = \sum_{k=0}^{\infty} k(k-1)s^{k-2} \cdot p_k > 0 \text{ za } s \in [0, 1].$$

Očitno je funkcija g tudi strogo naraščajoča in prav tak je tudi njen odvod, saj po predpostavki izvzamemo edino porazdelitev števila potomcev, ki bi to preprečevala. Po eni izmed osnovnih lastnosti rodovnih funkcij, dobimo enakost $g'(1) = \sum_{k=0}^{\infty} k \cdot 1^{k-1} \cdot p_k = \mu$. Oglejmo

si najprej primer, ko je $\mu \leq 1$. Denimo, da obstaja $0 \leq a < 1$, da je $g(a) = a$. Ker je g zvezna v a in $\mathbb{1}$ ter odvedljiva v vsaki točki na $(a, 1)$, po Lagrangeovem izreku obstaja točka $b \in (a, 1)$, da je $g(1) - g(a) = g'(b)(1 - a)$. Upoštevši, da sta a in $\mathbb{1}$ fiksni točki funkcije g , dobimo $g'(b) = 1$. To pa ni v redu, saj je g' strogo naraščajoča in $\mu = g'(1) \leq 1$ po predpostavki.

Denimo zdaj, da je $\mu > 1$. Če je $p_0 = 0$, ima enačba $g(s) = s$ rešitvi $s = 1$ in $s = 0$. Ker je g strogo konveksna, drugih rešitev ni. Naj bo torej $g(0) = p_0 > 0$. Uporabimo Taylorjev izrek za funkcijo g in upoštevamo, da je $s \in (0, 1)$.

$$\begin{aligned} g(s) &= g(1) + g'(1)(s - 1) + \dots \\ &= g(1) + \mu(s - 1) + \dots < g(1) + (s - 1) = s, \end{aligned}$$

za nek s dovolj blizu $\mathbb{1}$ – tako blizu, da je morebitni negativni del zanemarjenega ostanka Taylorjeve vrste po absolutni vrednosti manjši od vsote pozitivnega zanemarjenega dela Taylorjeve vrste in razlike, ki jo pridemo z oceno μ -ja. Tedaj torej velja strogi neenačaj v zgornji enačbi. Označimo ta dovolj velik s z s^* . Velja torej $g(0) - 0 > 0$ in $p(s^*) - s^* < 0$ ter $s^* \in (0, 1)$. Ker je $g(s) - s$ zvezna, obstaja $t \in (0, s^*)$, da je $g(t) - t = 0$. \square

4.2.2 Povezava procesov razvejanja in slučajnih grafov

Vrnimo se zdaj k slučajnim grafom in poskusimo razbrati, kaj smo v bistvu ugotovili. V primeru Erdős-Rényijevega modela je graf, ki je element množice $G(n, p)$, skonstruiran tako, da vzamemo n vozlišč in med vsakima dvema vozliščema naredimo povezavo z verjetnostjo p neodvisno. Če si torej izberemo nek koren (neko konkretno vozlišče), lahko konstrukcijo enačimo s procesom razvejanja. Koren bo imel z verjetnostjo $(1 - p)^{n-1}$ nič otrok, z verjetnostjo $p(1 - p)^{n-2}$ enega in tako naprej. Enako velja za vse njegove morebitne otroke. V tem primeru je

$$\mu = \sum_{k=0}^n k \cdot p^k (1 - p)^{n-k}.$$

Oziroma ko pošljemo število vozlišč proti neskončnosti

$$\mu = \sum_{k=0}^{\infty} k \cdot p^k (1 - p)^{n-k}.$$

Če je $\mu \leq 1$, se bo proces razvejanja ustavil z verjetnostjo $\mathbb{1}$. Slučajni graf bo torej za majhne p sestavljen iz disjunktnih dreves. Erdős in Rényi sta pokazala [7], da to velja za $0 < p \ll 1/n$. Če je $\mu > 1$,

se bo proces razvejanja končal z verjetnostjo manjšo od 1. V tem primeru lahko torej pričakujemo, da se bodo v grafu začela pojavljati vse večja drevesa in sčasoma tudi cikli oziroma splošne velike povezane komponente.

4.2.3 Zgled – epidemološki model

Oglejmo si epidemološki model Reeda in Frosta [8]. Predpostavimo, da imamo populacijo velikosti n .

- Na začetku je okužena ena oseba.
- Oseba, ki je okužena bolezen bodisi preživi in postane imuna, bodisi zaradi bolezni umre – v vsakem primeru je odstranjena iz modela.
- Zdrave osebe, se od okužene osebe nalezejo z verjetnostjo p , neodvisno od drugih okužb.

Ta model lahko obravnavamo s pomočjo Erdős-Rényijevih grafov. Pri izbruhu epidemije nas zanima predvsem nejna pričakovana razsežnost (to je kako hitro se bo bolezen širila in koliko bo okuženih). V skladu s tem namreč zdravstvene ustanove sprejemajo ukrepe za njeno zaježitev. Na podlagi dosedanjega širjenja bolezni in podatkov iz preteklosti lahko določimo verjetnost p , da se posameznik naleze od okuženega osebka. To je torej verjetnost v Erdős-Rényijevem modelu. Razsežnost epidemije bo enaka velikosti povezane komponente v $G(n, p)$; ta bo ena sama, saj smo po predpostavki začeli z eno okuženo osebo. Če bo $p < 1/n$ se bo okužilo majhno število oseb – proces razvejanja se bo ustavil.

4.2.4 Zaključek

Ogledali smo si najosnovnejši model slučajnih grafov. Čeprav se na podlagi vpeljave modela zdi, da o njegovi strukturi ne moremo povedati veliko, nam rodovne funkcije omogočajo pridobiti kar nekaj informacij. Konstrukcijo slučajnega grafa lahko namreč enačimo s procesom razvejanja. Na podlagi rezultatov o rodovnih funkcijah, smo zmožni napovedati kako povezan bo nek slučajni graf in kakšne bodo njegove povezane komponente (to je naprimer ali bo graf gozd in ali bo morda vseboval cikle). Na koncu smo si ogledali še primer uporabe izpeljanih rezultatov.

LITERATURA

- [1] Anne Bouillard, *Random graphs*, lecture notes, <http://www.di.ens.fr/~bouillard/PACS/randomGraphs.pdf>.
- [2] Fan Chung, *A whirlwind tour of random graphs*, a survey article in *Encyclopedia on Complex Systems*, Springer, 2008, <http://www.math.ucsd.edu/~fan/wp/randomg.pdf>.
- [3] P. Erdős, A. Rényi, *On random graphs*, I. *Publ. Math. Debrecen* 6, 290-297, 1959.
- [4] H. W. Watson, F. Galton, *On the Probability of the Extinction of Families*, *Journal of the Anthropological Institute of Great Briatin*, volume 4, 138-144, 1875.
- [5] T. Apostol, *Mathematical analysis*, Addison-Wesley, 361-366, 1974.
- [6] H. Jeffreys, B. Swirles, *Methods of mathematical physics*, 3rd edition, Cambridge University Press, 49-50, 1988.
- [7] P. Erdős and A. Rényi, *On the evolution of random graphs*, *Magyar Tud. Akad. Mat. Kutató Int. Közl.* 5, 17-61, 1960.
- [8] H. Abbey, *An examination of the Reed-Frost theory of epidemics*, *Hum. Biol.*, 1952.

PROBLEM ŠTIRIH BARV

Larisa Gostenčnik

5.1 UVOD

Leta 1852 je južnoafriški matematik Francis Guthrie barval zemljevid angleških grofij tako, da sta bili grofiji, ki imata skupno mejo, pobarvani z različnima barvama. Pri tem je ugotovil, da za barvanje celotnega zemljevida zadostujejo že samo štiri barve. To je v pismu opisal tudi svojemu bratu Fredericku, ki je bil študent Augustusa de Morgana na univerzi v Londonu, ki mu je problem tudi predstavil. Tako je nastal znameniti izrek štirih barv, ki pravi, da so štiri barve dovolj, da pobarvamo poljuben zemljevid, kjer so sosednje komponente zemljevida različnih barv. Poglavje je strukturirano tako, da si v prvem podrazdelku 5.2 ogledamo podoben problem, ki pravi, da lahko poljuben zemljevid pobarvamo s petimi barvami. V drugem podrazdelku 5.3 definiramo pojem neizogibne konfiguracije in se seznamimo z novo idejo, kako dobiti neizogibno množico konfiguracij. V tretjem podrazdelku 5.4 definiramo reducibilnost konfiguracij, obravnavamo C in D-reducibilnost ter dokažemo, da je Birkhoffov diamant C-reducibilen. V četrtem, zadnjem podrazdelku 5.5, pa predstavimo idejo dokaza izreka o štirih barvah bolj podrobno.

5.1.1 *Problem štirih barv*

Naša naloga je, da pobarvamo zemljevid s samo štirimi barvami. Očitno je videti, da so tri barve premalo za takšno barvanje. Prav tako je dokaj enostavno pokazati, da je pet barv dovolj, kar bomo pokazali v nadaljevanju. Pred končnim dokazom samega izreka o štirih barvah je obstajalo veliko nepravilnih verzij le tega. Eno izmed njih je leta

1879 napisal Kempe in je kar 11 let veljala za pravilno. Drugo je leta 1890 predstavil Tait in je veljala za pravilno 1 leto, dokler je ni ovrigel Tutte.

5.1.2 Prevedba na problem v teorijo grafov

Problem štirih barv se da prevesti tudi v teorijo grafov in ga z njeno pomočjo tudi dokazati. Zemljevidi so ravninski grafi, zato so dualni grafi zemljevidov tudi ravninski grafi. Tem grafom nato barvamo vozlišča, ki nadomestijo lica zemljevida in želimo, da je kromatično število novih grafov manjše ali enako 4. Dual ravninskega grafa je ravninski, vendar pa ni nujno enostavni. Če odstanimo večkratne povezave in zanke se kromatično število grafa ne spremeni, zato obravnavamo le enostavne grafe. Prav tako se bomo osredotočili le na triangulacije, to so ravninski grafi, katerih lica so trikotniki. To naredimo zato, ker lahko poljuben enostavni graf na vsaj treh točkah dopolnimo do triangulacije z dodajanjem novih povezav. Če bomo vozlišča tako dobljene triangulacije lahko pobarvali s štirimi barvami, bomo s štirimi barvami lahko pobarvali tudi osnovni graf, saj se z brisanjem povezav samo zmanjšuje število omejitev pri barvanju. Omenimo še očitno dejstvo, da so grafi z manj kot tremi točkami 4-obarvljivi.

5.2 PROBLEM PETIH BARV

Trditvev 5.1. Naj bo n_k število vozlišč stopnje k v triangulaciji T . Potem velja

$$\sum_{k \geq 2} (6 - k)n_k = 12.$$

Dokaz. Vemo, da velja $n = \sum_{k \geq 2} n_k$. Dalje, če je d_i stopnja vozlišča i , $i = 1, 2, \dots, n$, potem velja $2e = \sum_{i=1}^n d_i$, kjer je e število povezav grafa T . Predpostavimo, da velja $d_i = k$ za vse $i \in I$ in $d_i \neq k$ za vse druge i . To pomeni, da je I množica vseh vozlišč stopnje k v triangulaciji T . Potem velja enakost $\sum_{i \in I} d_i = k|I|$. Ker velja $|I| = n_k$, sledi $\sum_{i=1}^n d_i = \sum_{k \geq 2} kn_k$. V triangulaciji T je vsako lice dolžine 3. Če seštejemo vse povezave vseh lic, dobimo vsoto $3f$. V tem izračunu je vsaka povezava šteta dvakrat, saj je vsaka povezava na robu dveh lic. Torej je $3f = 2e$ in dobimo $3f = \sum_{k \geq 2} kn_k$. Ob uporabi Eulerjeve formule $n - e + f = 2$ dobimo:

$$\sum_{k \geq 2} n_k - \frac{1}{2} \sum_{k \geq 2} kn_k + \frac{1}{3} \sum_{k \geq 2} kn_k = 2.$$

Ko to formulo poenostavimo, dobimo željeno formulo iz trditve. \square

Posledica 5.2. Vsaka triangulacija T vsebuje vsaj eno vozlišče stopnje 2, 3, 4 ali 5.

Dokaz. Iz zgornje trditve vemo, da velja enačba $\sum_{k=2}^5 (6-k)n_k - \sum_{k \geq 6} (k-6)n_k = 12$. Če prenesemo drugo vsoto na desno stran enačbe, dobimo $\sum_{k=2}^5 (6-k)n_k = 12 + \sum_{k \geq 6} (k-6)n_k$. Desna stran enačbe je vedno pozitivna, zato mora biti tudi leva. Torej mora biti $n_k \geq 1$ za vsaj en $k = 2, 3, 4, 5$. \square

Sedaj lahko s pomočjo zgornje posledice dokažemo izrek o petih barvah.

Izrek 5.3. Vsaka triangulacija je 5-obarvljiva.

Dokaz. Izrek dokazujemo s protislovjem. Predpostavimo, da vsaka triangulacija ni 5-obarvljiva. T naj bo tak minimalni protiprimer. Po prejšnji posledici vemo, da T vsebuje vsaj eno vozlišče stopnje 2, 3, 4 ali 5. Poglejmo si primer za vsako vozlišče posebej:

- T vsebuje vozlišče v stopnje 2. Ker je T enostavni graf, je izomorfen K_3 in je očitno 5-obarvljiv.
- T vsebuje vozlišče v stopnje 3. $T - v$ je triangulacija, ki je očitno 5-obarvljiva. Pobarvamo jo lahko s tremi barvami, vozlišče v pa s četrto barvo in tako imamo 5-barvanje grafa T in nimamo minimalnega protiprimera.
- T vsebuje vozlišče v stopnje 4. $T - v$ ni triangulacija, zato mu dodamo dodatno povezavo. $T - v$ z dodano povezavo lahko pobarvamo s štirimi barvami, vozlišče v pa s peto barvo in tako T ima 5-barvanje in ni minimalni protiprimer.

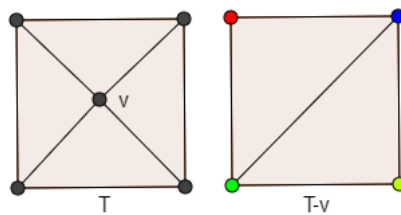


Tabela 18: Na sliki vidimo primer barvanja za graf $T - v$, kjer je v vozlišče stopnje 4.

- T vsebuje vozlišče v stopnje 5. Vozlišča grafa $T - v$ so pobarvana z različnimi barvami $c_1 c_2 c_3 c_4 c_5$. Radi bi se znebili ene barve. Pogledamo si podgraf $H_{c_1 c_3}$ grafa $T - v$. To je graf induciran z vozlišči, pobarvanimi z barvama c_1 in c_3 . Točki u_1 in u_3 sta na isti povezani komponenti grafa $H_{c_1 c_3}$, če obstaja v v $T - v$ pot od

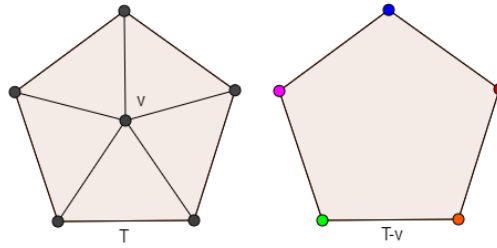


Tabela 19: Na sliki vidimo primer barvanja za graf $T - v$, kjer je v vozlišče stopnje 5.

u_1 do u_3 , ki poteka le po vozliščih pobarvanih z barvama c_1 in c_3 . Tej poti rečemo veriga. Če u_1 in u_3 nista na isti povezani komponenti grafa $H_{c_1c_3}$, lahko vozlišču u_3 zamenjamo barvo in ga pobarvamo z barvo c_1 . Tako imamo 4-barvanje grafa $T - v$ in s tem tudi 5-barvanje grafa T . Podobno velja za točki u_2, u_5 in podgraf $H_{c_2c_5}$. Do težave pridemo, če obstajata hkrati verigi c_1c_3 in c_2c_5 , kar vidimo na sliki 3. T je triangulacija, zato obstaja

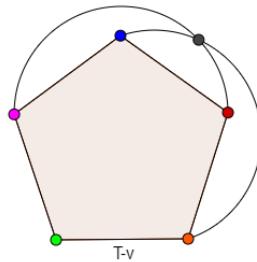


Tabela 20: Verigi c_1c_3 in c_2c_5 .

takšna vložitev grafa $T - v$ v ravnino, da $T - v$ nima povezav znotraj cikla. Verigi c_1c_3 in c_2c_5 se morata nekje sekati. Ker je $T - v$ ravninski, se morata sekata v vozlišču. To ni možno, saj bi potem moralo biti to vozlišče pobarvano z dvema različnima barvama. Torej obe verigi ne moreta hkrati obstajati in eni točki na grafu $T - v$ lahko vedno spremenimo barvo in tako dobimo 4-barvanje. Vozlišče v pobarvamo s peto barvo in imamo 5-barvanje.

□

Te verige imenujemo Kempejeve verige. Kempejev dokaz izreka o štirih barvah je deloval na istem principu verig kot dokaz o petih barvah. Pri vozliščih stopnje 2 in 3 je dokaz trivialen, pri vozliščih stopnje 4 uporabimo argument z verigami, zaplete pa se pri vozliščih

stopnje 5. Vemo, da za cikel okrog vozlišča v zadostujejo 4 barve. Zanima nas, kaj naredimo, če pride do situacije, kjer se v ciklu ne moremo znebiti še ene dodatne barve. Imamo verigi u_2u_5 in u_2u_4 . Ker so vozlišča u_2 , u_4 in u_6 povezana z verigami ne moremo izločiti nobene izmed barv c_2 , c_4 in c_6 . Opazimo, da med u_3 in u_5 zagotovo ni verige, zato lahko vozlišče u_3 pobarvamo z barvo c_5 . Prav tako med vozlišči u_1 in u_3 ni verige, zato lahko u_1 pobarvamo z barvo c_4 . Tako smo se na 5-ciklu znebili barve c_1 in c_3 . Vendar verigi med vozlišči u_2u_5 in u_2u_4 vsebujeta isto barvo, zato se lahko v sekata v vozliščih pobarvanih s c_2 . Tako lahko dobimo novo verigo med vozliščema u_1 in u_4 , ki sta obe pobarvani z barvo c_4 , kar nam onemogoči zamenjavo barve c_1 na u_1 . Tako ne moremo dobiti 4-barvanja grafa T . Vidimo, da je Kempejev dokaz za izrek štirih barv res napačen. Kljub temu, pa se njegova ideja o verigah uporablja v pravilnem dokazu. Da bi ga lahko dokazali pa se moramo seznaniti z novimi idejami in pojmi, ki so uporabljeni v končnem, pravilnem dokazu.

5.3 NEIZOGIBNE KONFIGURACIJE

Definicija 5.4. *Konfiguracija* v triangulaciji je del triangulacije, ki leži v nekem ciklu. Število vozlišč v tem ciklu imenujemo velikost cikla.

Definicija 5.5. Množico konfiguracij imenujemo *neizogibna*, če v vsaki triangulaciji nastopa vsaj ena izmed konfiguracij te množice.

Vemo že, da je množica vozlišč stopnje 2, 3, 4 in 5 neizogibna množica konfiguracij. Cilj dokaza izreka o štirih barvah je bil poiskati takšno neizogibno množico konfiguracij, da bi lahko za vsak element te množice pokazali, da ne nastopa v minimalnem protiprimeru, kar pomeni, da tak protiprimer ne obstaja. To smo naredili že za primer 5-barvanja. Za vozlišča stopnje manj od 5 je to enostavno pokazati, za vozlišča stopnje 5 pa je že precej zahtevno. Zato bi bilo potrebno poiskati konfiguracije, ki bi jo nadomestile. Taka množica bi morala biti neizogibna.

Nova ideja, kako dobimo neizogibno množico konfiguracij se imenuje prenašanje naboja, ki si jo je zamislil Heesch. Da dokažemo, da je množica neizogibna, dokazujemo s protislovjem. Predpostavimo, da obstaja graf, ki ne vsebuje nobene konfiguracije iz te množice. Vozliščem grafa priredimo števila, ki jih poimenujemo naboji. Vsota vseh nabojev v grafu je pozitivno število. V nadaljnjem prenašamo naboje med sosednjimi vozlišči tako, da se celoten naboj ohranja. Če lahko ob predpostavki, da triangulacija ne vsebuje nobene konfiguracije iz naše množice poiščemo tak postopek prenašanja naboja, da v našem grafu najdemo konfiguracijo iz začetne množice konfiguracij,

smo prišli v protislovje. Tako je začetna množica resnično neizogibna množica konfiguracij.

5.4 REDUCIBILNOST

Definicija 5.6. Konfiguracija je *reducibilna*, če ne more biti vsebovana v nobenem minimalnem protiprimeru.

V primeru Kempejevih verig smo videli kako reduciramo vozlišča stopnje 2, 3, 4 in 5 v primeru 5-barvanja. V nadaljevanju bomo pokazali kako reduciramo bolj zapletene konfiguracije. Predstavili bomo C-reducibilnost in D-reducibilnost in si ta pojma pogledali na primeru Birkhoffovega diamanta.

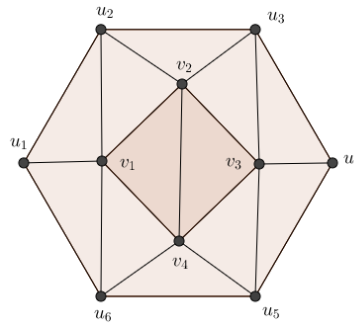


Tabela 21: Birkhoffov diamant

Najprej bi se lotili naivne metode za določanje reducibilnosti. Osredotočimo se na cikel, ki vsebuje konfiguracijo. To je cikel $u_1u_2u_3u_4u_5u_6$. Ker ne vemo, kako izgleda triangulacija izven tega cikla, preverimo vsa možna barvanja pri 4-barvanju cikla $T - C$. Če se da T pobarvati s 4 barvami, potem ta konfiguracija je reducibilna. Naj bo T' triangulacija brez konfiguracije, $T' = T - C$. Barvamo T' : $c_2c_1c_2c_3c_4c_3$. Sedaj moramo pobarvati še samo konfiguracijo znotraj cikla, torej vozlišča v . Pobarvamo jih z barvami $c_1c_3c_4c_2$ in imamo 4-barvanje Birkhoffovega diamanta. Obstajajo pa barvanja cikla, ki jih ne moremo prenesti na cel T . Takšno barvanje je $c_2c_1c_3c_1c_3c_1$. Vozlišči v_2 in v_4 sta obe povezani z vozlišči, ki sta pobarvani z barvama c_1 in c_3 . Iz tega sledi, da morata biti pobarvani z barvama c_2 ali c_4 . Ker sta vozlišči v_2 in v_4 sosedni, lahko brez škode za splošnost pobarvamo v_2 s c_2 in v_4 s c_4 . Tako je vozlišče v_3 sosedno z vozlišči, ki so pobarvane z vsemi štirimi barvami, zato ga moramo pobarvati s peto barvo in nimamo 4-barvanja. Na tem mestu si pomagamo s Kempejevimi verigami. Pogledamo si podgraf $H_{c_1c_4}$, induciran z vozlišči, pobarvanimi z barvama c_1 in c_4 . Obravnavamo tri različne možnosti:

- Vozlišče u_4 je v različni komponenti kot u_2 in u_6 . Vozlišču u_4 tako lahko spremenimo barvo in ga pobarvamo s c_4 . Potem so vozlišča v pobarvana z barvami $c_3c_2c_1c_4$ in imamo 4-barvanje grafa T .
- Vozlišče u_4 je v isti komponenti kot u_2 , a v drugi komponenti kot u_6 . Tako lahko vozlišču u_6 spremenimo barvo in ga pobarvamo z barvo c_4 in imamo 4-barvanje grafa T .
- Vozlišča u_2, u_4, u_6 so vsa v isti komponenti. Potem u_3 ne more biti v isti komponenti kot u_1 ali u_5 . Tako lahko zamenjamo barvo vozlišča u_3 in ga pobarvamo z barvo c_2 in vozlišča v pobarvamo z barvami $c_4c_3c_4c_2$. Spet imamo 4-barvanje.

Če je konfiguracija reducibilna preko naivne metode ali Kempejevih verig, rečemo da je D-reducibilna. Testiranje reducibilnosti na ta način je sicer preprosto, a lahko ima veliko časovno zahtevnost. Velja tudi, da niso vse konfiguracije D-reducibilne. Tako lahko preverjamo reducibilnost še na tretji način. Iz T odstranimo konfiguracijo C , za katero želimo pokazati, da je reducibilna. Nato v $T - C$ identificiramo določena vozlišča v ciklu konfiguracije in dodamo nove povezave, da dobimo novo triangulacijo T' . Za to triangulacijo T' bo veljalo, da lahko njeno poljubno 4-barvanje razširimo na $T - C$. Če se to 4-barvanje razširi na T vemo, da je C reducibilna. Konfiguracije, ki so reducibilne na ta način imenujemo C-reducibilne in pojem C-reducibilnost. Poglejmo to idejo na primeru Birkhoffovega diamanta. Na sliki 5 odstranimo konfiguracijo in nato identificiramo vozlišča u_3 in u_5 ter dodamo povezavo med u_1 in u_3 , da dobimo triangulacijo. Na sliki

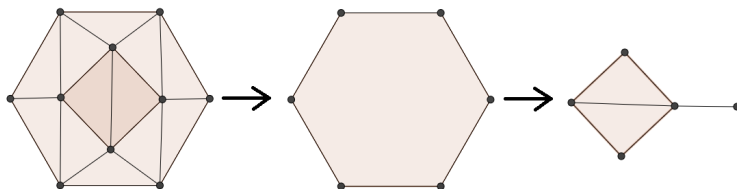


Tabela 22: Slika nam prikazuje, kako dobimo novo triangulacijo.

6 pobarvamo T' s štirimi barvami in nato prenesemo to barvanje na cikel, ki vsebuje konfiguracijo in nato na cel Birkhoffov diamant.

Lema 5.7. *Birkhoffov diamant je C-reducibilen.*

Dokaz. Obstaja 6 različnih barvanj vozlišč $u_1u_2u_3u_5u_4u_6$ v grafu $T' = T - C$. Ta so: $c_1c_2c_3c_1c_2$, $c_1c_2c_3c_2c_2c_2$, $c_1c_2c_3c_4c_2$, $c_1c_2c_3c_4c_4$, $c_1c_2c_3c_1c_4$, $c_1c_2c_3c_2c_4$. Za vsa barvanja se da enostavno pokazati, da se jih da razširiti na 4-barvanje na T . Problem nastane pri barvanju $c_1c_2c_3c_2c_2$. Ker

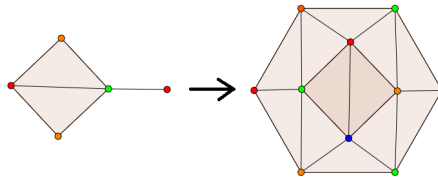


Tabela 23: Prenos barvanja triangulacije na Birkhoffov diamant.

ne gre na naiven način producirati 4-barvanja, se tega lotimo s pomočjo Kempejevih verig. Gledamo podgraf $H_{c_2c_4}$, induciran z barvama c_2 in c_4 . Spet obravnavamo različne možnosti. Če so u_2, u_4, u_6 vse v isti komponenti, potem u_1 in u_5 spadata k različnim komponentam $H_{c_1c_3}$ in lahko vozlišče u_5 pobarvamo z barvo c_1 . Vozlišča v tako pobarvamo s $c_4c_1c_4c_3$. Tako dobimo barvanje vozlišč u v Birkhoffovem diamantu $c_1c_2c_3c_2c_1c_2$ in njegovo 4-barvanje. Če u_2 ni v isti komponenti kot u_4 in u_6 , lahko u_2 pobarvamo s c_4 in vozlišča v s $c_3c_2c_1c_4$. Tako dobimo 4-barvanje. Podobno obravnavamo primer, ko u_6 ni v isti komponenti kot u_2 in u_4 . Če vozlišče u_4 ni v isti komponenti kot u_2 ali u_6 , potem pobarvamo u_2 in u_6 z barvo c_4 in dobimo barvanje vozlišč v s $c_3c_2c_4c_1$. Tako dobimo barvanje vozlišč u s $c_1c_4c_3c_2c_1c_4$. Spet smo dobili 4-barvanje Birkhoffovega diamanta in s tem tudi njegovo C-reducibilnost. \square

Očitno je, da lahko C-reducibilnost testiramo hitreje kot D-reducibilnost, ker imamo z identifikacijo vozlišč konfiguracije manj možnih kombinacij barvanja.

5.5 PROBLEM ŠTIRIH BARV

Seznanili smo se z idejami, ki so bile potrebne za dokaz izreka o štirih barvah, katerega sta zapisala Appel in Haken s pomočjo računalnika ter objavila leta 1976. Sedaj pa predstavimo idejo dokaza bolj podrobno.

Izrek 5.8. *Vsaka triangulacija je 4-obarvljiva.*

Dokaz. Osnovna ideja dokaza je enaka kot pri dokazu problema petih barv. Imamo neizogibno množico konfiguracij in za vsako izmed teh konfiguracij pokažemo, da je neizogibna. V dokazu sta upoštevala, da velikost cikla, ki vsebuje konfiguracijo, ni večja od 14, saj je bila za večje cikle časovna zahtevnost prevelika. Vedela pa sta, da mora biti velikost cikla vsaj 12, saj je E.H.Moore našel triangulacijo, v kateri je imela najmanjša C-reducibilna konfiguracija cikel velikosti 12 (kasneje se je izkazalo, da se je Moore motil). Nato je bilo potrebno producirati neizogibno množico konfiguracij. To je bilo narejeno z računalnikom

z metodo prenašanja naboja. Predpostavila sta, da množica ni neizogibna in za vsako triangulacijo T brez vozlišč stopnje manjše od 5 predpisala naboj $60(6 - \deg(v))$ za vsako vozlišče v in tako je bil z upoštevanjem trditve iz začetka skupen naboj v grafu 720. Nato sta s prenašanjem naboja skušala priti v protislovje in našla primerno kandidatko. Končna neizogibna množica je tako imela 1879 konfiguracij. Potem je ta neizogibna množica konfiguracij morala biti preverjena za reducibilnost. Najprej sta jo preverila za D-reducibilnost. Če je konfiguracija padla na tem testu, sta jo preverila še za C-reducibilnost. Ko je bila vsaka konfiguracija preverjena za reducibilno, je bil izrek dokazan. Če pa je konfiguracija padla na testu za reducibilnost ali test ni bil končan v zahtevanem času, saj sta le tega omejila, potem je bilo potrebno popraviti neizogibno množico konfiguracij in cel postopek ponoviti od začetka. \square

5.6 ZAKLJUČEK

Tako je bil problem štirih barv dokazan s pomočjo računalnika 124 let po tem, ko je bil prvič predstavljen Augustusu de Morganu. Dokaz izreka je zelo dolg in ga nekateri matematiki še danes ne priznavajo za pravičnega, saj je dokazan s pomočjo računalnika. Zaradi njegove dolžine so se Robertson, Sanders, Seymour in Thomas odločili napisati svoj dokaz, ki je bil objavljen leta 1996. Ideja je podobna kot pri dokazu Appela in Hakna, poenostavili so algoritem prenašanja naboja in tako zmanjšali moč neizogibne množice na 633 konfiguracij. S tem so zmanjšali tudi časovno zahtevnost algoritma za 4-barvanje grafov.

LITERATURA

- [1] D.A. Holton in J. Sheehan *Australian Mathematical Society Lecture Series*. 7, (1993) 49–74.
- [2] *Four-color theorem*, [ogled 7. 4. 2016], dostopno na <http://mathworld.wolfram.com/Four-ColorTheorem.html>.
- [3] *The Four Color Theorem*, [ogled 7. 4. 2016], dostopno na <http://people.math.gatech.edu/~thomas/FC/fourcolor.html>.
- [4] *Izrek štiri barv*, v: Wikipedia: The Free Encyclopedia, [ogled 7. 4. 2016], dostopno na https://sl.wikipedia.org/wiki/Izrek_%C5%A1tiri_barv.
- [5] *Four color theorem*, v: Wikipedia: The Free Encyclopedia, [ogled 7. 4. 2016], dostopno na https://en.wikipedia.org/wiki/Four_color_theorem.

Pojem homomorfizma srečamo v algebri kot preslikavo, ki ohranja operacijo. Podobno je tudi pri grafih – homomorfizem je preslikava med grafoma, ki ohranja relacijo povezanosti vozlišč. S pomočjo homomorfizmov lahko preučujemo različne lastnosti grafov, na primer barvanja, ožine in lastnosti produktov grafov.

V nadaljevanju najprej predstavimo osnovne definicije in nekaj ilustrativnih zgledov. V razdelku 6.2 spoznamo relacijo, ki jo podajajo homomorfizmi. Nato si ogledamo lastnosti retrakta, ki je le posebna vrsta homomorfizma. V razdelku 6.4 pridobljeno znanje povežemo z barvanji grafov, nato pa v zadnjem razdelku spoznamo še pojem jedra. Vsebina je večinoma povzeta po [1].

6.1 OSNOVNI POJMI

Oglejmo si formalno definicijo homomorfizma, kot ga je prvi študiral Sabidussi v pedesetih in zgodnjih šestdesetih letih prejšnjega stoletja ([6]).

Definicija 6.1. Naj bosta G in H grafa. Preslikava $f: V(G) \rightarrow V(H)$ je *homomorfizem*, če velja $uv \in E(G) \implies f(u)f(v) \in E(H)$.

V nadaljevanju pogosto zlorabimo notacijo in pišemo homomorfizem kar kot $f: G \rightarrow H$.

Spomnimo se še nekaj drugih definicij.

Definicija 6.2. Naj bosta G in H grafa. Preslikava $f: V(G) \rightarrow V(H)$ je *izomorfizem*, če velja $uv \in E(G) \iff f(u)f(v) \in E(H)$. Če je še dodatno $G = H$, preslikavi pravimo *avtomorfizem*.

Pravi podgraf grafa G je podgraf, ki ni prazen in hkrati ni enak grafu G .

Relacijo povezanosti v grafu G označimo kot $u \sim_G v$ in pomeni, da je $uv \in E(G)$.

Opomba 6.3. Direktno iz definicij sledi naslednje:

- Vsak izomorfizem ali avtomorfizem grafov je tudi homomorfizem.
- Če grafa G in H nimata zank, potem za vsak homomorfizem $f: G \rightarrow H$ sledi: $uv \in E(G) \implies f(u) \neq f(v)$.

Ves čas obravnavamo enostavne (in končne) grafe.

Zgled 6.4. Da si homomorfizme lažje predstavljamo, si oglejmo nekaj zgledov.

- Vložitev podgrafa v graf je vedno homomorfizem.
- Naj bo G dvodelni graf in $V(G) = A \cup B$ ustrezna particija vozlišč (A, B disjunktni in neprazni). Tedaj je preslikava $f: G \rightarrow K_2$, definirana kot

$$f(u) = \begin{cases} 1, & \text{če } u \in A \\ 2, & \text{če } u \in B, \end{cases}$$

kjer K_2 enačimo s povezavo z vozliščema 1 in 2. Preverimo, da je to res homomorfizem. Če je $uv \in E(G)$, potem je (brez škode za splošnost) $u \in A, v \in B$. Sledi, da je $f(u) = 1, f(v) = 2$ in torej $f(u)f(v) \in E(K_2)$.

Kadar graf G vsebuje vsaj eno povezavo, lahko definiramo tudi homomorfizem v obratni smeri, to je $g: K_2 \rightarrow G$, kot $g(1) = a, g(2) = b$, kjer sta $a \in A, b \in B$ ter $ab \in E(G)$.

Med grafoma G in K_2 torej obstajata homomorfizma v obe smeri, vendar si grafa nista izomorfna (čim ima G vsaj 3 vozlišča).

- Označimo vozlišča cikla C_7 zaporedoma z A, B, C, D, E, F, G in vozlišča cikla C_5 zaporedoma z A, B, C, D, E . Oglejmo si preslikavo $h: C_7 \rightarrow C_5$, definirano kot

$$\begin{aligned} h|_{\{A,B,C\}} &= \text{id} \\ h(D) &= h(F) = D \\ h(E) &= h(G) = E. \end{aligned}$$

Po definiciji hitro vidimo, da je h homomorfizem. Analogno bi lahko definirali homomorfizem $C_{n+2} \rightarrow C_n$ za splošen $n \geq 3$.

Vprašamo se lahko, ali v tem zadnjem primeru obstaja homomorfizem tudi v obratni smeri, tj. $C_n \rightarrow C_{n+2}$. Občutek nam pravi, da to ne bo mogoče, ker bi morali manjši cikel nekako "pretrgati". Oglejmo si problem kar v splošnem. Najprej potrebujemo naslednjo definicijo.

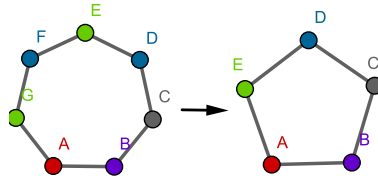


Tabela 24: Homomorfizem med različno velikima cikloma.

Definicija 6.5. *Ožina* grafa je dolžina najkrajšega cikla vsebovanega v njem. Če graf ne vsebuje nobenega cikla, je njegova ožina neskončno. Podobno definiramo tudi *sodo* in *liho ožino* kot dolžino najkrajšega sodega oz. lihega cikla v grafu.

V primeru cikla je ožina seveda enaka njegovi dolžini. Ožina poljubnega polnega grafa na $n > 2$ točkah je tri, ker vedno vsebuje trikotnik. V drevesih ni nobenega cikla, zato je njihova ožina enaka neskončno. Kot vemo, je ožina invarianta grafa. Sedaj pa posplošimo ugotovitev iz zgornjega zglada ([3]).

Trditev 6.6. *Naj bosta G in H (G ni dvodelen) ter $f: G \rightarrow H$ homomorfizem med njima. Tedaj je liha ožina grafa G večja ali enaka lihi ožini grafa H .*

Dokaz. Naj bo $C = c_1c_2 \dots c_k$ najkrajši lihi cikel v G . Dokazati moramo, da v grafu H obstaja lih cikel dolžine $\leq k$. Ker graf G ni dvodelen, ga ne moremo slikati v povezavo izomorfno K_2 (to sledi iz posledice 6.19, ki jo bomo spoznali v nadaljevanju). Ker je C cikel, je tudi $f(C)$ cikel, čeprav je morda izrojen (lahko so kakšne povezave in vozlišča združena). Torej $f(C)$ v grafu H podaja vsaj kakšen cikel, in sicer morajo biti vsi ti cikli očitno dolžine $\leq k$. Razmisliti moramo le, da se ne more zgoditi, da bi pri tem dobili same cikle sode dolžine. Gotovo dobimo vsaj kak cikel, ker je C lihe dolžine in ga ne moremo slikati v nek acikličen graf (spet po posledici 6.19). Štejmo dolžino cikla tako, da upoštevamo tudi izrojene povezave, ki se vse pojavijo sodokrat. Če bi bil $f(C)$ sestavljen iz samih sodih ciklov (nekako združenih med sabo), bi sledilo, da je $f(c_1)f(c_2) \dots f(c_k)$ izrojeni cikel sode dolžine, kar pa ni možno, ker je k lih. \square

Opomba 6.7. Zanimivo je opaziti:

- Zahtevati moramo, da graf G ni dvodelen, sicer ga lahko homomorfno slikamo v povezavo znotraj grafa H .

- Podobna trditev ne velja za sodo ožino, ker lahko najkrajši sodi cikel v celoti preslikamo v povezavo znotraj grafa H , ne glede na to, kakšna je ožina tega grafa.

Iz trditve direktno sledi:

Posledica 6.8. *Naj bosta G in H (G ni dvodelen). Če je liha ožina grafa H strogo večja od lihe ožine grafa G , potem ne obstaja homomorfizem $G \rightarrow H$.*

Posledica pove, da res ne obstaja homomorfizem iz C_n v C_{n+2} , kadar je n liho število.

V nadaljevanju potrebujemo nekaj pojmov.

Če je $f: G \rightarrow H$ homomorfizem grafov, za $v \in V(H)$ definiramo $f^{-1}(v) = \{u \in V(G) ; f(u) = v\}$ in ga imenujemo *vlakno* homomorfizma f . Če sta grafa enostavna, tvori $f^{-1}(v)$ neodvisno množico v grafu G , tj. množico vozlišč, ki so paroma nepovezana. Vlakna homomorfizma f določajo particijo oz. razčlenitev π množice $V(G)$, ki jo imenujemo *jedro* homomorfizma f . Particijo sestavljajo neodvisne množice.

Zgled 6.9. Pomen pojmov si oglejmo na zgledih dvodelnega grafa in cikla.

- Vlakni homomorfizma f sta $f^{-1}(1) = A$ in $f^{-1}(2) = B$, ki sta res neodvisni množici. Jedro homomorfizma f je torej $\pi = \{A, B\}$, ki je res particija množice $V(G)$.

Vlakna homomorfizma g so oblike

$$g^{-1}(u) = \begin{cases} \{1\}, & \text{če } u \in A \\ \{2\}, & \text{če } u \in B, \end{cases}$$

kar sta tudi neodvisni množici. Jedro je particija $V(K_2) = \{1\} \cup \{2\}$.

- Vlakna homomorfizma h so množice z enim ali dvema elementoma in podajajo particijo $V(C_7) = \{A\} \cup \{B\} \cup \{C\} \cup \{D, F\} \cup \{E, G\}$, kot je z barvami nakazano na sliki 24.

6.2 RELACIJA \rightarrow

Definiramo relacijo \rightarrow na sledeč način: $G \rightarrow H$ (beri: "ima homomorfizem v") natanko tedaj, ko obstaja homomorfizem iz G v H .

Trditev 6.10. *Relacija \rightarrow je refleksivna in tranzitivna.*

Dokaz. Identična preslikava $G \rightarrow G$ je očitno homomorfizem, zato je relacija refleksivna.

Naj bosta $f: G \rightarrow H$ in $g: H \rightarrow Z$ homomorfizma med grafi. Preverimo, da je $g \circ f: G \rightarrow Z$ tudi homomorfizem. Naj bosta $u, v \in V(G), u \sim_G v$. Ker je f homomorfizem, je $f(u) \sim_H f(v)$. Ker je g homomorfizem je $g(f(u)) \sim_Z g(f(v))$, kar smo želeli dokazati. \square

Relacija \rightarrow ni nujno antisimetrična (glej primer zgoraj z dvodelnim grafom), torej tudi ni delna urejenost.

Definicija 6.11. Grafa G in H sta *homomorfno ekvivalentna*, če obstajata homomorfizem $G \rightarrow H$ in homomorfizem $H \rightarrow G$.

Očitno je homomorfna ekvivalenca ekvivalenčna relacija, zato podaja ekvivalenčne razrede grafov. Nanje lahko naravno razširimo relacijo \rightarrow . Opazimo, da je relacija \rightarrow delna urejenost na ekvivalenčnih razredih homomorfne ekvivalentnosti grafov.

Če sta grafa homomorfno ekvivalentna, še nista nujno izomorfna. Primer: dvodelni graf in K_2 (glej zgled zgoraj).

6.3 RETRAKTI

V nadaljevanju potrebujemo tudi homomorfizme, ki na nekem podgrafu mirujejo. Oglejmo si njihovo formalno definicijo ([7]).

Definicija 6.12. (*Prava*) *retrakcija* je homomorfizem $f: G \rightarrow H$, kjer je H (pravi) podgraf G in je $f|_H$ identična preslikava. Tedaj H imenujemo (*pravi*) *retrakt* grafa G .

Oglejmo si preprost zgled retrakta.

Zgled 6.13. Naj bo X graf kot na sliki 25. Definirajmo preslikavo $f: X \rightarrow C$, kjer je C zunanji 5-cikel grafa X , kot

$$\begin{aligned} f|_C &= \text{id}_C \\ f(F) &= B \\ f(G) &= C \\ f(H) &= D \\ f(I) &= E \\ f(J) &= A. \end{aligned}$$

Tako definirana preslikava je očitno homomorfizem grafov. Ker pa je f na C identiteta, je celo retrakcija. Torej je C retrakt grafa X . Podobno bi lahko graf X retrahirali tudi na notranji 5-cikel. Geometrijsko si lahko predstavljamo potek retrakcije tako, da notranji 5-cikel zavrtimo v pozitivni smeri, ga povečujemo in prečne povezave med obema cikloma prilepimo na zunanji cikel. Opozoriti velja, da morda bolj naravna predstava, ko bi notranji cikel le napihnil in torej slikali

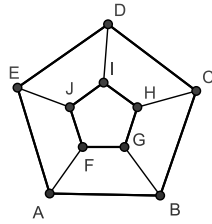
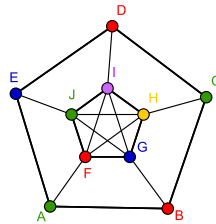
Tabela 25: Graf, ki ga lahko retrahiramo na cikel C_5 .

Tabela 26: Barvanje danega grafa podaja rekrakcijo.

$F \mapsto A, G \mapsto B$ itd. ni retrakcija (ker niti ni homomorfizem, saj sta na primer vozlišči A in F v začetnem grafu povezani, v končnem pa ne).

6.4 POVEZAVA Z BARVANJI GRAFOV

Spomnimo se ([9]), da je barvanje vozlišč grafa G preslikava b iz $V(G)$ v množico barv, pri čemer sosednji vozlišči ne smeta biti pobarvani z enako barvo. Če lahko G pobarvamo s k različnimi barvami, rečemo, da je G k -obarvljiv in b k -barvanje. Najmanjše število k , da lahko G pobarvamo s k različnimi barvami imenujemo kromatično število grafa G in označimo s $\chi(G)$. Barvanje grafa določa particijo vozlišč na neodvisne množice (v neki množici so vozlišča iste barve). Smiselno se zdi, da lahko to particijo povežemo z jedrom nekega homomorfizma. S tem v mislih si oglejmo še en zgled retrakcije.

Zgled 6.14. Naj bo H graf na sliki 26. Če vozlišča enake barve slikamo v enako pobarvana vozlišča v notranji 5-kliki, očitno dobimo homomorfizem, ki je poleg tega še retrakcija na kliko.

Zgled lahko posplošimo ([3]):

Trditev 6.15. Če ima graf kliko velikosti n , potem vsako n -barvanje grafa podaja retrakcijo na n -kliko.

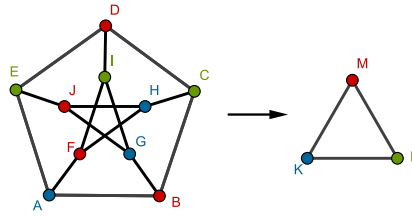


Tabela 27: Barvanje Petersenovega grafa homomorfizem.

Dokaz. Naj bo G graf s n -kliko $K_n \subseteq G$ in b n -barvanje grafa (barve označimo z $\{1, 2, \dots, n\}$). Ker je K_n klika, morajo biti vsa njena vozlišča pobarvana z različnimi barvami, torej lahko njena vozlišča brez škode označimo z $\{1, 2, \dots, n\}$ (ustrezno glede na barvo vozlišč). Definirajmo preslikavo $f: G \rightarrow K_n$ kot $f(u) = b(u)$. Najprej preverimo, da je preslikava f homomorfizem. Če je $uv \in E(G)$, sta vozlišči gotovo pobarvani z različnimi barvami. Njuni sliki $f(u), f(v)$ ležita v K_n kliku, torej sta gotovo povezani. Preslikava f je po definiciji identiteta na kliku K_n , zato je retrakcija. \square

Če graf ne vsebuje klike, potem homomorfizem, ki ga podaja barvanje, seveda ne bo retrakcija na poln graf. Razmislimo pa lahko, da vseeno dobimo homomorfizem, in poskušamo opaziti povezavo s kromatičnim številom grafa.

Zgled 6.16. Vemo, da lahko Petersenov graf pobarvamo s tremi barvami (kot na primer na sliki 27). Barvanje določa particijo vozlišč na neodvisne množice. Če definiramo homomorfizem, ki ima dano particijo za jedro, dobimo homomorfizem iz Petersenovega grafa v poln graf na 3 vozliščih, to je v $K_{\text{število barv}}$.

Ugotovitve iz zgleda posplošimo ([8]):

Trditev 6.17. *Kromatično število grafa G je enako najmanjšemu naravnemu številu r , da obstaja homomorfizem iz G v K_r .*

Dokaz. Naj bo $\chi(G) = r$. Dokažimo, da obstaja homomorfizem $G \rightarrow K_r$ in ne obstaja homomorfizem $G \rightarrow K_k$ za $k < r$.

Denimo, da obstaja homomorfizem $f: G \rightarrow K_k$. Preslikava ima k vlaken, ki tvorijo particijo $V(G)$ na neodvisne množice, torej določajo k -barvanje grafa G . Torej v tem primeru sledi $r = \chi(G) \leq k$, kar je protislovje.

Ker G lahko pobarvamo z r barvami, je preslikava $G \rightarrow K_r$, ki v isto vozlišče slika natanko vozlišča iste barve, homomorfizem. \square

Posledica 6.18. Naj bosta G in H grafa in $f: G \rightarrow H$ homomorfizem med njima. Tedaj je $\chi(G) \leq \chi(H)$.

Dokaz. Po prejšnji trditvi obstaja homomorfizem $g: H \rightarrow K_{\chi(H)}$. Tedaj je $g \circ f: G \rightarrow K_{\chi(H)}$ tudi homomorfizem in po prejšnji trditvi sledi, da je $\chi(G) \leq \chi(H)$. \square

Če zadnjo posledico negiramo, sledi:

Posledica 6.19. Če za dana grafa G in H velja $\chi(G) > \chi(H)$, potem ne obstaja noben homomorfizem iz G v H .

6.5 JEDRA

Nekateri grafi (ali njihovi podgrafi) se pri slikanju s homomorfizmom obnašajo na poseben način – se na nek način ohranjajo. Pravimo jim jedra in jih formalno definiramo na sledeč način.

Definicija 6.20. Graf G je *jedro* (anlg. core), če je vsak homomorfizem $G \rightarrow G$ bijekcija.

Podgraf $H \subseteq G$ je *jedro grafa* G , če je H jedro in obstaja homomorfizem $G \rightarrow H$. Tedaj označimo $H = G^\bullet$.

Zgoraj definiran pojem jedra se razlikuje od definicije v uvodu. Vendar je iz konteksta vedno razvidno, ali govorimo o jedru – particiji ali o jedru – (pod)grafu.

Zgled 6.21. Oglejmo si nekaj zgledov.

- Poln graf K_n je jedro. Poljuben homomorfizem samega vase je namreč le permutacija vozlišč, zato je bijekcija.
- Kolesa W_{2n+1} so jedra. Naj bo $f: W_{2n+1} \rightarrow W_{2n+1}$ poljuben homomorfizem. Nobeno od vozlišč s cikla se ne more slikati v središče in ker je središče povezano z vsemi vozlišči na ciklu, je edina možnost, da se središče slika samo vase. Preostali del preslikave f pa je potem le homomorfizem cikla C_{2n+1} . V nadaljevanju bomo videli, da so lihi cikli jedra, iz česar sledi, da je f bijekcija.
- Petersenov graf je jedro (dokaz preko lastnosti Kneserjevih grafov).
- Sodi cikli C_{2n} ($n \geq 2$) niso jedra. Ker so sodi cikli dvodelni grafi, obstaja homomorfizem iz cikla v povezavo $K_2 \subset C_{2n}$, ki očitno ni injektiven, torej tudi ni bijektiven (glej sliko 28).
- Poljuben dvodelen graf z vsaj tremi vozlišči in vsaj eno povezavo ni jedro.

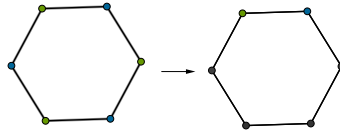


Tabela 28: Če slikamo tako, kot nakazujejo barve vozlišč na sliki, dobimo homomorfizem $C_{2n} \rightarrow C_{2n}$, ki ni bijektiven.

- Polni graf brez ene povezave ni jedro. Vozlišči, ki nista povezani, lahko pobarvamo z isto barvo, nato pa homomorfizem definiramo tako, da je na ostalih vozliščih identiteta, na nepovezanih vozliščih pa slikamo v izbranega predstavnika barvnega razreda. Preslikava je homomorfizem, ki ni injektiven (glej sliko 29).

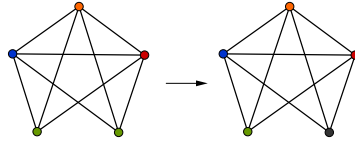


Tabela 29: Če slikamo tako, kot nakazujejo barve vozlišč na sliki, dobimo homomorfizem, ki ni bijektiven.

Opazimo, da direktno iz definicije sledi, da če je G^\bullet jedro grafa G , obstajata homomorfizem $G^\bullet \rightarrow G$ (vložitev) in homomorfizem $G \rightarrow G^\bullet$ (homomorfizem iz definicije). To lastnost bomo potrebovali v nadaljevanju.

Navedimo ekvivalentni definiciji jedra, ki nam bosta pomagali bolje razumeti jedra ([2] in [4]).

Trditev 6.22. *Naslednje trditve so ekvivalentne:*

1. Graf G je jedro.
2. Ne obstaja homomorfizem iz G v njegov pravi podgraf.
3. Graf G nima nobenega pravega retrakta.

Dokaz. Implikaciji (1) \implies (3) in (3) \implies (2) sta očitni. Dokazati moramo le še (2) \implies (1). Naj bo $f: G \rightarrow G$ poljuben homomorfizem. Po predpostavki (2) mora biti surjektiven. Ker je graf G končen, sledi tudi bijektivnost. \square

Oglejmo si še dodatno lastnost jedra danega grafa.

Trditev 6.23. *Naj bo H jedro grafa G . Tedaj je H retrakt G .*

Ta lastnost jedra seveda ne karakterizira, ker ni nobenega razloga, da bi bil poljuben retrakt tudi jedro grafa. Oglejmo si sedaj dokaz trditve.

Dokaz. Ker je H jedro G , obstaja homomorfizem $f: G \rightarrow H$, da je $f|_H$ bijekcija. Torej je $f|_H: H \rightarrow H$ avtomorfizem. Zato obstaja njegov inverz g , ki je tudi avtomorfizem. Tedaj je po trditvi 6.10 preslikava $f \circ g: G \rightarrow H$ homomorfizem grafov. Poleg tega je $(f \circ g)|_H(v) = f(g(v)) = v$, torej je $(f \circ g)|_H = \text{id}_H$, iz česar sledi, da je H retrakt od G . \square

V nadaljevanju bi radi videli, da vsak graf ima jedro, ki je v nekem smislu enolično. Potrebujemo naslednjo lemo.

Lema 6.24. *Naj bosta G in H jedri. Grafa G in H sta homomorfno ekvivalentna natanko tedaj, ko sta izomorfna.*

Dokaz. (\Rightarrow) Naj bosta $f: G \rightarrow H$ in $g: H \rightarrow G$ homomorfizma med grafoma. Potem je $f \circ g: H \rightarrow H$ tudi homomorfizem, ki je bijektiven, saj je H jedro. Podobno sklepamo, da je $g \circ f$ bijekcija. Sledi, da sta si preslikavi f in g inverzni in obe bijektivni, torej sta grafa G in H izomorfna.

(\Leftarrow) Očitno, saj je izomorfizem hkrati tudi homomorfizem in ima inverz, ki je tudi homomorfizem. \square

Iz leme sledi, da je relacija \rightarrow delna urejenost na izomorfnih razredih jeder. Pred dokazom izreka o obstoju in enoličnosti jedra, si oglejmo še povezavo med jedri grafov in homomorfno ekvivalenco.

Trditev 6.25. *Grafa G in H sta homomorfno ekvivalentna natanko tedaj, ko sta njuni jedri izomorfni.*

Dokaz. Dokazati moramo: grafa G in H sta homomorfno ekvivalentna natanko tedaj, ko sta njuni jedri homomorfno ekvivalentni. Potem trditev velja po lastnosti 6.24.

(\Rightarrow) Če obstaja homomorfizem $G \rightarrow H$, potem imamo sledeče zaporedje homomorfizmov:

$$G^\bullet \rightarrow G \rightarrow H \rightarrow H^\bullet,$$

ki tvori homomorfizem iz G^\bullet v H^\bullet . Podobno iz obstoja homomorfizma $H \rightarrow G$ izpeljemo obstoj homomorfizma $H^\bullet \rightarrow G^\bullet$.

(\Leftarrow) Če obstaja homomorfizem $G^\bullet \rightarrow H^\bullet$, potem imamo sledeče zaporedje homomorfizmov:

$$G \rightarrow G^\bullet \rightarrow H^\bullet \rightarrow H,$$

ki tvori homomorfizem iz G v H . Spet lahko podobno sklepamo tudi za obratno obrnjene homomorfizme. \square

Dokažimo še naslednji izrek, ki zagotavlja obstoj jedra v vsakem grafu.

Izrek 6.26. Vsak graf G ima jedro, ki je induciran podgraf grafa G in je enoličen do izomorfizma natančno.

Dokaz. Naj bo \mathcal{P} množica vseh takih podgrafov G , da obstaja homomorfizem iz G vanj. Množica je neprazna (ker je identiteta homomorfizem iz G v G) in končna (ker je graf G končen, ima končno podgrafov). Torej ima vsaj en najmanjši element $H \subseteq G$ glede na inkluzijo grafov (to je kandidat za jedro grafa G). Zaradi minimalnosti ne obstaja noben homomorfizem iz H v njegov pravi podgraf. Po trditvi 6.22 je H jedro.

Dokazati moramo le še enoličnost. Denimo, da sta H_1, H_2 jedri grafa G . Označimo pripadajoča homomorfizma z $f_i: G \rightarrow H_i, i = 1, 2$. Tedaj je $f_1|_{H_2}$ homomorfizem iz H_2 v H_1 in $f_2|_{H_1}$ homomorfizem iz H_1 v H_2 . Torej sta H_1 in H_2 homomorfno ekvivalentna in zato po lemi 6.24 izomorfna. \square

Čeprav zdaj vemo, da jedro grafa obstaja in je do neke mere enolično, nam to ne pomaga pri ugotavljanju, ali je graf sam sebi jedro. Izkaže se celo, da je to NP-poln problem [5].

Zgled 6.27. Za ilustracijo poiščimo jedra grafov iz zgleda 6.21, ki niso sami sebi jedra.

- Če je G dvodelni graf (z vsaj 3 vozlišči in vsaj eno povezavo) ali cikel sode dolžine, obstaja homomorfizem $G \rightarrow K_2 \subset G$ in K_2 je jedro ter hkrati induciran podgraf. Torej je poljubna povezava znotraj grafa G njegovo jedro in seveda so vse povezave med samo izomorfne.
- Jedro polnega grafa K_n brez ene povezave je induciran podgraf, ki je izomorfen K_{n-1} . V zgledu 6.21 smo opisali homomorfizem $K_n \rightarrow K_{n-1}$, poleg tega pa je K_{n-1} jedro, saj je poln graf.

Že v uvodnem zgledu smo srečali nekaj družin jeder – polni grafi, lihi cikli in liha kolesa. Vendar smo le za polne grafe znali elementarno pokazati, da so jedra. Upamo, da lahko najdemo še kakšno večjo družino jeder. S tem v mislih si oglejmo naslednjo definicijo.

Definicija 6.28. Če je G takšen graf, da ima vsak njegov pravi podgraf kromatično število $< \chi(G) = k$, ga imenujemo *kromatično kritični ali k-kritični*.

Zgled 6.29. Oglejmo si nekaj zgledov.

- Vsak pravi podgraf polnega grafa K_n ($n \geq 1$) vsebuje vsaj eno povezavo manj, zato lahko vozlišči, ki zdaj nista več sosednji pobarvamo z isto barvo in porabimo največ $n - 1$ barv. Torej ima pravi podgraf kromatično število strogo manjše od n . Torej so polni grafi n -kritični.
- Vsak pravi podgraf lihega cikla vsebuje vsaj eno povezavo manj od začetnega grafa. Torej je to podgraf poti P_{2n+1} , ki ima kromatično število $2 < 3 = \chi(C_{2n+1})$. Torej so lihi cikli 3-kritični.

Trditev 6.30. *Vsak kromatično kritični graf je jedro.*

Dokaz. Naj bo G kromatično kritični. Po posledici 6.19 sledi, da ne obstaja noben homomorfizem iz G v njegov pravi podgraf. Torej je jedro G lahko le graf sam. Ker po izreku 6.26 jedro obstaja, je dokaz končan. \square

Iz trditve sledi, da so polni grafi in lihi cikli jedra. Hkrati pa nam podaja še širši zgled jeder – vse kromatično kritične grafe. Takšne vrste rezultat je koristen tudi pri iskanju jeder v grafu, ki ni sam sebi jedro. Če namreč najdemo homomorfizem iz grafa v nek podgrafi, ki je kromatično kritični, je to po definiciji iskano jedro.

6.6 ZAKLJUČEK

Vidimo, da so homomorfizmi precej močno orodje pri delu z grafi. Zanimivo je, da lahko iz neobstoja homomorfizma med dvema grafoma sklepamo nekaj o njunih lastnostih (npr. o lihi ožini in kromatičnemu številu). Presenetljivi so tudi rezultati povezani z jedri, predvsem to, da vsak graf ima enolično jedro in da lahko iz primerjave jeder dveh grafov sklepamo na morebitno homomorfno ekvivalenco grafov.

V seminarju so predstavljene le osnove dela s homomorfizmi. Poleg tega so zelo razvite teorije, ki iščejo povezave med homomorfizmi in kromatičnimi števili produktov grafov ter enolično obarvljivimi grafi. Veliko je znanega tudi o jedrih različnih družin grafov.

LITERATURA

- [1] Godsil, Chris, and Gordon Royle. "Algebraic graph theory, volume 207 of Graduate Texts in Mathematics." (2001): 103–120.
- [2] Hell, Pavol, and Jaroslav Nešetřil. "The core of a graph." *Discrete Mathematics* 109.1 (1992): 117-126.
- [3] Hahn, Geňa, and Claude Tardif. "Graph homomorphisms: structure and symmetry." *Graph symmetry*. Springer Netherlands (1997): 107-166.
- [4] Foniok, Jan, and Jaroslav Nešetřil. "Graph Theory and Combinatorics" (2011).
- [5] Hell, Pavol and Jaroslav Nešetřil. "On the complexity of H - coloring." *J. Combin. Theory Ser. B*, 48 (1990): 92-110.
- [6] Sabidussi, Gert. "Graph derivatives." *Math. Z.* 76 (1961): 385 – 401.
- [7] Hell, Pavol. "Rétractions de graphes." Ph. D. Thesis, Université de Montréal, 1972.
- [8] Catlin, Paul. "Homomorphisms as generalizations of graph coloring." *Congr. Numer.* 50 (1985): 179 – 186.
- [9] Škrekovski, Riste. "Uporabna diskretna matematika." Lecture notes, Fakulteta za matematiko in fiziko, Univerza v Ljubljani (2016).

STISKANJE GRAFOV

Miha Eleršič

7.1 UVOD

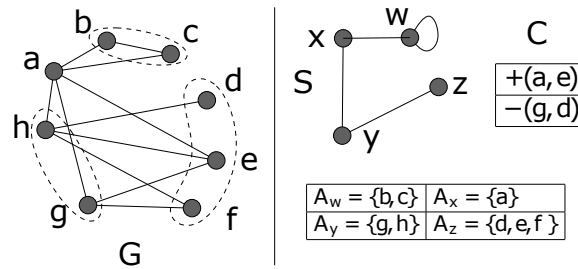
Grafi se uporabljajo za modeliranje pojavov in struktur iz realnega sveta. Nekateri takšni primeri pa so lahko zelo veliki. Primeri velikih grafov so naprimer grafi povezav med spletnimi stranmi [5], socialna omrežja, usmerjanje internetega prometa, podatki o nakupih v ku-povalnih centrih itd.

Veliko število vozlišč in povezav nas ovira pri odkrivanju novih znanj iz velikih grafov, saj obstoječi algoritmi niso zmožni opraviti svojega dela v praktičnem času. Kot rešitev si zamislimo stisnjen graf, ki z manjšim številom vozlišč in povezav predstavi glavne lastnosti originalnega grafa. Na stisnjenih grafih lahko nato uporabimo obstoječe algoritme za analizo grafov in pričakujemo podobne rezultate, kot če bi analizirali izvirni graf.

Običajno se za stiskanje grafov uporabljajo različne metode gru-čenja, saj imajo vozlišča iz iste gruče podobne lastnosti in jih lahko predstavimo s supervozlišči. Drug pristop je, da graf predstavimo z najpomembnejšimi vozlišči in povezavami, kjer nam pomembnost vo-zlišča dajejo različne mere, naprimer centralnosti [6, 7]. Omenimo še iskane pogostih podgrafov, kjer lahko pogoste podgrafe nadomestimo s kazalci na en zapis takega podgrafa [8]. Prvi dve metodi imata po-mankljivost, da izgubljata informacije (lossy compression), tretja pa je računsko zelo zahtevna in lahko za ogromne grafe postane neupo-rabna.

V naslednjih poglavjih je opsian pristop k stiskanju grafov povzet po [2]. Gre za pristop, ki uporabi teorijo informacije za izgradnjo sti-snjene predstavitev grafa, sestavljeno iz supervozlišč in povezav med

Tabela 30: Primer stisnjene predstavitev grafa



njimi. Da dosežemo stiskanje brez izgub pa dodamo še množico popravkov, ki nam pomaga pri izgradnji originalnega grafa.

7.2 STISNJENA PREDSTAVITEV GRAFA

Naj bo $G = (V_G, E_G)$ graf. Njegova stisnjena predstavitev je $R = (S, C)$, ki je sestavljena iz povzetka grafa $S = (\mathcal{V}_S, E_S)$ in množice popravkov C . Množica \mathcal{V}_S je razbitje množice V_G . Vsako supervozlišče povzetka grafa torej vsebuje enega ali več vozlišč originalnega grafa, vsako vozlišče originalnega grafa pa se pojavi v natanko enem supervozlišču. Množica E_S predstavlja povezave med super vozlišči. Povezava $(V_i, V_j) \in E_S$ pomeni povezave med vsakim vozliščem iz V_i z vsakim vozliščem iz V_j v grafu G . Množica popravkov C vsebuje dva tipa elementov: $+e$ pomeni dodajanje povezave iz grafa G , ki je povzetek ni zajel; $-e$ pa pomeni brisanje povezave, ki jo povzetek rekonstruira, a se ne nahaja v originalnem grafu.

Opazka 7.1. Predstavitev R enolično določi graf G .

Na sliki 30 je na primeru prikazan graf in njegova stisnjena reprezentacija. Število povezav v stisnjeni implementaciji se zmanjša iz 11 na 6, kjer so 4 povezave del povzetka grafa, 2 pa so vsebovane v množici popravkov.

7.3 DOLŽINA MINIMALNE PREDSTAVITVE

V teoriji informacije poznamo pojem minimalne predstavitev (angl. Minimal description length – MDL). Gre za formalizacijo Occamove britve. Če imamo več različnih možnih predstavitev istega podatka, potem najboljšo predstavitev definiramo kot tisto, ki doseže največje stiskanje oziroma je najkrajša [1].

Da lahko uporabimo to definicijo, moramo najprej definirati dolžino predstavitev grafa.

Definicija 7.2. Cena predstavitve grafa je

$$\text{cena}(R) = |E_s| + |C|.$$

Ceno zapisa vozlišč zanemarimo, saj je za velike grafe v primerjavi s ceno zapisa povezav zanemarljiva. Sedaj lahko definicijo MDL uporabimo na predstavitev stisnjenih grafov.

Definicija 7.3. Za graf G je njegova *minimalna predstavitev grafa* R tista, za katero velja

$$\nexists R' : \text{cena}(R') < \text{cena}(R).$$

Poglejmo si primer iskanja minimalne predstavitve grafa, če fiksiramo razbitje \mathcal{V}_S . Za vsak par vozlišč V_i in V_j s π_{ij} označimo množico vseh možnih povezav med vozlišči vsebovanimi v V_i in V_j . Velja $|\pi_{ij}| = |V_i| * |V_j|$. Z A_{ij} označimo povezave med vozlišči vsebovanimi v V_i in V_j , ki se pojavijo v G . Opazimo $A_{ij} = \pi_{ij} \cap E_G$. V stisnjem grafu imamo dve možnosti za predstavitev A_{ij} :

- V C dodamo vse povezave iz A_{ij} . To nas stane $|A_{ij}|$ povezav.
- V E_s dodamo povezavo (V_i, V_j) , v C pa dodamo vse odvečne povezave. To nas stane $(1 + |\pi_{ij}| - |A_{ij}|)$ povezav.

Izbira za pomasemezen par i in j ne vpliva na ceno drugih parov. Za vsak par lahko torej izberemo tisto možnost, ki ima nižjo ceno in s tem pridemo do končne predstavitve, ki ima najnižjo ceno za izbrano razbitje.

Opazka 7.4. Za fiksno razbitje \mathcal{V}_S je dolžina minimalne predstavitve enolično določena.

Ker pa število možnih razbitij eksponentno narašča, ne bomo iskali razbitja z njamanjšo ceno, vendar bomo pokazali, da požrešna metoda iskanja dobrega razbitja vrne zelo dobre rezultate.

7.4 APROKSIMACIJA PREDSTAVITVE GRAFA

Do sedaj smo govorili o predstavitev grafa, ki enolično rekonstruirajo originalen graf – stiskanje brez izgub. Za nekatere uporabe stiskanja grafov, kot je na primer vizualizacija, pa ta lastnost ni nujno potrebna. Definiramo predstavitev, ki omogoča stiskanje grafov z izgubami.

Naj bo N_v množica povezav iz v v grafu G in N'_v množica povezav iz v v grafu rekonstruiranem iz R_ϵ . Množica $|N_v - N'_v|$ predstavlja povezave v G , ki jih ni v predstavitvi R_ϵ . Množica $|N'_v - N_v|$ predstavlja odvečne povezave v predstavitvi R_ϵ , ki ji ni v originalnem grafu.

Definicija 7.5. ϵ -aproksimacija predstavitve grafa G je stisnjena predstavitev $R_\epsilon = (S_\epsilon, C_\epsilon)$, za katero velja:

$$\forall v \in G : |N_v - N'_v| + |N'_v - N_v| \leq \epsilon |N_v|.$$

Ker za vsako vozlišče velja, da bo pravilno predstavljen vsaj $(1 - \epsilon)$ delež povezav, potem velja tudi za celoten graf: predstavitev R_ϵ pravilno predstavi vsaj $(1 - \epsilon)$ delež vseh povezav originalnega grafa.

Konstruiranje aproksimacije predstavitve grafa je enostavno, če že imamo konstruirano navadno predstavitev grafa. Za vsak element množice C_ϵ pogledamo, če bi njegova odstranitev kršila pogoj iz definicije 7.5 in ga pogojno na to odstranimo.

Problem se pojavi, če smo odstranili že vse elemente C_ϵ in hkrati še nismo dosegli mejnega pogoja. Nadaljevati bi morali z odstranjevanjem povezav iz S_ϵ . Sedaj je potrebno preverjati pogoj iz definicije 7.5 za vsa vozlišča originalnega grafa vsebovana v obeh vozliščih iz S_ϵ , med katerima je povezava, ki jo želimo odstraniti. Zaradi tega je bolje aproksimacijo predstavitve grafa konstruirati direktno [2], vendar tega v tem članku ne bomo predstavili.

7.5 ALGORITEM

V tem razdelku predstavimo algoritem GREEDY, ki požrešno zgradi stisnjeno predstavitev vhodnega grafa.

7.5.1 Ideja

Imenujmo eno vozlišče predstavitve grafa supervozlišče. Supervozlišče je torej ena izmed množic razbitja originalnega grafa. Algoritem začne z razbitjem, ki v vsaki množici vsebuje le eno vozlišče originalnega grafa. Algoritem nato iterativno združuje po dve supervozlišči v eno večje supervozlišče, dokler tako združevanje zmanjšuje ceno predstavitve grafa.

7.5.2 Izračun najboljšega para za združevanje

Iz poglavja 7.3 vemo, da je cena predstavitve povezav med dvema supervozliščema V_i in V_j enaka $\min(|A_{ij}|, 1 + |\pi_{ij}| - |A_{ij}|)$. Označimo to ceno z c_{ij} . V_i in V_j sta sosednji, če $\exists(a, b) \in E_G : a \in V_i \wedge b \in V_j$. Skupna cena predstavitve enega supervozlišča V_i je vsota cen povezav z vsemi njegovimi sosedi. Označimo to ceno z c_i . Z $s(u, v)$ označimo razmerje v ceni predstavitve dela grafa če združimo u in v v novo supervozlišče w .

Definicija 7.6. Naj bo

$$s(u, v) = \frac{c_u + c_v - c_w}{c_u + c_v}.$$

Razlog za izbiro relativnega zmanjšanja cene namesto absolutnega je v tem, da absolutno zmanjšanje postavi previsoko oceno vozliščem z veliko sosedi, četudi nimajo veliko skupnih sosedov. Z relativno oceno se temu izognemo.

7.5.3 Realizacija

Algoritem razdelimo na tri faze: inicializacija, iterativno združevanje in izpis.

V fazi inicializacije vsem parom vozlišč na razdalji dve določimo vrednost funkcije $s(u, v)$. Gledamo samo vozlišča na razdalji dve, saj bolj oddaljena vozlišča nimajo skupnih sosedov in imajo zato zagotovo negativno vrednost funkcije s , kar pomeni večjo ceno predstavitve po združevanju. Za učinkovitejše delovanje vstavimo rezultate v kopico, ki ima na vrhu vedno največji element.

V fazi iterativnega združevanja vzamemo iz kopice največji element $s(u, v)$ in združimo supervozlišči u in v v novo vozlišče w . Nato iz kopice odstranimo vse vrednosti, ki vsebujejo u ali v . Nazadnje v kopico dodamo vse pare, ki vsebujejo novo supervozlišče w . S tem še vedno nismo pokrili vseh možnih sprememb cen, ki so že v kopici. Pregledati moramo še soseščino supervozlišč u in v , kar je ekvivalentno soseščini w , in po potrebi spremeniti njihovo vrednost in mesto v kopici. To ponavljamo dokler je največji element v kopici večji od 0.

V fazi izpisa je potrebno zgraditi množico popravkov, saj smo do sedaj upravljali le z razbitji. Postopek izgradnje sledi iz ospisa v poglavju 7.3. V zgledu 1 je psevdokoda zgoraj opisanega algoritma. Na sliki 7.5.3 je prikazanih nekaj vmesnih korakov algoritma GREEDY na grafu iz slike 30.

7.5.4 Algoritem RANDOMIZED

Algoritem GREEDY lahko pospešimo. Namesto računanja vrednosti funkcije $s(u, v)$ za vse pare, si izberemo eno vozlišče naključno, naj bo to vozlišče v . Nato pregledamo vrednost funkcije za vsak u na razdalji dve od v . Nato združimo v z najboljšim kandidatom. Ponavljamo dokler niso več mogoča ugodna združenja. Ta pristop nam seveda prinese slabše rezultate kot algoritem GREEDY, vendar prihrani na računski zahtevnosti. V podrobnosti algoritma se ne bomo spuščali.

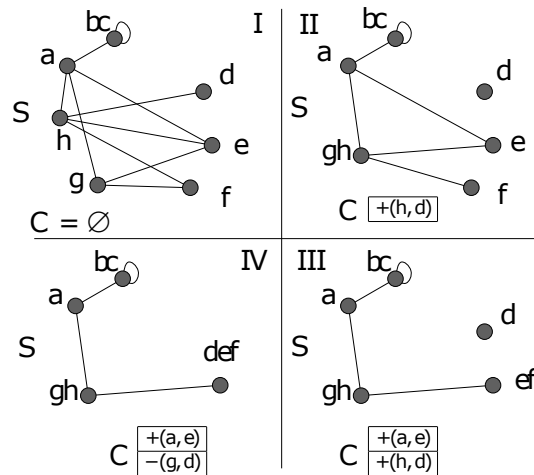
Algorithm 1 GREEDY

```

 $V_S = V_G; H = \emptyset$ 
for all  $(u, v)$  element  $V_S$  that are 2 hops apart do
  if  $(s(u, v) > 0)$  then insert  $(u, v, s(u, v))$  into  $H$ 
  end if
end for
while  $H$  is not empty do
  choose pair  $(u, v) \in H$  with the largest  $s(u, v)$  value
   $w = u \cup v$ 
   $V_S = (V_S - \{u, v\}) \cup w$ 
  for all  $x \in V_S$  that are within 2 hops of  $u$  or  $v$  do
    delete  $(u, x)$  and  $(v, x)$  from  $H$ 
    if  $(s(w, x) > 0)$  then
      insert  $(w, x, s(w, x))$  into  $H$ 
    end if
  end for
  for all pairs  $(x, y)$ , such that  $x$  or  $y$  is in  $N_w$  do
    delete  $(x, y)$  from  $H$ 
    if  $(s(x, y) > 0)$  then
      insert  $(x, y, s(x, y))$  into  $H$ 
    end if
  end for
end while
 $E_S = C = \emptyset$ 
for all pairs  $(u, v)$  such that  $u, v \in V_S$  do
  if  $(|A_{uv}| > (|\pi_{uv}| + 1)/2)$  then
    add  $(u, v)$  to  $E_S$ 
    for all  $(a, b) \in (\pi_{uv} - A_{uv})$  do
      add  $-(a, b)$  to  $C$ 
    end for
  else
    for all  $(a, b) \in A_{uv}$  do
      add  $+(a, b)$  to  $C$ 
    end for
  end if
end for
return  $((V_S, E_S), C)$ 

```

Tabela 31: Ilustracija postopka algoritma GREEDY



7.6 REZULTATI

V članku [2] so predstavljeni rezultati primerjave algoritmov GREEDY in RANDOMIZED in primerjava predstavljenih algoritmov z trenutno uveljavljenimi algoritmi za stiskanje grafov REF in GRAC.

7.6.1 Učinkovitost stiskanja

Na sliki 32 je prikazno testiranje stiskanja brez izgub algoritmov GREEDY in RANDOMIZED na naključnih grafih s ceno predstavitvije od 30.000 do 400.000. Vidimo, da oba algoritma prinašata dobro stiskanje, ki se z večanjem grafov izboljšuje. Kot pričakovano, GREEDY vedno doseže boljše stiskanje od RANDOMIZED.

7.6.2 Hitrost stiskanja

Na sliki 33 je prikazana primerjava hitrosti algoritmov GREEDY in RANDOMIZED na naključnih grafih z 30.000 do 400.000 vozlišči. Kot pričakovano je RANDOMIZED hitrejši.

7.6.3 Primerjava z drugimi algoritmi stiskanja

Na sliki 34 je prikazana primerjava med algoritmi GREEDY, RANDOMIZED, REF [3] in GRAC [4].

REF je popularen algoritem za stiskanje grafov, ki predstavljajo povezave med spletnimi stranmi. V primerjavi je uprabljena spremenjena različica, ki opravi le transformacijo grafa, brez kodiranja bitne

Tabela 32: Učinkovitost stiskanja

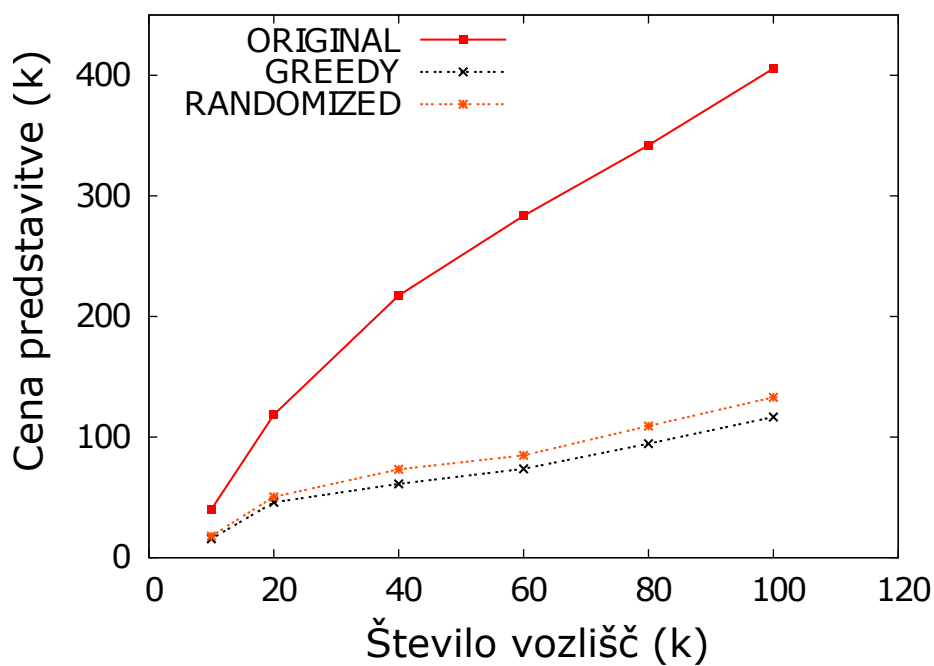


Tabela 33: Hitrost stiskanja

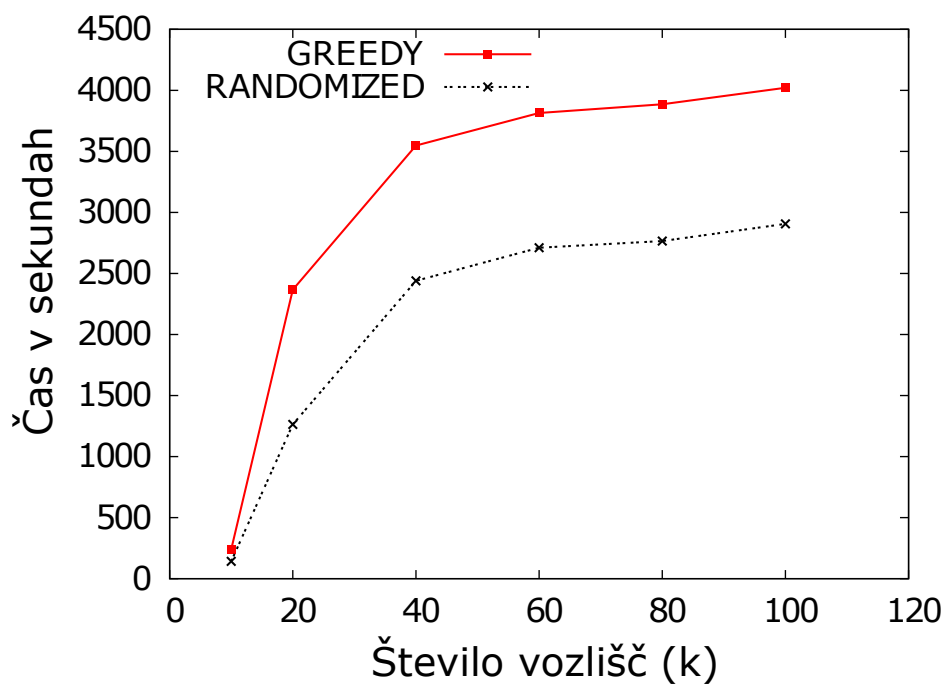
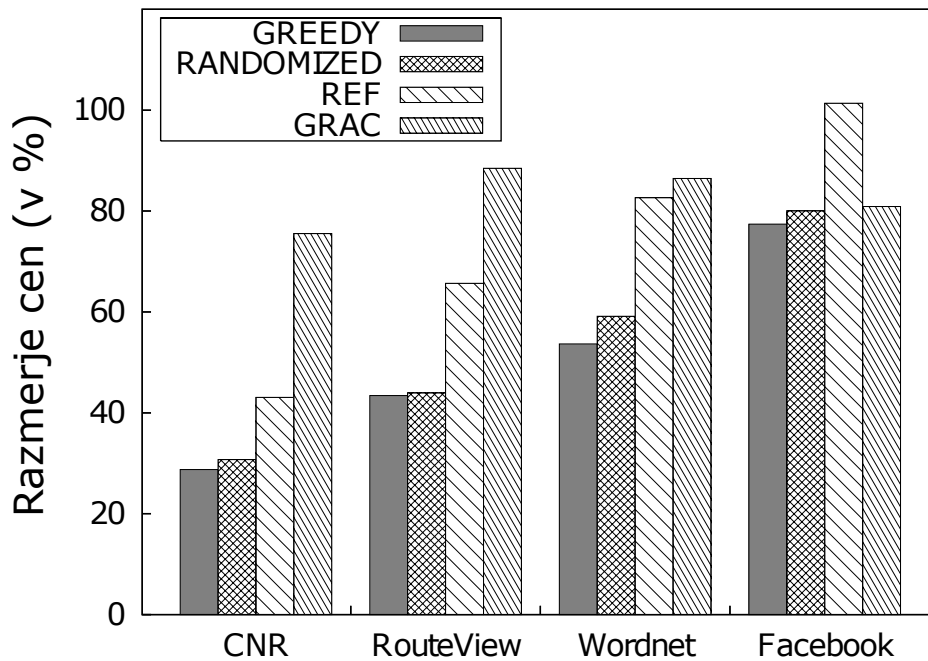


Tabela 34: Primerjava z drugimi algoritmi stiskanja



predstavitve grafa. Modifikacija je potrebna, saj naša ocena šteje število povezav in ne števila bitov v končni reprezentaciji.

GRAC je algoritem za gručenje grafov. Uteži za gručenje so bile nastavljene tako, da so minimizirale končno ceno predstavitve.

Testni podatki

- **CNR:** Podatki o povezavah med stranmi domene CNR s 100.000 vozlišči in 405.586 povezavami [9].
- **RouteView:** Predstavlja topologijo usmerjevalnikov svetovnega spleta. Med dvema vozliščema je povezava, če obstaja fizična povezava med usmerjevalnikoma. Graf ima 10.000 vozlišč in 21.000 povezav [10].
- **WordNet:** Iz baze besed *WordNet* je zgrajen graf, kjer povezava ustreza povezavi če imata besedi neko skupno točko. To so na primer sopomenke, nadpomenke in podobno. Nastali graf ima 76.853 vozlišč in 121.307 povezav [11].
- **Facebook:** Predstavlja graf, zgrajen leta 2005 iz socialnega omrežja Facebook študentov univerze Cornell. Vozlišča so študenti, povezava pa obstaja če sta dva študenta na omrežju označena kot prijatelja. Vsebuje 14.562 vozlišč in 601.735 povezav [2].

Opazimo, da se primerjana algoritma REF in GRAC bolje obneseta na svojih specializiranih področjih. GRAC je relativno boljši na socialnem omrežju, REF pa se bolje obnese pri stiskanju povezav med spletnimi stranmi. Vidimo, da tako RANDOMIZED kot GREEDY premagata obstoječa algoritma za stiskanje grafov na vseh testnih primerih.

7.7 ZAKLJUČEK

V članku je opisana stisnjena predstavitev grafa in algoritem za njen izračun, povzeto po [2]. Stisnjena predstavitev grafa je manjša, a ima podobne lastnosti kot originalen graf, kar nam omogoča uporabo obstoječih algoritmov za analizo in prikaz velikih grafov. Stisnjena predstavitev je sestavljena iz povzetka grafa in seznama popravkov. Skupaj enolično določita začetni graf. Predstavljen je bil način kako stiskanje brez izgub posplošiti na stiskanje z izgubami. Za izračun predstavitve sta predstavljena algoritma GREEDY in njegova hitrejša različica RANDOMIZED. Na testnih podatkih iz realnega sveta sta oba algoritma pokazala boljše rezultate od obstoječih algoritmov za stiskanje. Presentljivo je, da tako enostavna konstrukcija premaga svoje prednike. To nakazuje, da je področje stiskanja grafov še mlado, in da bodo prihodnji algoritmi še veliko boljši.

LITERATURA

- [1] J. Rissanen *Modeling by shortest data description*, Automatica 14 (5), 465–658, 1978.
- [2] S. Navlakha, R. Rastogi, N. Shrivastava, *Graph summarization with bounded error*, SIGMOD '08: Proceedings of the 2008 ACM SIGMOD international conference on Management of data, New York, 2008.
- [3] P. Boldi, S. Vigna, *The webgraph framework: Compression techniques*, WWW, 595–602, 2004.
- [4] I. Dhillon, Y. Guan, B. Kulis. *A fast kernel-based multilevel algorithm for graph clustering*, KDD, 629–634, 2005.
- [5] S. Brin, L. Page, *The anatomy of a large-scale hypertextual web search engine*, WWW, 107–117, 1998.
- [6] S. Gert, *The centrality index of a graph*, Psychometrika 31(4),581–603, December 1966.
- [7] M.E.J. Newman, M. Girvan, *Finding and evaluating community structure in networks*, Physical Review E 69 ,2004.
- [8] M. Fiedler, C. Borgelt: *Subgraph support in a single large graph*, ICDM '07: Proceedings of the Seventh IEEE International Conference on Data Mining Workshops, Washington, DC, USA, IEEE Computer Society, 399–404, 2007.
- [9] P. Boldi and S. Vigna, *The WebGraph Framework I: Compression Techniques*, Proc. of the Thirteenth International World Wide Web Conference (WWW 2004), ACM Press, Manhattan, USA, 2004, <http://law.di.unimi.it/datasets.php>
- [10] *RouteView*, University of Oregon Route Views Project, 2001, <http://www.routeviews.org/>
- [11] V. Batagelj, A. Mrvar *Pajek datasets*, 2006 <http://vlado.fmf.uni-lj.si/pub/networks/data/>

Rebeka Merhar

Po Vizingovem izreku imajo lahko kubični grafi kromatičen indeks 3 ali 4. Zanimiv problem je iskanje kubičnih grafov, ki imajo kromatičen indeks 4. Ideja, da bi iskali take grafe je prišla iz problema štirih barv. Tait se je lotil dokaovanja problema štirih barv tako, da je naredil dualen graf triangulacije. Graf, ki ga je dobil, je bil povezna, kubičen in brez mostov. Tait je problem štirih barv prevedel na tri barvanje povezav kubičnih grafov brez mostov.

Do leta 1975 so bile poznane samo štiri take skupine kubičnih grafov, ki imajo kromatičen indeks 4 in sicer

- Petersenov graf, ki je prikazan na sliki 36,
- dva Blanuševa grafa,
- Szekerejev graf, ki je prikazan na sliki 35,
- Descartesov graf, ki je prikazan na sliki 40.

Ker so bili taki grafi zelo zmuzljivi, jih je Martin Gradner v delu [1] poimenoval *snarki*. Izraz si je sposodil iz pesmi *Lov na snarka*, ki jo je napisal Lewis Carroll. Slika 35 je vzeta iz vira [5].

V prvem razdelku definiramo snarke in povemo nekaj njihovih lastnosti. V drugem razdelku se posvetimo Issacovim snarkom, definiramo točkasti in zvezdasti produkt ter predstavimo dve konstrukciji snarkov. V podrazdelku predstavimo še cvetlične snarke ter dvojni zvezdast snark. V zadnjem razdelku na kratko opišemo še k -snarke.

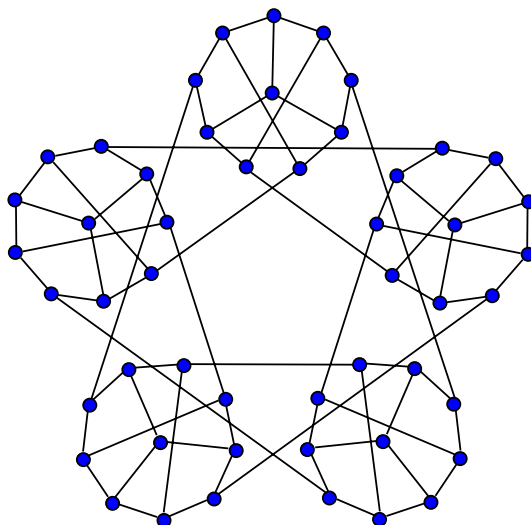


Tabela 35: Slika prikazuje Szekerjev graf.

8.1 OSNOVNE DEFINICIJE

Celotna seminarska naloga je povzeta po viru [2].

Definicija 8.1. Graf G je *ciklično po povezavah povezan*, če je potrebno odstraniti vsaj k povezav, da graf razpade na dve komponenti pri čemer obe komponenti vsebujeta cikel.

Definicija 8.2. *Snark* je kubičen graf, ki je ciklično po povezavah 4 povezan, ima ožino vsaj 5 in ima kromatičen indeks 4.

Definicija 8.3. Naj bo G graf in naj bo A neka množica povezav. Potem je A *ciklična prerezna množica*, če ima graf $G - A$ dve komponenti in vsaka od teh komponent vsebuje cikel.

Zgornje definicije so vzete po delu [2]. Slika 36 je vzeta iz vira [6].

Trditev 8.4. *Petersenov graf je najmanjši snark in edini snark na 10 vozliščih.*

Dokaz. Predpostavimo, da je G snark. Naj bo A ciklično prerezna množica po povezavah grafa G . Torej ima graf $G - A$ dve komponenti in vsaka od niju vsebuje cikel. Ker je ožina grafa G 5, morata komponenti grafa $G - A$ vsebovati vsaj 5 vozlišč. Torej sledi, da je $|V(G)| \geq 10$. Ker iščemo najmanjši možni snark in ker vemo, da na 10 vozliščih obstaja vsaj en snark, sledi, da ima G 10 vozlišč. Ker je

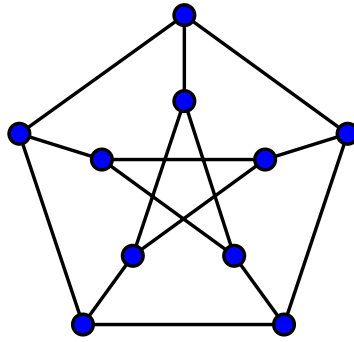


Tabela 36: Slika prikazuje Petersenov graf.

graf G kubičen in ker vemo, da ima 10 vozlišč, sledi, da ima G 15 povezav.

G je sedaj snark na 10 vozliščih in naj bo A spet neka ciklično po povezavah prerezna množica. Zanima nas velikost množice A . Vemo, da je velikost A manjša od 15, saj ima G 15 povezav. Ker je G snark je vsaj 4 ciklično po povezavah povezan. To pomeni, da je velikost $|A| \geq 4$. Predpostavimo da je $|A| = 4$. Potem $G - A$ vsebuje dve komponenti z 5 vozlišči. V obeh komponentah imamo 4 vozlišča stopnje 2 in 1 stopnje 3. Vozlišč take stopnje pa ne moremo povezati tako, da bi imeli cikel.

Če je $|A| > 5$, potem sledi, da je $|V(G)| > 10$. To pa je v protislovju z minimalnostjo grafa G . Torej je $|A| = 5$. Naj bo $A = \{i i', i = 1, 2, 3, 4, 5\}$. Ker je $E(G) = 15$ ima graf $G - A$ dva 5 cikla. Brez škode za splošnost je eden izmed teh ciklov $(1, 2, 3, 4, 5)$. Sedaj moramo še vozlišča $1', 2', 3', 4', 5'$ povezati tako, da dobimo 5 cikel in zadostimo zahtevi, da je G snark. Če bi povezali $1' 2'$ bi nam $(1, 2, 2', 1', 1)$ dalo 4 cikel. Kar je v nasprotju s predpostavko, da ima G po definiciji snarkov ožino 5. Če povežemo $1' 5'$ spet dobimo 4 cikel $(1, 1', 5', 5, 1)$. Torej je $1'$ povezan z $3'$ in $4'$. Podobno vidimo, da je $2' 4'$ in $2' 5'$ ter $3' 1'$. Torej smo res dobili $G = P$, kjer je P Petersenov graf.

□

Hipoteza 8.5. Vsak snark se da skrčiti na Petersenov graf.

Ne obstajajo snarki, ki bi imeli 12, 14 ali 16 vozlišč. Na 18 vozliščih obstajata natanko dva snarka, to sta Blanuševa snarka. Šest snarkov obstaja na 20 vozliščih in dvajset snarkov obstaja na 22 vozliščih. Ker

so snarki kubični grafi, po lemi o rokovanju sledi, da ne obstajajo snarki, ki bi imeli liho mnogo vozlišč.

8.2 ISSACOVI SNARKI

Do leta 1975 so bile znane že prej omenjene skupine snarkov. Leta 1975 pa je Issac objavil članek [3] v katerem je opisal konstrukcijo dveh neskončnih razredov snarkov. Eden od teh razredov vsebuje vse do takrat poznane snarke, razen Petersenovega grafa. Drugi razred pa vsebuje med drugim tudi cvetlične snarke. V [2] je omenjeno, da je te neodvisno odkril leta 1972 Grinberg, vendar njegovo delo ni bilo objavljeno. Po viru [2] je leta 1976 Loupkhine opisal še eno konstrukcijo, ki iz poznananih snarkov skonstruira nove. Issac pa je odkril še en snark, ki ne sodi v nobeno od opisanih konstrukcij. To je tako imenovani dvojni zvezda snark.

8.2.1 Točkast produkt

Definicija 8.6. Naj bosta L in R povezana kubična grafa in naj bosta a in b vozlišči v grafu R . Naj bodo vozlišča r, s, x, t, u, y iz grafa L , pri čemer sta vozlišči x in y sosednji. Vozlišče x je povezano z vozliščema r in s , vozlišče y pa s t in u . Točkast produkt dveh povezanih kubičnih grafov L in R označimo z $L \cdot R$ in je definiran kot

1. odstranimo neki par sosednjih vozlišč x in y iz L ,
2. odstranimo neodvisni povezavi ab in cd iz R ,
3. povežemo $\{r, s\}$ z $\{a, b\}$ ter vozlišči $\{t, u\}$, z $\{c, d\}$ ali povežemo $\{r, s\}$ z $\{c, d\}$ ter $\{t, u\}$ z $\{a, b\}$.

Točkast produkt je prikazan na sliki 37.

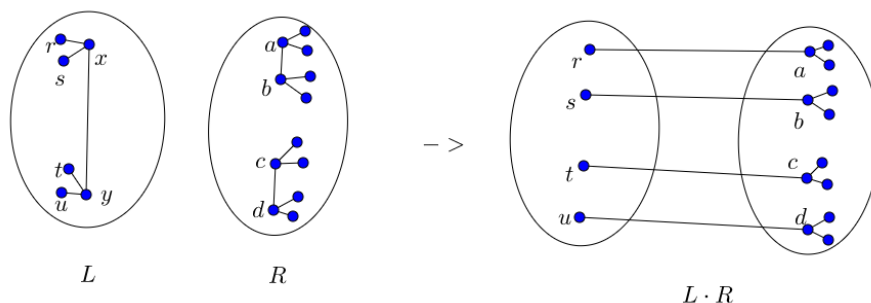


Tabela 37: Slika prikazuje točkast produkt med grafoma L in R .

Opomba 8.7. Točkasti produkt ni enolično določen. Odvisen je od izbire sosednjih vozlišč v grafu L in izbire neodvisnih povezav v R . Odvisen je tudi kako povežemo vozlišča, da dobimo $L \cdot R$.

Očitno je, da če je L reda m in R reda r , da je točkasti produkt $L \cdot R$ reda $m + r - 2$.

Izrek 8.8. *Predpostavimo, da velja $\chi'(L) = \chi'(R) = 4$. Potem sledi, da je $\chi'(L \cdot R) = 4$.*

Pri dokazu bomo potrebovali naslednjo lemo, ki je ne bomo dokazali.

Lema 8.9. *Naj bo G kubični multigraf, ki je 3 povezan po povezavah z barvami c_i , $i = 1, 2, 3$. Naj bo Z neka prerezna množica povezav v G in n_i število povezav pobarvanih z barvo c_i v prerezni množici Z . Potem so števila n_i kongruentna po modulu 2.*

Dokaz. Predpostavimo, da je $\chi'(L \cdot R) = 3$. Potem po prejšni lemi sledi, da so povezave ar, bs, ct, du lahko pobarvane na enega od naslednjih načinov

1. $(1, 1, 1, 1)$,
2. $(1, 1, 2, 2)$,
3. $(1, 2, 1, 2)$.

Podrobneje si oglejmo prvo barvanje, to je barvanje $(1, 1, 1, 1)$. Potem lahko povezavi ab in cd , ki sta v grafu R pobarvamo z barvo 1, vse ostale povezave v grafu R pa pobarvamo enako, kot so pobarvane povezave iz R , ki so v grafu $L \cdot R$. Tako smo dobili 3 barvanje povezav grafa R , kar pa je v nasprotju s predpostavko, da je $\chi' = 4$.

Podobno naredimo pri ostalih dveh barvanjih. Iz obeh barvanj lahko skonstruiramo pravilno 3 barvanje grafa R , kar je v protislovju s predpostavko, da je $\chi'(R) = 4$. \square

Po tej trditvi sledi, da če sta L in R snarka, je potem tudi $L \cdot R$ snark. Preveriti bi bilo še potrebno, da če imata L in R ciklično povezanost po povezavah 4, da ima potem tudi $L \cdot R$ ciklično povezanost po povezavah 4. In preveriti bi še morali, da ima $L \cdot R$ ožino 5.

Po tem postopku dobimo oba Blanuševa grafa. Szekerseov graf pa je $P(P(P(P(P \cdot P))))$.

Lahko pa je $L \cdot R$ snark, čeprav nista ne L in ne R snarka. Obstajajo tudi konstrukcije snarkov, ki skonstruirajo snarke iz grafov, ki nimajo kromatičnega indeksa 4. Te dobimo s pomočjo zvezdastega produkta.

8.2.2 Zvezdast produkt

Definicija 8.10. Naj bosta L in R povezana kubična grafa in naj bodo vozlišča r, s, t, u iz grafa L , pri čemer so vozlišča r, s in t povezana z u . Nadalje naj bodo vozlišča a, b, c, v iz grafa R in sicer taka, da je v povezan z vozlišči a, b in c . Zvezdast produkt $L * R$ dobimo takole

1. izbrišemo vozlišče u iz grafa L in vozlišče v iz grafa R skupaj z njunimi povezavami,
2. povežemo vozlišča $\{r, s, t\}$ z vozlišči $\{a, b, c\}$.

Primer zvezdastega produkta je prikazan na sliki 38.

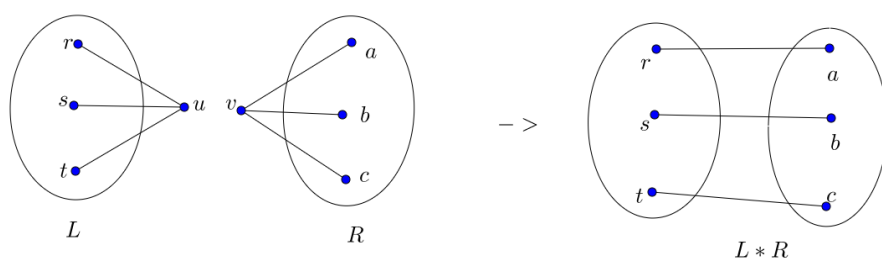


Tabela 38: Slika prikazuje zvezdast produkt med grafoma L in R .

Opomba 8.11. Zvezdasti produkt ni enolično določen.

Trditev 8.12. S pomočjo zvezdastega produkta lahko skonstruiramo snarke iz grafov, ki niso snarki.

Dokaz. Naj bosta S_1, S_2 snarka in A_1 ter A_2 povezana kubična grafa. Naj bo graf $G = S_2 \cdot (A_1 * (S_1 * A_2))$. Pri točkastem produktu zbrisemo dve sosednji vozlišči v S_2 in povezavo med njima ter eno povezavo iz A_1 in eno povezavo iz A_2 . V zvezdastem produktu pa izbrišemo dve različni vozlišči iz S_1 ter eno iz A_1 in eno iz A_2 . Dobimo graf, ki je prikazan na sliki 39. Ker je G 3-regularen graf ima po Vizingovem izreku ali kromatičen indeks enak 3 ali 4. Predpostavimo, da je $\chi'(G) = 3$ in vzamemo neko 3-barvanje po povezavah grafa G . Naj bodo $\{e_1, e_2, e_3, e_4\}$ povezaven kot vidimo na sliki 39. Po lemi, ki smo jo prej omenili, povezave $\{e_1, e_2, e_3, e_4\}$, ki jih pobarvamo v tem vrstnem redu na tri načine:

1. $(1, 2, 1, 2)$,
2. $(1, 1, 1, 1)$,
3. $(1, 1, 2, 2)$.

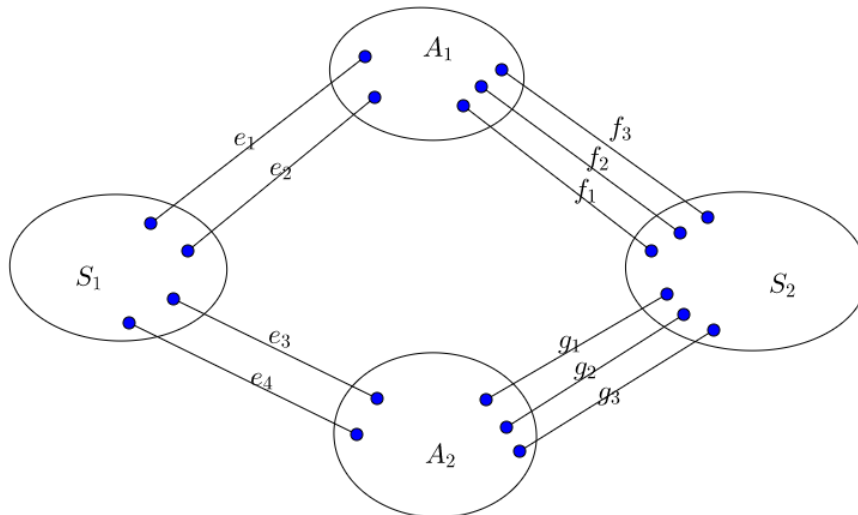


Tabela 39: Slika nam prikazuje graf $G = S_2 \cdot (A_1 * (S_1 * A_2))$.

Podrobneje si oglejmo barvanje $(1, 2, 1, 2)$. To barvanje nam inducira 3-barvanje grafa S_2 . To pa je v nasprotju s predpostavko, da je S_2 snark in ima torej kromatičen indeks 4. Torej sta možni samo barvanji $(1, 1, 1, 1)$ ter $(1, 1, 2, 2)$. Lahko smo torej prepričani, da sta povezavi e_1 in e_2 pobarvani z enako barvo ter, da sta tudi povezavi e_3 in e_4 pobarvani z enako barvo.

Če upoštevamo lemo in prerezno množico po povezavah $\{e_1, e_2, f_1, f_2, f_3\}$ sledi, da imajo f_1, f_2, f_3 različne barve. Podobno vidimo, da so g_1, g_2, g_3 pobarvane z različnimi barvami. To nam inducira 3-barvanje po povezavah grafa S_1 . Ker je S_1 po predpostavki snark in ima zato kromatičen indeks enak 4, pridemo v protislovje. Torej je predpostavka $\chi'(G) = 3$ napačna. Sledi, da je $\chi'(G) = 4$. \square

Torej lahko skonstruiramo nove snarke iz A_1 in A_2 , čeprav A_1 in A_2 nista snarka. Zahtevati pa moramo, da sta A_1 in A_2 izbrana tako, da ima graf G zahtevano ožino in ciklično povezanost. Ta proces lahko ponovimo in sicer definiramo graf $G' = G \cdot (A_3 * (S_3 * A_4))$. Na podoben način pokažemo, da je $\chi'(G) = 4$. Spet predpostavimo, da je $\chi'(G) = 3$ in z uporabo leme pridemo v protislovje, tako da dobimo 3-barvanje grafov, ki so snarki.

S pomočjo teh idej lahko skonstruiramo Descartesov graf, ki je na sliki 40.

Slika 40 je vzeta iz vira [7].

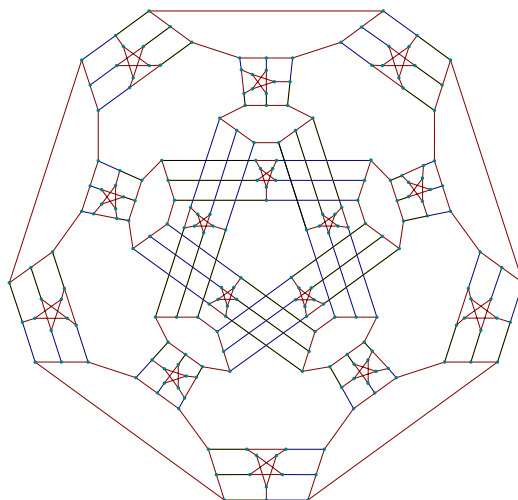


Tabela 40: Slika prikazuje Descartesov graf.

8.2.3 Cvetlični snarki

Definiramo graf J_k , kjer je k liho število strogo večje od 3. Naj bo $I(m) = \{0, 1, \dots, m-1\}$ množica indeksov. Definiramo množico vozlišč

$$V(J_k) = \{u_i, v_i : i \in I(k)\} \cup \{w_j : j \in I(2k)\}$$

in množico povezav

$$\begin{aligned} E(J_k) = & \{u_i u_{i+1}, u_i v_i, w_j w_{j+1} : i \in I(k), j \in I(2k)\} \\ & \cup \{v_i w_j : i \equiv j \pmod{k}, i \in I(k), j \in I(2k)\}. \end{aligned}$$

Primer takega grafa je prikazan na sliki 41, ki je vzeta iz vira [4].

Opomba 8.13. Če dovolimo $k = 3$, potem J_3 vsebuje trikotnik in ni snark. Ampak, če skrčimo trikotnik na eno vozlišče, dobimo Peterse-
nov graf.

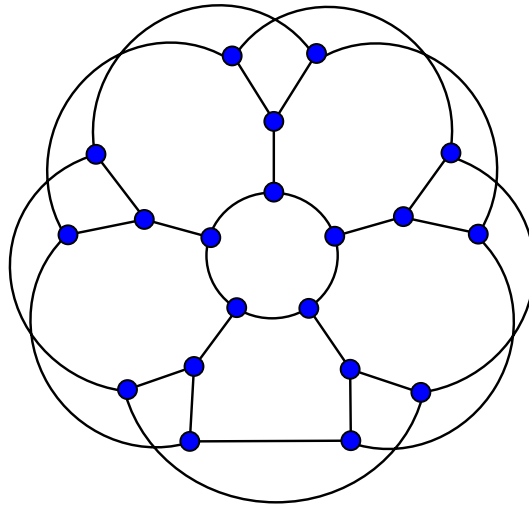


Tabela 41: Slika prikazuje cvetlični snark.

8.2.4 Dvojni zvezdast snark

Issac je našel še en snark, ki ne spada pod nobenega od prej omenjenih razredov. To je dvojni zvezdast snark. Graf je prikazan na sliki 42, ki je vzeta iz vira [8].

8.3 k -SNARKI

Definicija 8.14. Za graf rečemo, da je k -snark, če je k regularen, ciklično vsaj $k + 1$ po povezavah povezan, ki ne vsebuje klike reda $k + 1$ minus 1-faktor ali klike reda k .

Hipoteza 8.15. Vsak k -snark vsebuje podgraf, ki se da skrčiti na Petersenov graf.

8.4 ZAKLJUČEK

V seminarski nalogi smo predstavili snarke in motivacijo, da so jih začeli iskati. Predstavili smo snarke, ki so bili znani preden je Issac v svojem delu predstavil konstrukcije snarkov. Definirali smo točkasti produkt in dokazali, da če sta grafa L in R snarka je tudi $L \cdot R$ snark. S pomočjo zvezdastega produkta pa smo konstruirali snarke iz grafov,

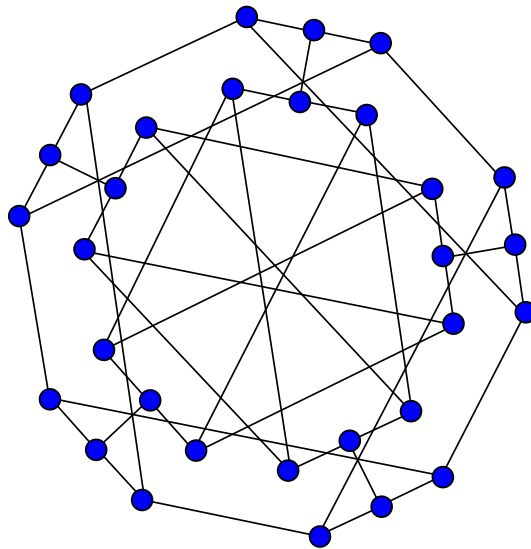


Tabela 42: Slika prikazuje dvojni zvezdast snark.

ki niso nujno snarki. Opisali smo še cvetlične snarke in dvojni zvezdast snark. Na kratko smo še omenili k -snarke. Seminarska naloga predstavlja nek problem, ki v osnovi izgleda lahek, iskanje vseh kubičnih grafov, ki bi imeli kromatičen indeks 4. Problem se je izkazal za težkega in kar nekaj časa je trajalo preden je bila najdena konstrukcija takih grafov. Še sedaj imamo dve odprti hipotezi in sicer, če se dajo snarki skrčiti na Petersenov graf in, če k -snarki vsebujejo podgraf, ki se da skrčiti na Petersenov graf. Snarki so se izkazali za res zmuzljive grafe.

LITERATURA

- [1] M. Gardner, *Scientific American.*, Mathematical Games, 234, št. 40, (1976), 126-130 in št. 9, (1976), 210-211.
- [2] D. A. Holton, J. Sheehan, *Australian Mathematical Society Lecture Series 7.*, The Petersen Graph, (1993), 79-107.
- [3] R. Issac, Infinite families of nontrivial graphs which are not Tait colourable, *Amer. Math. Monthly*, 82, (1972), 221-239.
- [4] Kokoro, *Wikimedia Commons*, CC BY-SA 3.0, dostopno na <https://commons.wikimedia.org/w/index.php?curid=7399602>.
- [5] Kokoro, *Wikimedia Commons*, CC BY-SA 3.0, dostopno na <https://commons.wikimedia.org/w/index.php?curid=7398789>.
- [6] Leshabirukov, *Wikimedia Commons*, CC BY-SA 3.0, dostopno na <https://commons.wikimedia.org/w/index.php?curid=5788203>.
- [7] MathsPoetry, *Wikimedia Commons*, CC BY-SA 3.0, dostopno na https://commons.wikimedia.org/wiki/File:Blanche_Descartes_snark.svg.
- [8] Ed Jr. Pegg, iz *MathWorld-A Wolfram Web Resource*, ustvaril Eric W. Weisstein, dostopno na <http://mathworld.wolfram.com/DoubleStarSnark.html>.

HIPOHAMILTONOVI GRAFI

Gregor Rus

Glavna tema seminarja bo reševanje problema, ki ga je predstavil Rene Sousselier v svojem članku [1].

V podeželskem klubu je predsednik organiziral jubilejno srečanje. Nanj je želel povabil vse člane kluba. Vsi člani kluba so enakovredni, zato se je odločil, da bodo sedeli pri okrogli mizi. Vendar se v klubu vsi ne razumejo med seboj. Vsak član ima v klubu nekaj prijateljev in nekaj članov, s katerimi se ne razume. Predsednik si je želel, da bi se na srečanju vsi imeli dobro, zato je poskušal urediti vrstni red tako, da bi imel vsakdo imel kot svoja soseda svoja prijatelja.

Potem, ko se je nekaj časa trudil narediti ustrezni sedežni red, je na pomoč poklical prijatelja matematika. Po kratkem razmisleku mu je matematik podal naslednji odgovor: „Željen sedežni red je nemogoč. Lahko pa ti uspe, če enega od članov prepričaš, da ne pride.“ Predsednik je vprašal: „Katerega pa moram prepričati, da ne pride?“ In dobil odgovor: „Vseeno je, katerikoli bo dober. Mimogrede, če bi imel manj članov kluba, tega problema ne bi imel.“

Predsednik je našel izgovor, da ga na večerjo ni bilo in uspel sestaviti ustrezni vrstni red.

Glavno vprašanje, ki si ga je zadal avtor članka je, kolikšno je število članov kluba. Prav tako se je vprašal, če je s tem podatkom tudi določeno, kdo koga ne mara (do simetrija natančno).

Za odgovor na vprašanji si bomo ogledali družino *hipohamiltonovih grafov*. V seminarju bomo raziskali tudi druge lastnosti te družine.

9.1 OSNOVNE DEFINICIJE

Najprej ponovimo definicijo *hamiltonovega grafa*.

Definicija 9.1. Naj bo G poljuben graf in $|V(G)| = n$. Če obstaja v grafu cikel $v_1v_2 \cdots v_n$ skozi vsa vozlišča grafa, tak cikel imenujemo *hamiltonov cikel*. Če ima G hamiltonov cikel, je G *hamiltonov*.

Za lažje označevanje grafov v nadaljevanju seminarja, vpeljimo še naslednjo oznako.

Opomba 9.2. Naj bo G graf in $u \in V(G)$. Z $G - u$ označimo graf, ki ga dobimo, če G izbrišemo u in vse povezave, ki imajo eno krajišče v u .

Problem podan v uvodu bomo obravnavali s pomočje teorije grafov. Za vozlišča grafa izberimo člane kluba. Če sta dva člana prijatelja, povežimo vozlišča, ki prikazujeta ta dva člana. Če bi takšen graf imel hamiltonov cikel, bi lahko vse člane posedli za mizo v zaporedju hamiltonovega cikla. Graf članov in prijateljstev kluba G pa ima lastnost, da je hamiltonov, če mu izbrišemo vozlišče in vse povezave, ki imajo eno krajišče v v in sicer za poljuben $v \in V(G)$.

Definicija 9.3. Naj bo G graf, ki ni hamiltonov. Če velja, da je za vsak $v \in V(G)$, $G - v$ je hamiltonov, je G *hipohamiltonov graf*.

Rešitev našega problema bo najmanjši hipohamiltonov graf. Za iskanje najmanjšega hipohamiltonovega grafa bomo potrebovali nekaj lastnosti hipohamiltonovih grafov. Najprej pogledjmo mejo za najmanjšo stopnjo v grafu.

Lema 9.4. Če je G hipohamiltonov, potem je $\delta(G) \geq 3$, kjer $\delta(G)$ označuje najmanjšo stopnjo v grafu.

Dokaz. Denimo najprej, da obstaja v grafu izolirano vozlišče u . Potem za vsako drugo vozlišče u grafa G , graf $G - v$ ne more vsebovati hamiltonovega cikla, ker ni povezan. Če obstajav grafu vozlišče u z $\deg(u) = 1$, potem u ne leži na nobenem ciklu, zato graf G ne more biti hipohamiltonov. Sedaj lahko predpostavimo $\delta(G) \geq 2$. Vzemimo u in v sosednji vozlišči v G . Ker je $G - v$ hamiltonov, je u del hamiltonovega cikla v $G - v$. Torej je u sosedn z vsaj dvema vozliščema na ciklu in z v . □

Tudi največjo stopnjo v hipohamiltonovem grafu znamo omejiti.

Lema 9.5. Če je G hipohamiltonov, potem je $\Delta(G) \leq \lfloor \frac{1}{2}(n-1) \rfloor$, kjer $\Delta(G)$ označuje največjo stopnjo v grafu.

Dokaz. Vzemimo vozlišče $v \in V(G)$ v grafu, ki ima stopnjo večjo kot $\lfloor \frac{1}{2}(n-1) \rfloor$. Graf $G - v$ je hamiltonov s ciklom u_1, \dots, u_m . Ker ima v vsaj $\lfloor \frac{1}{2}(n-1) \rfloor + 1$ sosedov, je sosednje z dvema zaporednima vozliščema v ciklu, npr. z u_i in u_{i+1} . Potem je $u_1, \dots, u_i, v, u_{i+1}, \dots, u_m$ hamiltonov cikel v G , kar je protislovje. □

Pokažemo lahko tudi, da je vsak hipohamiltonov graf 3-povezan po vozliščih.

Lema 9.6. *Naj bo G hipohamiltonov graf. Potem je G po vozliščih 3-povezan.*

Dokaz. Če odstranimo poljubno vozlišče iz grafa, je tako dobljeni graf hamiltonov. Torej lahko potem odstranimo še vsaj eno vozlišče in je graf še vedno povezan. \square

Zanima nas tudi, vsaj koliko vozlišč potrebujemo v grafu.

Lema 9.7. *Naj bo G hipohamiltonov. Potem je $|V(G)| \geq 10$.*

Dokaz. Z upoštevanjem 9.4 in 9.5 mora biti $|V(G)|$ vsaj 7, sicer dobimo, da je maksimalna stopnja manjša od minimalne. V primeru $|V(G)| = 7$ želimo dobiti kubični graf na 7 vozliščih, za katerega (po lemi o rokovanju) vemo, da ne obstaja. Torej $|V(G)| \geq 8$.

Recimo, da je $|V(G)| = 8$. Iz 9.4 in 9.5 dobimo, da je G kubičen graf. Izberemo vozlišče iz $V(G)$ in ga označimo z v_0 . Med preostalimi vozlišči dobimo cikel, ker je G hipohamiltonov. Vozlišča cikla označimo z v_1, v_2 do v_7 . Tudi vozlišče v_0 je stopnje 3. Če je vozlišče v_0 sosednje z dvema zaporednima vozliščema ali z v_1 in v_7 , je G hamiltonski (torej ni hipohamiltonov). Premislimo najprej, da v G ni trikotnikov.

Če bi v G imeli trikotnik, ki ne vsebuje v_0 , bi imeli vozlišče stopnje 4. Če pa obstaja v G trikotnik, ki vsebuje v_0 , oblike v_i, v_{i+1}, v_0 , potem je G tudi hamiltonski.

Sedaj lahko brez škode za splošnost predpostavimo, da sta zaporedni vozlišči na ciklu s katerimi v_0 ni povezano, v_1 in v_2 . Potem ima v_0 sosedne v_3, v_5 in v_7 . Ker je graf kubičen morajo biti v njemu še povezave $v_1 \sim v_4$ in $v_2 \sim v_6$ (drugače bi imeli trikotnik). V tem primeru imamo v G hamiltonov cikel $v_1, v_4, v_3, v_2, v_6, v_5, v_0, v_7$. Skico našega primera si lahko na sliki 43.

Pokazati moramo še, da ne moremo imeti hipohamiltonovega grafa na 9 vozliščih. Za lažje razumevanje spremljamo sliko 44.

Stopnja vseh vozlišč je med 3 in 4 po prejšnjih lemah. Ker ne obstaja kubični graf na liho mnogo vozliščih, obstaja vsaj eno vozlišče stopnje 4. Vozlišče stopnje 4 označimo z v_0 . Ostala vozlišča tvorijo cikel. Označimo jih z od v_1 do v_8 , podobno kot v prejšnjem primeru. Brez izgube splošnosti lahko predpostavimo, da je v_0 povezano z lihimi vozlišči.

Če je vozlišče v_2 povezano s katerim izmed sodih vozlišč, dobimo cikel čez vseh 9 vozlišč. Torej je v_2 povezano bodisi z v_5 ali z v_7 . Predpostavimo z v_5 . Podoben razmislek velja tudi za vozlišče v_8 . Torej je v_8 povezano s v_3 (z v_5 ne more biti zaradi največje stopnje). Potem mora biti v_6 povezano z v_1 in v_4 s v_7 . Dobljen graf je dvodelen (v

eni množici so liha vozlišča, v drugi pa soda). Graf $G - v_1$ je potem podgraf $K_{5,3}$ in zato ne more biti hamiltonov. \square

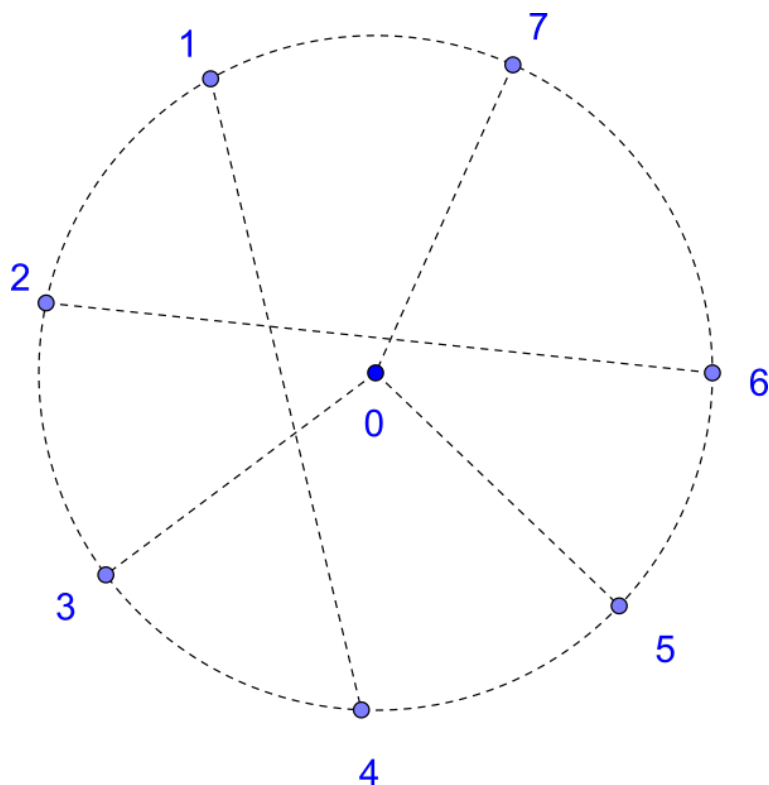


Tabela 43: Poskus konstrukcije hipohamiltonovega grafa na 8 vozliščih.

Sedaj vemo, da je minimalno število vozlišč v hipohamiltonovem grafu 10. Izkaže se, kot nam pove naslednji izrek, da hipohamiltonov graf na 10 vozliščih obstaja in da je določen do izomorfizma natančno.

Izrek 9.8. Če je G hipohamiltonov in $|V(G)| = 10$, potem je G izomorfen Petersenovemu grafu.

Dokaz. Naj bo G hipohamiltonov in naj bo $|V(G)| = 10$. Imamo dve možnosti: bodisi v G obstaja vozlišče stopnje 4, bodisi je G kubičen graf. Najprej si oglejmo možnost, kjer je v G vozlišče stopnje 4.

Podobno kot v prejšnjem dokazu označimo vozlišče stopnje 4 z v_0 , poljubno vozlišče izmed preostalih z v_1 in potem njegove sosedne na ciklu po vrsti (vemo, da imamo cikel na vseh ostalih): Ker je vozlišče v_0 stopnje 4, je sosednje s 4 vozlišči na ciklu. Brez izgube splošnosti predpostavimo, da je sosednje z v_1, v_3, v_5, v_7 . Tudi preostala vozlišča morajo imeti še vsaj enega sosedo. Če je v grafu povezava med v_2 in v_9 , potem dobimo cikel $v_1, v_2, v_9, v_8, v_7, v_6, v_5, v_4, v_3, v_0, v_1$. Če je v grafu povezava med v_2 in v_8 , dobimo cikel $v_0, v_1, v_9, v_8, v_2, v_3, v_4, v_5, v_6, v_7, v_0$.

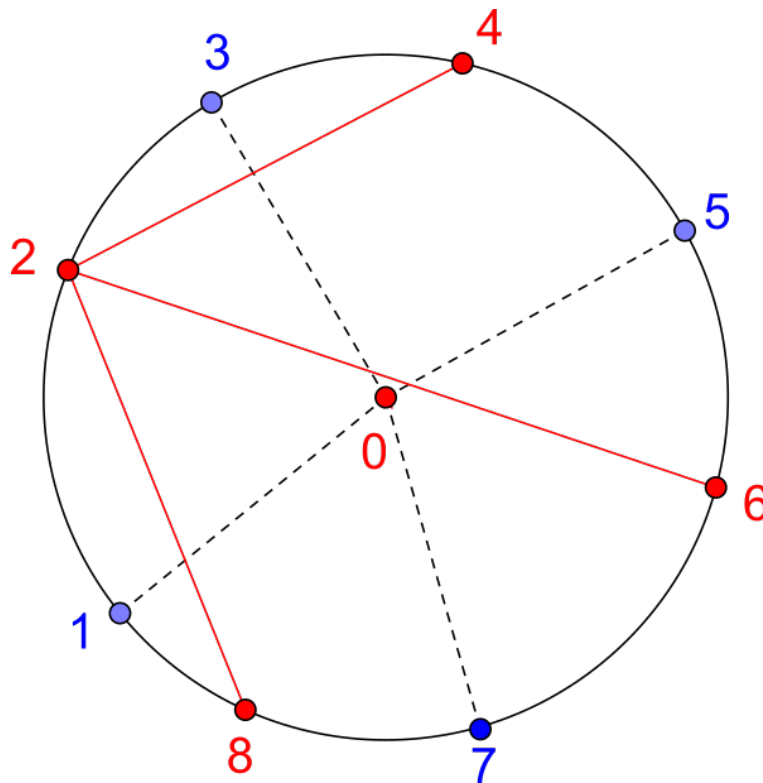


Tabela 44: Hipohamiltonov graf na 9 vozliščih bi bil dvodelen.

Ostali možnosti sta simetrični. Tudi vsa izmed teh pet vozlišč podajo simetrično situacijo. Torej mora biti vsako izmed vozlišč v_2, v_4, v_6, v_8, v_9 povezano z vsaj enim izmed vozlišč v_1, v_3, v_5 ali v_7 , posledično je vsaj eno izmed teh povezano z dvemi. To pa je protislovje, ker je maksimalna stopnja v grafu 4. Sliko razmislek nam predstavlja slika 45. Torej je hipohamiltonov graf na 10 vozliščih kubičen.

Znano dejstvo je, da je Petersenov graf vozliščno tranzitiven in da ni hamiltonov. Ker je vozliščno tranzitiven, velja, da je $P - v$ izomorfen $P - u$ za vsak par vozlišč u in v v grafu. Zato je dovolj pokazati, da je $P - v$ hamiltonov za en v . Oglejmo si sliko 46. Hamiltonov cikel je na grafu obarvan. Torej P je hipohamiltonov.

Vsak hipohamiltonov graf G je 3-povezan po lemi 9.6. Ker pa je Petersenov graf edini 3-povezan kubični graf na 10 vozliščih [2], ki ni hamiltonov, je G izomorfen P . \square

Opomba 9.9. S pomočjo programske opreme Sage lahko tudi sami preverimo zadnjo karakterizacijo.

Na tak način smo odgovorili na vprašanje, ki je bilo zastavljeno v uvodu. V klubu je deset članov, vsak ima tri prijatelje, ki jih ponazarjajo povezave v Petersenovem grafu.

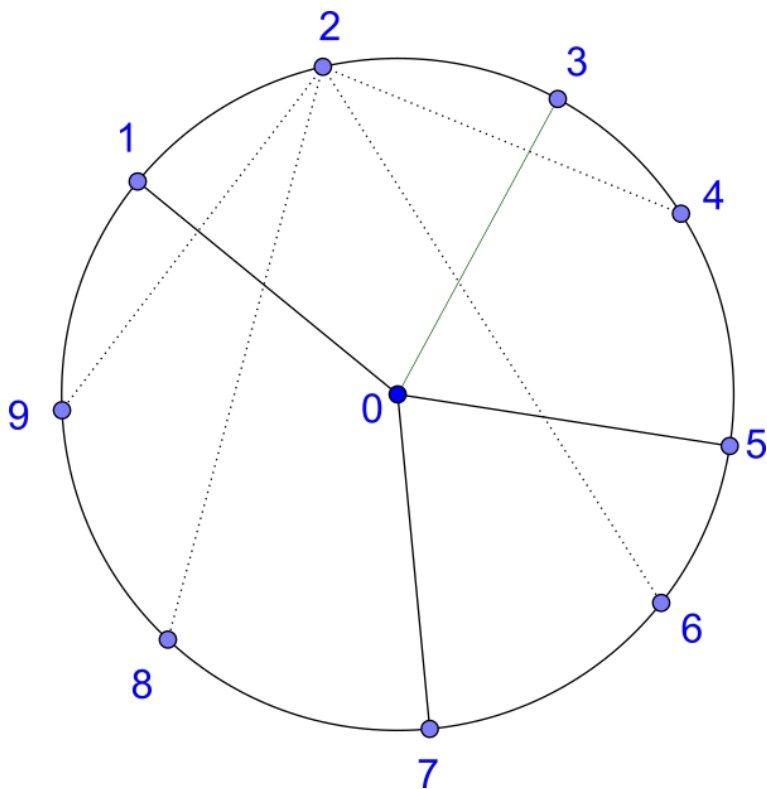


Tabela 45: Kako izgleda hipohamiltonov graf na 10 vozliščih.

9.2 RAZISKOVANJE HIPOHAMILTONOVIVH GRAFOV

Izkaže se, da obstaja neskončno mnogo hipohamiltonovih grafov. Pri raziskovanju (večinoma je potekalo s pomočjo računalnika) so ugotovili, da obstaja hipohamiltonov graf z n vozlišči, za vsak $n \geq 10$ z izjemo 11,12,14 in 17 [2] (zadnji rezultat je bil dolgo samo domneva). Trenutno je problem iskanja najmanjšega ravninskega hipohamiltonovega grafa. Najmanjši najden ima 42 vozlišč, vidimo ga na sliki 47, ki jo najdemo na internetu [4].

Poglejmo si eno izmed posplošitev Petersenovega grafa, ki generira neskončno družino hipohamiltonovih grafov. To posplošitev je prvi predstavil matematik Lindgren.

Definicija 9.10. Za $k \geq 1$ definiramo $L(k)$ kot graf z:

- $VL(k) = \{v_0, v_1, \dots, v_{6k+2}\} \cup \{v\}$,
- $EL(k) = \{v_i \sim v_{i+1}\} \cup \{v_{1+3j} \sim v_{5+3j}, v \sim v_{3j}; j \in \{0, 1, \dots, 2k\}\}$.

Opomba 9.11. Seštevanje v definiciji poteka modulu $6k + 2$.

Zlahka pokažemo, da je $L(1)$ hipohamiltonov, ker je izomorfen Petersenovemu grafu. Pokazati želimo še, da je $L(k)$ hipohamiltonov za vsak $k \geq 2$.

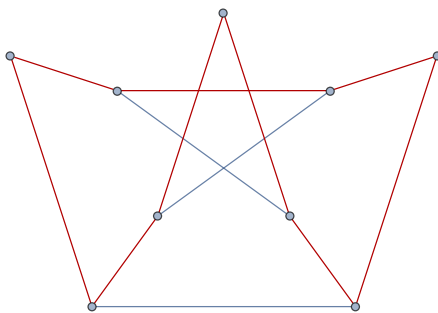


Tabela 46: Hamiltonova pot v Petersenovem grafu, ki mu izbrišemo eno vozlišče.

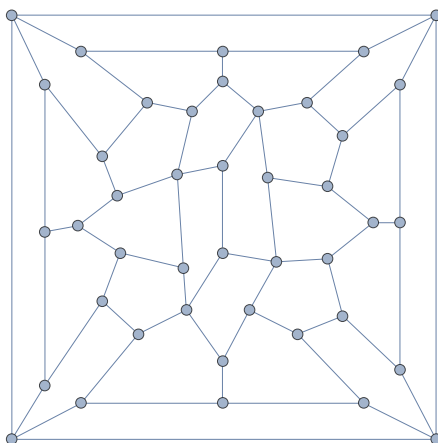


Tabela 47: Trenutno najmanjši najden ravninski hipohamiltonov graf.

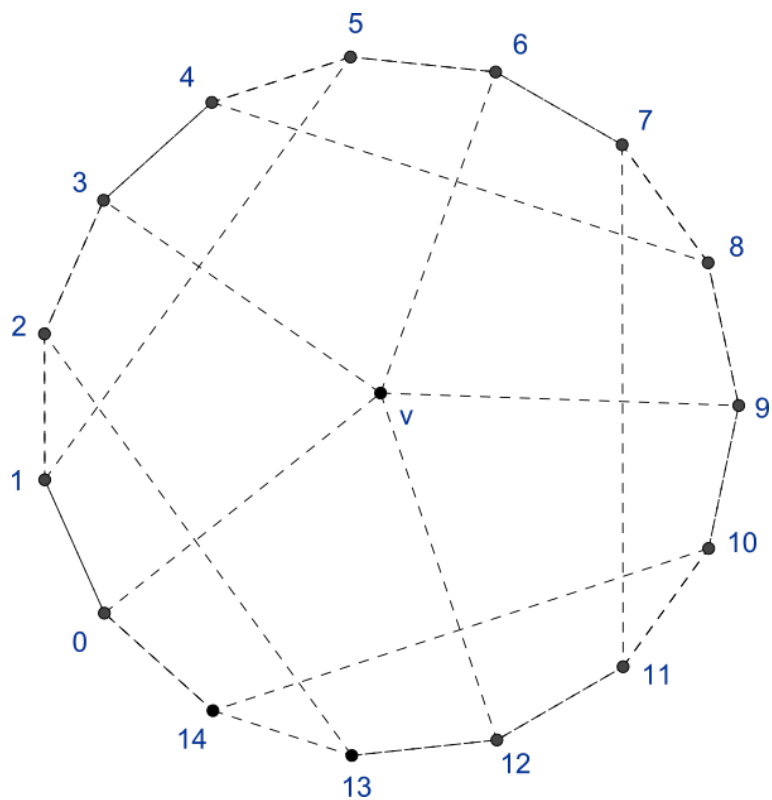
Trditev 9.12. Za vsak $k \geq 2$ je $L(k) - v$ hipohamiltonov, kjer je v odlikovano vozlišče grafa $L(k)$.

Dokaz. Po definiciji opazimo cikel $v_0, v_1, \dots, v_{6k+2}$. □

Opazimo, da so si simetrična vozlišča, ki so enakovredna modulu 3. Torej lahko pogledamo grafe $L(k) - v_0, L(k) - v_1$ in $L(k) - v_2$. Zaradi boljše predstave bomo poiskali hamiltonove cikle v teh grafih na primeru $k = 2$, za lažjo predstavo pa si oglejmo še sliko 48, sliko grafa $L(2)$.

- V grafu $L(2) - v_0$ opazimo cikel: $v_1, v_5, v_4, v_8, v_9, v_{10}, v_{14}, v_{13}, v_{12}, v_{11}, v_7, v_6, v, v_3, v_2, v_1$.
- V grafu $L(2) - v_1$ opazimo cikel: $v_2, v_{13}, v_{14}, v_0, v, v_{12}, v_{11}, \dots, v_3, v_2$.
- V grafu $L(2) - v_2$ opazimo cikel: $v_1, v_0, v_{14}, \dots, v_6, v, v_3, v_4, v_5, v_1$.

Analogno dobimo za vsak k . Dokaz, da $L(k)$ ni hamiltonov je sicer lahek, a precej obsežen. Najdemo ga v [3]. Tako smo našli neskončno družino hipohamiltonovih grafov.

Tabela 48: Graf $L(2)$.

9.3 ZAKLJUČEK

Teorija grafov nam ponuja ogromno problemov in smeri, ki jih lahko raziskujemo. Tako smo v tem seminarju spoznali še eno zanimivo lastnost grafov. Zanimivo je, da se kot poseben primer spet pojavi Petersenov graf, ki v tej vlogi nastopa v mnogih problemih. Videli smo tudi, kako lahko matematiko uporabimo v realnem (oz. na pol realnem) življenju. Morda vso to dokazovanje in učenje definicij le ni tako neuporabno.

LITERATURA

- [1] R. Sousselier, *Revue Française de Recherches Opérationelle* 7, Probleme no. 29: Le cercle des irascibles, (1963): 405-406.
- [2] D. A. Holton, J. Sheehan, *Australian Mathematical Society Lecture Series* 7, The Petersen Graph, (1993), 214-244.
- [3] W. F. Lindgren, *American Mathematical Monthly (Mathematical Association of America)* 74, An infite class of hypohamiltonian graphs, (1967), 1087-1089.
- [4] Weisstein, Eric W. "Wiener-Araya Graph." From MathWorld—A Wolfram Web Resource. <http://mathworld.wolfram.com/Wiener-ArayaGraph.html>

LIMITE GRAFOV

Sven Ceršk

Pri preučevanju velikih omrežij pogosto ne moremo podati celotnega grafa, ki predstavlja omrežje, saj bi bil ta prevelik. Zato opazujemo naključno generirane manjše grafe, ki imajo podobne lastnosti kot celotno omrežje.

Glavno vprašanje te seminarske naloge je kdaj sta si dva grafa podobna, saj na ta način lahko nekemu velikemu omrežju, ki ga ne moremo natančno predstaviti, priredimo manjše (naključne) grafe, ki izkazujejo podobne lastnosti kot celotno omrežje.

Ideje predstavljene v tej seminarski nalogi so uporabne tudi na področju ekstremalne teorije grafov, ki se ukvarja z iskanjem ekstremalnih (maksimalnih ali minimalnih) grafov, ki zadoščajo neki lastnosti. Primer je vprašanje "Koliko 4-ciklov ima graf, ki ima več kot polovico vseh možnih povezav?" Pogosto nas zanima le asimptotično obnašanje določenih lastnosti pri naraščajoči velikosti grafa. Tako spet naletimo na podobne probleme kot pri preučevanju velikih omrežij.

10.1 LIMITE GRAFOV

Opazujemo zaporedje grafov z naraščajočim številom vozlišč in želimo ugotoviti ali je takšno zaporedje konvergentno. Konvergenco si v tem kontekstu lahko interpretiramo kot to, da so si grafi z naraščajočo velikostjo bolj in bolj podobni. Osredotočili se bomo na zaporedja grafov, kjer je gostota povezav z naraščajočo velikostjo neničelna. To pomeni, da stopnje vozlišč z naraščajočo velikostjo grafa prav tako naraščajo. Takšnim grafom pravimo gosti grafi (dense graphs). Nasprotno za redke grafe (sparse graphs) velja, da je stopnja vozlišč omejena z neko konstanto neodvisno od velikosti grafa. Razlike v obnašanju gostih in redkih grafov so zelo velike, zato opisani pristop deluje le na gostih grafih.

V nadaljevanju si bomo ogledali, kako definirati podobnost in konvergenco zaporedji grafov.

10.2 GOSTOTA HOMOMORFIZMOV IN KONVERGENCA GRAFOV

Definicija 10.1 (Homomorfizem). Naj bosta H in G grafa. Preslikava $\phi : V(H) \rightarrow V(G)$ je homomorfizem, če velja $\{u, v\} \in E(H) \implies \{\phi(u), \phi(v)\} \in E(G)$.

S $\text{Hom}(H, G)$ označimo množico vseh homomorfizmov iz grafa H v G .

$$\text{Hom}(H, G) = \{\phi : V(H) \rightarrow V(G) \mid \phi \text{ je homomorfizem}\}.$$

Število vseh homomorfizmov med H in G označimo s

$$\text{hom}(H, G) = |\text{Hom}(H, G)|.$$

Definicija 10.2 (Gostota homomorfizmov). Za grafa H in G definiramo gostoto homomorfizmov $t(H, G)$ kot kvocient med številom vseh homomorfizmov iz H v G in številom vseh možnih preslikav $V(H) \rightarrow V(G)$

$$t(H, G) = \frac{\text{hom}(H, G)}{|V(G)|^{|V(H)|}}.$$

Če vzamemo nek določen graf H in si ogledamo gostoto homomorfizmov iz tega grafa v druge grafe $t(H, \cdot)$, lahko preostale grafe ločimo glede na to gostoto.

Če pogledamo dva grafa G_1 in G_2 in za njiju velja, da je gostota $t(H, G_1)$ blizu gostote $t(H, G_2)$ to pomeni, da sta si vsaj v nekem pogledu podobna. Namesto enega grafa H lahko opazujemo gostote več grafov $(t(H_1, \cdot), t(H_2, \cdot), \dots)$, kar še natančneje potrjuje podobnost grafov G_1 in G_2 , če $t(H_1, G_1) \approx t(H_1, G_2)$, $t(H_2, G_1) \approx t(H_2, G_2)$ in tako dalje. Namesto neke določene množice grafov $\{H_i\}$ pa lahko opazujemo tudi kar vse grafe.

Za zaporedje grafov $(G_n)_n$ lahko rečemo, da je konvergentno, če za vsak graf H velja, da je zaporedje gostot $(t(H, G_n))_n$ konvergentno.

S tem pa se odpre novo vprašanje, kako izgledajo takšna konvergentna zaporedja grafov. V ta namen si je dobro ogledati matrike sosednosti grafov v zaporedju.

Iz matrike sosednosti lahko narišemo sliko tako, da vsako enico zamenjamo s kvadratom črne barve, vsako ničlo pa s kvadratom bele barve. Primer je prikazan na sliki 49.

V primerih na sliki 50 so prikazani različni tipi konvergentnih grafov na 10, 50 in 100 vozliščih. V prvi vrsti so naključni grafi z verjetnostjo povezave $\frac{1}{2}$, v drugi grafi na vozliščih $V = \{1, 2, \dots, n\}$, kjer

$$\begin{bmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix}$$

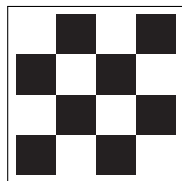


Tabela 49: Primer slike 4-cikla.

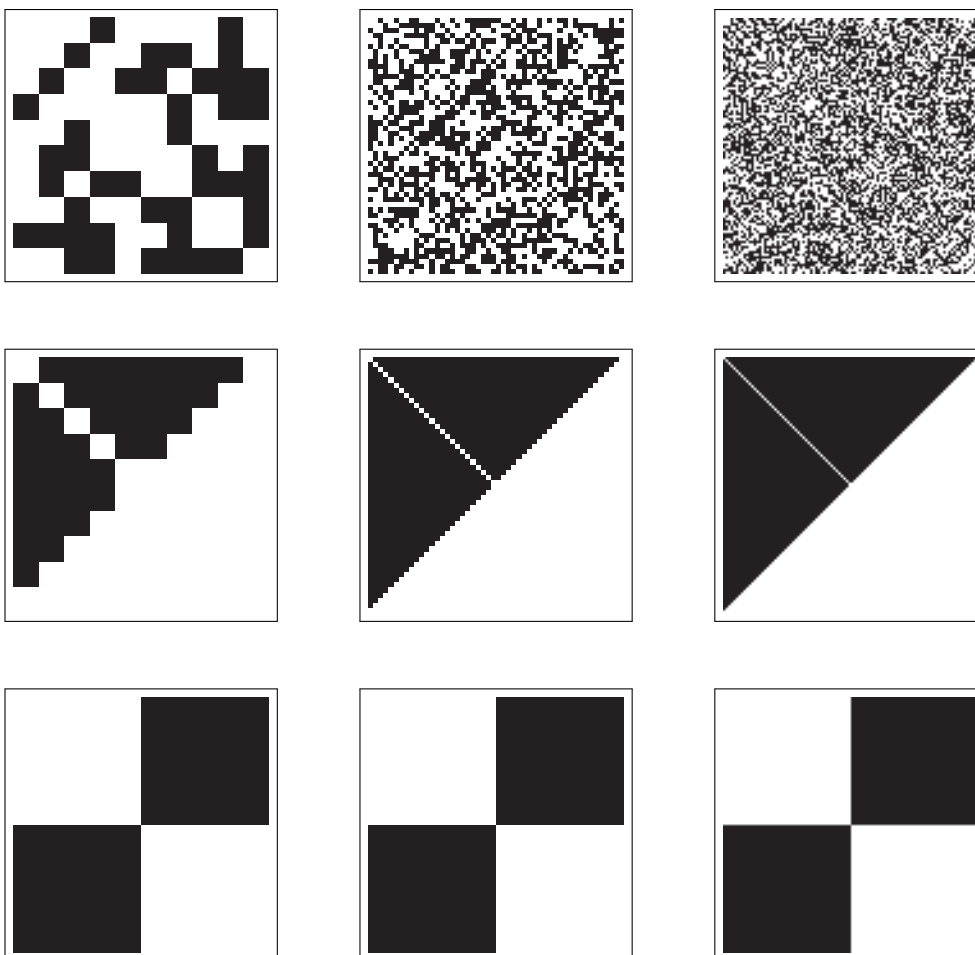


Tabela 50: Slike različnih tipov grafov na 10, 50 in 100 vozliščih.

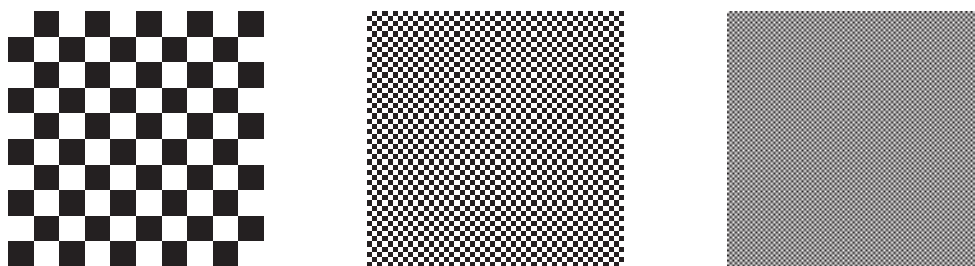


Tabela 51: Slike dvodelnih grafov z drugače razporejenimi vozlišči.

sta vozlišči i in j povezani, če $i + j \leq n$ in pa v tretji vrstici dvodelni grafi. Vidimo, da so si slike v isti vrstici kar podobne.

Seveda pa je slika odvisna od razporeditve vozlišč v matriko, tako lahko za dvodelne grafe dobimo tudi drugačne slike (slika 51). V nadaljevanju bomo natančneje definirali metriko nad grafi, ki bo poskrbela, da ima zaporedje grafov enolično določeno "sliko".

Pred tem pa si oglejmo še, kaj bi lahko bila limita takšnih zaporedij grafov.

10.3 GRAFONI

Vsako sliko matrike sosednosti si lahko predstavljamo kot funkcijo na pravokotniku $[0, 1]^2$ tako, da sliko matrike povečamo ali pomanjšamo na velikost pravokotnika in točkam v črnih kvadratih priredimo vrednost 1, točkam v belih pa 0. Zdaj lahko definiramo splošnejši objekt *grafon* na sledeči način.

Definicija 10.3 (Označen grafon). Funkciji $W : [0, 1]^2 \rightarrow [0, 1]$ pravimo *označen grafon*, če je Lebesgue-ovo merljiva in simetrična (za vsaka u in v velja $W(u, v) = W(v, u)$).

Opomba 10.4. Za definicijo Lebesgue-ovo merljivih funkcij glej [5] in [6].

Poseben primer grafonov so vsi končni grafi, ki jih kot grafone predstavimo po zgoraj opisanem postopku. Ostale grafone si lahko predstavljamo tudi kot neskončne utežene grafe z vozlišči na intervalu $[0, 1]$.

Takšnim grafonom pravimo *označeni*, ker je predstavitev grafa v tej obliki odvisna od oštevilčenja njegovih vozlišč (kot smo videli zgoraj na primeru dvodelnih grafov). Nad grafi lahko definiramo ekvivalenčno relacijo, ki pravi, da sta dva grafa ekvivalentna, če lahko iz enega dobimo drugega s preštevilčenjem vozlišč (sta izomorfna). Podobno naredimo za grafone.

Definicija 10.5 (Neoznačen grafon). Neoznačen grafon je ekvivalenčni razred označenih grafonov, ki jih dobimo iz označenega grafona z obrnljivimi preslikavami, ki ohranjajo mero:

$$[W] = \{W^\phi : (x, z) \mapsto W(\phi(x), \phi(y)) \mid \phi \text{ obrnljiva, ohranja mero}\}.$$

Opomba 10.6. Preslikava f ohranja mero, če za vsako merljivo merljivo množico A velja $\mu(f^{-1}(A)) = \mu(A)$, kjer je μ mera.

Povezavo med grafoni in naključnimi grafi nam zagotovijo sledeče definicije.

Definicija 10.7. Naj bo H poln graf utežen z utežjo $w : V(H)^2 \rightarrow [0, 1]$. Za preslikavo w velja $w(u, v) = w(v, u)$, saj obravnavamo enostavne grafe.

Z $\mathbb{G}(H)$ označimo naključen (neutežen) graf na vozliščih $V(H)$, kjer za vsako povezavo $\{u, v\}$ velja, da se pojavi z verjetnostjo $w(u, v)$.

$$\begin{aligned}\mathbb{G}(H) &= (V, E) \\ V &= V(H) \\ P(\{u, v\} \in E) &= w(u, v)\end{aligned}$$

Definicija 10.8. Naj bo W grafon in S končna podmnožica intervala $[0, 1]$. Potem s $\mathbb{H}(S, W)$ označimo utežen poln graf z vozlišči S in utežjo $w = W|_{S \times S}$ ($\forall u, v \in S : w(u, v) = W(u, v)$) in definiramo

$$\mathbb{G}(S, W) = \mathbb{G}(\mathbb{H}(S, W)) \text{ ter}$$

$\mathbb{G}(n, W) = \mathbb{G}(U, W)$, kjer je U enakomerno naključno izbrana podmnožica $[0, 1]$ moči n .

Z zgornjimi definicijami lahko iz poljubnega grafona W dobimo naključne grafe različnih velikosti in tudi zaporedje naključnih grafov po naraščajoči velikosti. Na grafone lahko razširimo tudi definicijo gostote homomorfizmov. To si oglejmo na primeru.

Zgled 10.9. Vzemimo grafon W

$$t(\bullet\bullet, W) = \int_{[0,1]^2} W(x, y) dx dy.$$

Grafone si lahko predstavljamo kot utežene polne neskončne grafe na vozliščih $[0, 1]$ in zato gostoto povezav v takšnem grafu dobimo z integriranjem uteži (verjetnosti) po vseh povezavah, t.j. $[0, 1]^2$.

Podobno

$$t(\bullet\bullet\bullet, W) = \int_{[0,1]^4} W(x_1, x_2)W(x_2, x_3)W(x_3, x_4)W(x_4, x_1) dx_1 dx_2 dx_3 dx_4.$$

Tu integriramo po vseh možnih 4-ciklih.

Na podoben način lahko zapišemo $t(H, W)$ za vse grafe H . Razlog, da utežem na grafonu pravimo tudi verjetnosti, se nahaja v definiciji 10.8.

10.4 PROSTOR GRAFONOV

Nad prostorom grafonov definiramo metriko, ki bo služila merjenju podobnosti dveh grafonov (in tudi grafov).

Definicija 10.10 (Cut distance). Razdaljo δ_{\square} med dvema grafonoma W in U definiramo kot

$$\delta_{\square}(W, U) = \inf_{\substack{\phi: [0,1] \rightarrow [0,1] \\ \text{obrnjljiva,} \\ \text{ohranja mero}}} \sup_{\substack{S, T \subseteq [0,1] \\ \text{merljivi}}} \left| \int_{S \times T} (W(x, y) - U^{\phi}(x, y)) \, dx dy \right|.$$

Ta metrika meri največjo razliko med dvema grafonoma W in U po vseh merljivih škatlah pravokotnika $[0, 1]^2$, za to poskrbi

$$\sup_{\substack{S, T \subseteq [0,1] \\ \text{merljivi}}} \left| \int_{S \times T} (W(x, y) - U^{\phi}(x, y)) \, dx dy \right|.$$

Takšen izraz je še vedno odvisen od razporeditve vozlišč na intervalu $[0, 1]$, zato naredimo infimum po vseh možnih obrnjljivih preslikavah, ki ohranjajo mero.

Izrek 10.11. Naj bo W grafon in naj bo za vsak $n \in \mathbb{N}$ R_n naključen graf $\mathbb{G}(n, W)$. Za zaporedje grafov $(R_n)_n$ velja, da skoraj gotovo konvergira proti grafonu W glede na metriko δ_{\square} .

Opomba 10.12. Skoraj gotovo pomeni, da za zaporedje $(R_n)_n$ velja

$$P \left(\lim_{n \rightarrow \infty} \delta_{\square}(R_n, W) = 0 \right) = 1.$$

Posledica 10.13. Za vsak grafon W obstaja neko zaporedje grafov G_n , ki konvergira proti grafonu W glede na metriko δ_{\square} .

Izrek 10.14. Vsako zaporedje grafov $(G_n)_n$, za katero velja

$$\forall H : (t(H, G_n))_n \text{ je konvergentno zaporedje števil,}$$

je konvergentno glede na metriko δ_{\square} in konvergira proti nekemu grafonu.

Dokaz izreka se nahaja v [1]. Iz posledice in izreka vidimo, da so grafoni dobri kandidati za objekte, ki naj bi predstavljali limite grafov.

Funkcija δ_{\square} je v resnici le psevdometrika, saj obstajajo različni grafoni, za katere velja, da je razdalja med njimi glede na δ_{\square} enaka 0. Takšni grafoni so si po lastnostih zelo podobni. Grafonoma W in U za katera velja $\delta_{\square}(W, U) = 0$ pravimo *šibko izomorfna*, saj za njiju velja $t(H, W) = t(H, U)$ za poljuben graf H .

Definicija 10.15. Nad prostorom grafonov definiramo ekvivalenčno relacijo $W \sim U \iff \delta_{\square}(W, U) = 0$.

Z \mathcal{G} označimo kvocientni prostor prostora grafonov glede na relacijo \sim .

Izrek 10.16. $(\mathcal{G}, \delta_{\square})$ je metrični prostor.

Izrek 10.17 (Lovász – Szegedy). Prostor $(\mathcal{G}, \delta_{\square})$ je kompakten.

Posledica 10.18. Ker je $(\mathcal{G}, \delta_{\square})$ metrični prostor je kompakten natanko tedaj, ko je poln in popolnoma omejen. Prostor končnih grafov je v tem prostoru gosta podmnožica.

10.5 PREPROST PRIMER

V uvodu smo zapisali vprašanje “Koliko 4-ciklov ima graf, ki ima več kot polovico vseh možnih povezav?” Odgovor na podobno vprašanje nam da Erdőssev teorem [2], ki pravi, da je gostota homomorfizmov iz 4-cikla v graf G večja kot četrta potenca gostote povezav v grafu G .

$$t(\square, G) \geq t(\bullet\bullet, G)^4.$$

In posebej, če $t(\bullet\bullet, G) \geq \frac{1}{2}$, potem $t(\square, G) \geq \frac{1}{16}$.

Na ta problem lahko gledamo tudi kot na optimizacijski problem, ki zahteva, da minimiziramo $t(\square, G)$ med vsemi grafi G , ki zadoščajo pogoju $t(\bullet\bullet, G) \geq \frac{1}{2}$.

Izkaže se, da noben končen graf ne doseže minimuma $\frac{1}{16}$. Hitro pa lahko vidimo, da minimum doseže konstanten grafon $W \equiv \frac{1}{2}$, saj

$$\begin{aligned} t(\bullet\bullet, W) &= \int_{[0,1]^2} W(x, y) dx dy \\ &= \int_{[0,1]^2} \frac{1}{2} dx dy \\ &= \frac{1}{2}. \end{aligned}$$

in

$$\begin{aligned} t(\square, W) &= \int_{[0,1]^4} W(x_1, x_2) W(x_2, x_3) W(x_3, x_4) W(x_4, x_1) dx_1 dx_2 dx_3 dx_4 \\ &= \int_{[0,1]^4} \left(\frac{1}{2}\right)^4 dx_1 dx_2 dx_3 dx_4 \\ &= \frac{1}{16}. \end{aligned}$$

Zdaj, ko imamo grafon $W \equiv \frac{1}{2}$, ki doseže minimum, pa lahko z njim poiščemo grafe, ki so blizu minimuma, torej aproksimiramo rešitev. Dobra izbira grafa na n vozliščih je $G(n, W)$ (glej definicijo 10.8).

10.6 ZAKLJUČEK

Ogledali smo si, kako lahko definiramo konvergenco za zaporedja grafov, in iskali naravne limitne objekte. V ta namen smo definirali splošnejše objekte – *grafone*. Nad prostorom grafonov smo nato definirali metriko in opisali, kako poljuben graf predstavimo kot grafon. S takšnimi definicijami konvergenca v metričnem prostoru grafonov ustreza konvergenci, ki smo jo opisali na začetku (preko gostote homomorfizmov).

Na začetku nismo mogli odgovoriti na vprašanje, kaj so limite zaporedij grafov. Zdaj pa imamo metrični prostor grafonov, v katerem so grafi gosta podmnožica, kar pomeni, da so limite konvergentnih zaporedij grafov ravno grafoni.

V seminarju smo zapisali osnovne definicije in omenili nekaj izrekov. Ključni izrek je, da je prostor grafonov kompakten metrični prostor. Ta izrek je ekvivalenten *Szemerédijevi lemi o regularnosti*, ki govori o strukturi velikih grafov [2, 7].

LITERATURA

- [1] L. Lovász, *Large networks and graph limits*, American Mathematical Soc., 2012.
- [2] D. Glasscock, *What is a Graphon*, Notices of the AMS vol. 62 no. 1, American Mathematical Soc., 2013.
- [3] A. Grzesik, *Flag algebras in extremal graph theory*, PhD thesis, Jagiellonian University, Krakow, 2014.
- [4] C. Lee, *Lecture 5: Graph limits and flag algebra*, dostopano 11. 4. 2016. [Dostopno na http://math.mit.edu/~cb_lee/18.318/lecture5.pdf.]
- [5] Wikipedia, *Lebesgue measure* — *Wikipedia, The Free Encyclopedia*, dostopano 29. 6. 2016. [Dostopno na https://en.wikipedia.org/wiki/Lebesgue_measure.]
- [6] Wikipedia, *Measurable function* — *Wikipedia, The Free Encyclopedia*, dostopano 29. 6. 2016. [Dostopno na https://en.wikipedia.org/wiki/Measurable_function.]
- [7] Wikipedia, *Szemerédi regularity lemma* — *Wikipedia, The Free Encyclopedia*, dostopano 15. 7. 2016. [Dostopno na https://en.wikipedia.org/wiki/Szemerédi_regularity_lemma.]

11.1 UVOD

V sledeči seminarski nalogi bomo predstavili t.i. Johnsonove grafe ter posplošene Johnsonove grafe oz. J-grafe. Johnsonovi grafi nosijo ime po Ameriškem matematiku Selmer Martin Johnsonu, ki se je med drugim ukvarjal s problemom trgovskega potnika ter uporabo linearnega programiranja za reševanje le-tega. Prav tako pa se je Johnson ukvarjal s teorijo kodov ter asociativnimi shemami, od kjer so se kasneje razvili Johnsonovi grafi. Asociativne sheme so se najprej razvile v statistiki, kmalu pa so se razširile tudi na področje kombinatorike, algebre ter teorije kodov. Pri asociativnih shemah gre za to, da set elementov razdelimo na določeno število particij, nato pa opredelimo relacijo med dvema particijama, glede na izbrano lastnost. V teoriji kodov je tako znana Hammingova shema, kjer particije oz. razrede seta predstavljajo n -elementni binarni vektorji. Za dva vektorja nato lahko rečemo da sta 1-povezana, če se razlikujeta v 1 bitu, 2-povezana, če se razlikujeta v 2 bitih in tako naprej. Pri Johnsonovi shemi pa gre za to, da poljubni set X razdelimo na vse k -elementne podmnožice tega seta, dve podmnožici pa sta v relaciji, ko njun presek vsebuje natanko $k - 1$ elementov, pri čemer je k podan parameter. Hitro lahko opazimo, da so grafi idealna vizualizacija asociativnih shem, kjer vozlišča grafov predstavljajo particije seta, povezave pa relacije med njimi. Pri posplošenih Johnsonovih grafih oz. J -grafih, pa gre, kot že samo ime pove, za razširitev osnovne definicije Johnsonovih grafov, tako da sta sedaj dve k -elementi particiji v relaciji, če njun presek vsebuje natanko i elementov, kjer sta k in i podana parametra. Tako celotno družino grafov krepko razširimo in med drugim tako tudi definiramo t.i. Kneserjeve grafe, kjer je vhodni parameter k poljuben, i pa je enak

0; torej particiji sta v relaciji, natanko tedaj ko sta med seboj disjunktni.

V prvem delu najprej predstavimo zgoraj omenjene J -grafe in njihove osnovne lastnosti, saj predstavljajo lastnosti J -grafov tudi osnove za lastnosti Johnsonovih grafov, ker spadajo v isto družino. Predstavimo tudi nekaj pomembnih lastnosti J -grafov, ki se seveda prenesejo tudi na Johnsonove grafe, med drugim Izrek o Izomorfizmu ter Izrek o vozliščni tranzitivnosti. V drugem delu pa se bolj podrobno posvetimo Johnsonovim grafov, izpeljemo osnovne lastnosti Johnsonovih grafov, nato pa se osredotočimo na dve lastnosti Johnsonovih grafov, ki ne sledita iz J -grafov - premer in Hamiltonskost. Tako izpeljemo premer Johnsonovega grafa v odvisnosti od začetnih parametrov k, i ter pokažemo da je poljubni Johnsonov graf (ki zadostuje osnovnim omejitvam, opisanim v poglavju 1.3) tudi Hamiltonski.

11.2 J -GRAFI

Družina simetričnih grafov, ki nosi ime J -grafi oz. posplošeni Johnsonovi grafi, ima oznako $J(n, k, i)$ in je določena na naslednji način. Naj bo Ω končna množica moči n , potem vozlišča grafa $J(n, k, i)$ predstavljajo vse podmnožice množice Ω , moči k . Dve vozlišči pa sta povezani, če je moč njunega preseka enaka i .

11.2.1 Lastnosti J -grafov

Iz same definicije J -grafov sledi nekaj osnovnih lastnosti, s kratkim premislekom pa lahko dodatno izpeljemo še nekaj omejitev in zahtev, katerim morajo J -grafi zadostovati.

Lastnosti J -grafov:

1. $|V(J(n, k, i))| = \binom{n}{k}$,
2. Graf $J(n, k, i)$ je r -regularen, za $r = \binom{k}{i} \binom{n-k}{k-i}$,
3. $|E(J(n, k, i))| = \frac{1}{2} \binom{n}{k} \binom{k}{i} \binom{n-k}{k-i}$,
4. $n \geq 2k + i$.

Točke 1, 2 in 3 so posledica same definicije J -grafov. Namreč vseh k -elementnih podmnožic množice $|\Omega| = n$ je natanko $\binom{n}{k}$. Po definiciji sta dve vozlišči povezani, če je njun presek moči i . Izračun števila sosedov za posamezno vozlišče tudi sledi iz same definicije. Za posamezno vozlišče tako izberemo i elementov, ki bodo vsebovani v preseku (izmed k elementov, ki določajo vozlišče), ter $k - i$ elementov,

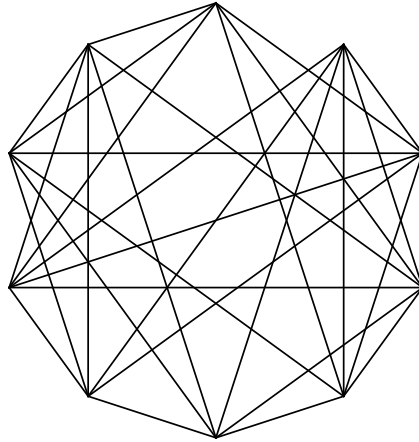


Tabela 52: $J(5, 2, 1)$, ki predstavlja komplement Petersenovega grafa oz. Kneserjevega $K(5, 2)$ grafa.

ki se v preseku ne bodo nahajali, izmed preostalih $n - k$ elementov v množici Ω .

Točka 4 pa sledi iz želje, da bo naš graf povezan, saj če zahteva iz toče 4 ni zadoščena, bo imel naš J -graf natanko 0 povezav. Z razmislekom vidimo, da če imamo 2 vozlišči, bo v njunem preseku (po definiciji) i elementov. Dodatno pa se mora v vsakem vozlišču nahajati še $k - i$ elementov, ki vozliščema niso skupni. Torej potrebujemo množico Ω moči minimalno $i + 2(k - i) = 2k - i$.

11.2.2 Izomorfizem J -grafov

Pomembno lastnost J -grafov opiše naslednji izrek o izomorfnosti.

Izrek 11.1. Za poljuben J -graf za katerega velja $n \geq 2k + i$ velja naslednja trditev:

$$J(n, k, i) \cong J(n, n - k, n - 2k + i). \quad (18)$$

Dokaz. Označimo dva izbrana grafa na sledeč način:

$$G_1 = J(n, k, i) \quad (19)$$

in

$$G_2 = J(n, n - k, n - 2k + i). \quad (20)$$

Želimo poiskati izomorfizem ϕ , ki bo slikal $G_1 \rightarrow G_2$. Da to drži, je potrebno pokazati, da je ϕ bijekciji ter da ohranja sosednost.

Glede na zahteve izomorfizma je tako primeren kandidat za ϕ kar funkcija, ki množici priredi njen komplement. Preverimo torej da $\phi : A \rightarrow A^c$ izpolnjuje pogoje za izomorfizem.

Grafa G_1 in G_2 imata po definiciji binomskih izrekov enako število vozlišč, namreč:

$$\binom{n}{k} = \binom{n}{n-k}. \quad (21)$$

Torej je ϕ bijekcija, saj je po definiciji komplement injektivna funkcija in je torej tudi bijektivna.

Preveriti moramo torej da ϕ ohranja sosednost. V J -grafih je sosednost odvisna le od moči preseka dveh vozlišč, torej preverimo kaj se zgodi preseku vozlišč $v_1, v_2 \in V(J)$, ko nad grafom uporabimo ϕ :

$$|v_1^c \cap v_2^c| = |(v_1 \cup v_2)^c| = n - |v_1 \cup v_2| = n - (|v_1| + |v_2| - |v_1 \cap v_2|) = n - 2k + i. \quad (22)$$

□

11.2.3 Vozliščna tranzitivnost J -grafovo

Izrek 11.2. Če velja $n \geq k \geq i$ potem grupa avtomorfizmov grafa $J(n, k, i)$ vsebuje grupo izomorfno simetrični grupi S_n . $J(n, k, i)$ je posledično vozliščno tranzitiven.

Dokaz. Dokazati želimo, da za poljubno izbrani vozlišči $u, v \in J(n, k, i)$ obstaja izomorfizem, ki bo u slikal v v .

Izberimo $\phi \in S_n$ in definirajmo njegovo delovanje nad vozlišči kot:

$$\phi(V), V \in \Omega = \{\phi(v) : v \in V\}. \quad (23)$$

Ker $\phi \in S_n$ je ϕ injektivna in zato tudi bijektivna nad $J(n, k, i)$. Z razmislekom hitro vidimo da ϕ tudi ohranja sosednost. Če si izberemo vozlišči $v, u \in V(J(n, k, i))$, je jasno, da se bodo elementi v njunem preseku ostali tud v preseku $\phi(v), \phi(u)$, hkrati pa bodo elementi, ki so bili vozliščema tuji, zaradi bijekcije ostali tuji. Torej iz $v \sim u$ res sledi $\phi(v) \sim \phi(u)$ in ϕ res ohranja sosednost.

Sedaj lahko pokažemo, da za poljubni vozlišči u, v , kjer velja, $u = \{u_1, u_2, \dots, u_k\}$, $v = \{v_1, v_2, \dots, v_k\}$, $u, v \in V(J(n, k, i))$ obstaja preslikava ϕ tako da $\phi(v) = u$. ϕ izberemo tako, da velja $\phi(v_1) = u_1, \phi(v_2) = u_2, \dots, \phi(v_k) = u_k$. Presostane name še množica elementov $\Omega \setminus v$, katero pa lahko poljubno slikamo na množico $\Omega \setminus u$, saj sta enakih moči. S tem smo naš izrek dokazali. □

11.3 JOHNSONOVİ GRAFI

Johnsonovi grafi predstavljajo eno od družin grafov v J -grafih. Gre za J -grafe oblike:

$$J(n, k, k - 1), \quad (24)$$

ki jih označimo z $J(n, k)$. Tukaj velja omeniti še eno od družin grafov vsebovanih v J -grafih - Kneserjeve grafe, oblike $J(n, k, 0)$, ki jih označimo s $K(n, k)$.

11.3.1 Lastnosti Johnsonovih grafov

Johnsonovi grafi predstavljajo družino grafov v J -grafih, zato so tudi njihove lastnosti podobne lastnostim J -grafov, z nekaj izjemami, ki sledijo iz povezave $i = k - 1$ ki velja za Johnsonove grafe. Lastnosti Johnsonovih grafov:

1. $|V(J(n, k))| = \binom{n}{k}$,
2. Graf $J(n, k)$ je r -regularen, za $r = k(n - k)$,
3. $|E(J(n, k, i))| = \frac{k}{2}k(n - k)\binom{n}{k}$,
4. $n \geq k + 1$,
5. $J(n, k) \cong J(n, n - k)$.

Prve 4 točke sledijo neposredno iz lastnosti J -grafov, pri čemer namesto i v enačbe vstavimo $k - 1$.

Točka 5 sledi iz izreka o izomorfizmu J -grafov. Če Johnsonov graf zapišemo na sledeč način:

$$J(n, k) = J(n, k, k - 1) \quad (25)$$

in v izrek o izomorfizmu grafov vstavimo $i = k - 1$ dobimo:

$$J(n, k, k - 1) \cong J(n, n - k, n - 2k + k - 1) = J(n, n - k, n - k - 1) = J(n, n - k). \quad (26)$$

Vidimo, da izrek o izomorfizmu drži tudi za Johnsonove grafe, saj je $J(n, n - k, n - k - 1)$, ki je rezultat izomorfizma ponovno Johnsonov graf.

11.3.2 Premer Johnsonovih grafov

Iz same definicije Johnsonovih grafov, lahko hitro izpelejmo tudi njihov premer, ki v praksi pride prav pri uporabi Johnsonovih grafov v računalništvu, kjer se uporabljajo med drugim tudi pri iskanju popolnih kodov, kot je to opisano v [8], kjer se avtor ukvarjajo z obstojem popolnih kodov v razdaljno tranzitivnih grafih.

Izrek 11.3. *Premer grafa $J(n, k)$ je enak $\min\{n, n - k\}$.*

Dokaz. Izberimo si poljubno vozlišče $v \in V(J(n, k))$, $v = \{v_1, v_2, \dots, v_k\}$ in ga fiksirajmo. To lahko storimo, saj je Johnsonov graf vozliščno tranzitiven. Sedaj želimo poiskati poljubno vozlišče $u \in V(J(n, k))$, tako da bo razdalja med u in v maksimalna.

Vozlišče v je fiksirano s k elementi - $\{v_1, v_2, \dots, v_k\}$. Vozlišče u pa želimo izbrati tako, da bo imelo z v kar največ tujih elementov, ki jih izbiramo iz množice $\Omega \setminus k$.

Tujih elementov je $n - k$, kar pomeni, da se lahko zgodi, da imamo premalo elementov, da bi lahko izbrali u , ki bi se od v razlikoval po vseh elementih. Posledica tega je, da imamo v presku vozlišč $\max\{0, 2k - n\}$ skupnih elementov.

Tukaj lahko ločimo dva primera, če se zgodi, da imamo v Ω dovolj elementov, da se u in v razlikujeta v vseh komponentah, potem smo našli vozlišči na maksimalni oddaljenosti. Po definiciji Johnsonovih grafov vemo, da se dve sosednji vozlišči razlikujeta v natanko eni komponenti, torej bosta vozlišči u in v na razdalji k , saj nam vsak premik po grafu omogoča, da popravimo le eno komponento, torej bomo potrebovali k korakov, da pridemo iz v do u .

Kaj pa v primeru da v Ω ni mogoče izbrati dveh vozlišč, ki bi se razlikovali v k komponentah? Kot smo napisali zgoraj, to pomeni, da je v njunem preseku $2k - n$ skupnih elementov. V tem primeru pa namesto grafa $J(n, k)$ gledamo izomorfen graf $J(n, n - k)$. Ker je izomorfizem Johnsonovih grafov definiran kot komplement množice, ki definira vozlišče (Izrek 1.2), se bosta u in v sedaj razlikovali v vseh $n - k$ komponentah. Torej res drži, da je premer Johnsonovega grafa definiran kot $\min\{k, n - k\}$. \square

11.3.3 Hamiltonov cikel v Johnsonovih grafih

Izrek 11.4. *$J(n, k)$ ima Hamiltonski cikel za $n \geq 3$.*

Dokaz. Zgornjo trditev bomo pokazali z indukcijo nad številom vozlišč v Johnsonovem grafu $J(n, k)$, z dokazom prirejenim po članku "Cutsets and Hamiltonian cycles of Johnson graphs"[6].

- $n = 3$: Za $n = 3$ obstajata dva Johnsonova grafa: $J(3, 1)$ ter $J(3, 2)$, ki pa sta si po Izreku 1.1 izomorfna, torej je dovolj, da najdemo Hamiltonovo pot v $J(3, 1)$. Če definiramo

$$\Omega = \{1, 2, 3\}, \quad (27)$$

potem je $J(3, 1)$ kar cikel na treh vozliščih in imamo Hamiltonsko pot

$$\{1\} \rightarrow \{2\} \rightarrow \{3\} \rightarrow \{1\}. \quad (28)$$

- $n = n + 1$: Definirajmo 2 Johnsonova grafa, na sledeč način:

$$\begin{aligned} C_1 &= J(n-1, k-1), \\ C_2 &= J(n-1, k). \end{aligned} \quad (29)$$

Po indukcijski predpostavki sta C_1 in C_2 Hamiltonska oz. vsebujeta Hamiltonov cikel. Sedaj želimo pokazati, da lahko na podlagi obstoja Hamiltonovega cikla v C_1 ter C_2 dokažemo obstoj Hamiltonovega cikla v $J(n, k)$. Za lažjo predstavo definirajmo

$$\Omega = \{a_2, a_3, a_4, \dots, a_n\}. \quad (30)$$

Sedaj pa v vsako vozlišče v grafu C_1 dodajmo element a_1 ter poglejmo C_1 in C_2 v kontekstu grafa $J(n, k)$.

Graf C_1 je sedaj izomorfen ciklu skozi $\binom{n-1}{k-1}$ vozlišč v $J(n, k)$, ki vsebujejo element a_1 . Graf C_2 pa je izomorfen ciklu skozi $\binom{n-1}{k}$ vozlišč v $J(n, k)$, ki elementa a_1 ne vsebujejo. Torej sta cikla C_1 ter C_2 disjunktna, hkrati pa skupaj pokrijeta

$$\binom{n-1}{k-1} + \binom{n-1}{k} = \binom{n}{k} \quad (31)$$

vozlišč, kar pa so ravno vse vozlišča v $J(n, k)$. Če nam sedaj uspe dokazati obstoj vozlišč a, b, c, d , za katera velja:

- $a, b \in C_1; c, d \in C_2$,
- $a \sim b, a \sim c, d \sim c, d \sim b$.

potem smo izrek uspešno dokazali, saj lahko nato cikla C_1 in C_2 preko vozlišč a, b, c, d združimo v en cikel, ki bo šel skozi vsa vozlišča v $J(n, k)$.

S premislekom hitro najdemo iskana vozlišča definirana na naslednji način:

- $a = \{a_1, a_2, a_3, \dots, a_k\}$,
- $b = \{a_1, a_3, a_4, \dots, a_k\}$,
- $c = \{a_2, a_3, a_4, \dots, a_k + 1\}$,
- $d = \{a_3, a_4, a_5, \dots, a_k + 2\}$.

□

11.4 ZAKLJUČEK

V seminarski nalogi predstavimo družino posplošenih Johnsonovih oz. J -grafov, njihove osnovne lastnosti, nato pa preidemo na Johnsonove grafe, ki so ena od pod družin J -grafov. Lastnosti ter dokazi so povzeti po [1, 2, 3, 4]. Pri Johnsonovih grafih izpeljemo še njihov premer ter pokažemo da so vsi Johnsonovi grafi tudi Hamiltonski. Dokaz Hamiltonskosti je prirejen po [6, 7].

LITERATURA

- [1] Godsil, C. and Royle, G.F., *Algebraic Graph Theory*, Graduate Texts in Mathematics, 2001.
- [2] Biggs, N., *Algebraic Graph Theory*, Cambridge Mathematical Library, 1993.
- [3] Andries E. Brouwer, Willem H. Haemers , *Spectra of graphs*, 2011.
- [4] Sabina Pintar, *Johnsonovi in kneserjevi grafi*, Fakulteta za matematiko in fiziko, Univerza v Ljubljani, 2014.
- [5] Elena Konstantinova, *Lecture notes on some problems on Cayley graphs*, Zbirka Izbrana poglavja iz matematike, št. 10, 2012.
- [6] Wantao Ning, Qiuli Li, Heping Zhang, *Cutsets and Hamiltonian cycles of Johnson graphs*, School of Mathematics and Statistics, Lanzhou University.
- [7] Brian Alspach, *Johnson graphs are Hamilton-connected*, School of Mathematical and Physical Sciences University of Newcastle Callaghan, 2011.
- [8] W. J. Martin, *Completely regular codes – a viewpoint and some problems*, Department of Mathematical Sciences and Department of Computer Science Worcester Polytechnic Institute, Worcester, Massachusetts