Svojstva slučajnog grafa

Račić, Lana

SVEUČILIŠTE U ZAGREBU
PRIRODOSLOVNO–MATEMATIČKI FAKULTET
MATEMATIČKI ODSJEK

Lana Račić

SVOJSTVA SLUČAJNOG GRAFA

Diplomski rad

Voditelj rada:
dr. sc. Nina Kamčev,
dr. sc. Vanja Wagner

Zagreb, rujan 2023.

Ovaj diplomski rad obranjen je dana ____________________ pred ispitnim povjerenstvom u sastavu:

1. ____________________________, predsjednik
2. ____________________________, član
3. ____________________________, član

Povjerenstvo je rad ocijenilo ocjenom _______________.

Potpisi članova povjerenstva:

1. ____________________________
2. ____________________________
3. ____________________________

Sadržaj

Uvod

Slučajni grafovi su područje matematike koje kombinira elemente diskretne matematike i vjerojatnosti. Na skupu svih grafova sa $n \in \mathbb{N}$ vrhova definira se vjerojatnosni prostor, a proučavaju se svojstva tipičnog grafa iz tog prostora kada n teži u beskonačnost. Najjednostavniji teorijski model je $\mathbb{G}_{n,p}$, u kojemu graf ima n vrhova, a svaki brid pojavljuje se nezavisno od ostalih sa vjerojatnošću $p \in \langle 0, 1 \rangle$.

Primjeri slučajnih grafova mogu se naći u biologiji, sociologiji i kemiji te posebno u fizici, gdje su se proučavali i kada je teorija slučajnih grafova u matematici tek započinjala. Grafovi sami po sebi uvijek su relevantni za računalnu znanost. Poznavanje teorije slučajnih grafova može ubrzati algoritme u čijoj podlozi je struktura grafa.

U matematici, slučajni grafovi pojavljuju se kao prirodno definirani objekti i kao korisni alati. Na primjer u turniru postojanje brida (u, v) znači da je u "nadjačao" v, međutim prirodno se može uvesti slučajnost ako stavimo da se (u, v) događa s vjerojatnošću p, a (v, u) inače. Ta konstrukcija zove se generalizirani turnir [10]. Pokazuje se da je za $n \leq \lfloor 2^{k/2} \rfloor$ slučajan graf $\mathbb{G}_{n,p}$ na n vrhova s pozitivnom vjerojatnošću takav da niti on sam niti njegov komplement ne sadrže k-kliku. To znači da je Ramseyev broj $R(k, k) > \lfloor 2^{k/2} \rfloor$. Vjerojatnosna metoda također omogućuje poboljšavanje nekih ograda, kao što se događa u Jansonovim nejednakostima.

Uzlet teorije slučajnih grafova krenuo je sa radovima Paulu Erdősa i Alfrédu Rényia 1960-ih. Jedan od njih je [4]. Od tada do danas interes za ovo područje nije prestao. Na primjer web stranica Dimensions daje 44000 publikacija sa riječima "random graph" u naslovu ili abstraktu, od kojih je otprilike tri tisuće objavljeno prošle godine[1].

Objavljene su i brojne knjige, od značajnog djela Béle Bollobása [2] do prošle godine objavljenoga pristupačnog udžbenika autora Alana Friezea and Michała Karońskog [5]. Od nedavnih napredaka, Jinyoung Park i Huy Tuan Pham 2022. godine objavili su prijedlog dokaza Khan-Kalaijeve slutnje [8]. Ova snažna i općenita tvrdnja, za koju ni Khan i Kalai nisu vjerovali da će se pokazati istinitom, poopćuje važne i teške teoreme o prijelaznim funkcijama slučajnih grafova [3].

[1] Ove brojke su tu da opišu količinu interesa za područje, pretraživački softver stranice Dimensions baziran je na umjetnoj inteligenciji i proces kojim su radovi uvršteni u rezultate pretrage nije transparentan. Ipak vizualni pregled rezultata i doživljena ozbiljnost stranice govore da se rezultati mogu koristiti za ove svrhe.

U ovome radu biti će riječi o samim temeljima teorije slučajnih grafova. Prvo poglavlje bavi se prirodnim pitanjem kada se neki fiksni graf H pojavljuje u slučajnom grafu $\mathbb{G}_{n,p}$. Otkriva se da se H pojavljuje naglo, i da vrijednost p nakon koje se H pojavljuje ovisi o gustoći bridova u H. U drugom poglavlju bavimo se činjenicom koja je otkrivena u glasovitom radu Erdősa i Rényija [4] te se oni o njoj izražavaju ovako, parafrazirano:

> Situacija može biti sažeta na sljedeći način: najveća komponenta $\mathbb{G}_{n,p}$ je reda veličine $\ln n$ kada $p \sim \frac{c}{n},\ c < 1$, reda veličine $n^{2/3}$ kada $p \sim \frac{1}{n}$ i reda veličine n kada $p \sim \frac{c}{n},\ c > 1$. Ovaj dvostruki "skok" veličine najveće komponente kada p prelazi $1/n$ je jedna od najupečatljivijih činjenica u vezi slučajnih grafova.

U trećem i posljednjem poglavlju bavimo se savršenim sparivanjima. Dokazuje se teorem u kojem se pokazuje vrijednost p u kojoj se $\mathbb{G}_{n,p}$ mijenja od grafa u kojem tipično ne postoji savršeno sparivanje do onoga gdje takvo sparivanje tipično postoji. Objašnjava se kako je ovo svojstvo povezano s postojanjem Hamiltonovog ciklusa, povezanosti i izoliranih vrhova.

Poglavlje 1

Slučajni grafovi

U ovome je poglavlju cilj upoznavanje s objektima i metodama proučavanja slučajnih grafova. Neke definicije i tvrdnje su standardne, no za razinu diplomskog rada prekompleksne za korištenje bez iskaza, takvi su na primjer definicija stabla, Cayleyeva formula, Čebiševljeva i Markovljeva nejednakost itd.

Prvo poglavlje iznosi sve važne definicije, počevši od definicija vezanih općenito za grafove. Tamo su navedeni i teoremi iz teorije grafova koje koristimo bez dokaza, kao i povijesni kontekst uz pojedine definicije i tvrdnje. Tu će biti naveden i teorem koji povezuje glavne modele slučajnih grafova $\mathbb{G}_{n,p}$ i $\mathbb{G}_{n,m}$.

U drugom potpoglavlju opisani su alati koji se često koriste u radu sa slučajnim grafovima. Tu se nalazi nekoliko riječi o asimptotskoj notaciji koja će se gotovo konstantno pojavljivati u ostatku rada, neke jednakosti i nejednakosti tehničke prirode, metode prvog i drugog momenta. Također se posebno promatra slučaj kada je slučajna varijabla X suma indikatorskih varijabli, jer će se često promatrati na primjer X je broj grafova nekog tipa u slučajnom grafu.

1.1 Osnovni pojmovi i strukture

Glavni objekt proučavanja su grafovi, konkretno u opsegu ovog rada neusmjereni grafovi.

Definicija 1.1.1. Graf $G = (V, E)$ *je uređeni par skupa vrhova* V *i skupa bridova* $E \subseteq \{\{i, j\};\ i, j \in V,\ i \neq j\}$.

Kažemo da brid $\{i, j\}$ *povezuje* vrhove i i j, odnosno da su vrhovi i i j *incidentni* s tim bridom. Za bridove kažemo da su *susjedni* ako dijele zajednički vrh, a za vrhove kažemo da su *susjedni* ako postoji brid koji povezuje ta dva vrha. Vrh $v \in V$ zovemo *izoliranim* ako nije incidentan niti sa jednim bridom.

Definicija 1.1.2. Put *duljine* $k \in \mathbb{N}$ *u grafu* $G = (V, E)$ *je niz vrhova* $(v_0, v_1, \ldots, v_k)$ *takav da* $\{v_{i-1}, v_i\} \in E$ *za svaki* $i = \{1, 2, \ldots, k\}$ *i* $v_i \neq v_j$ *kada* $i \neq j$*, osim eventualno* $v_0 = v_k$*. Put koji ima isti početni i završni vrh naziva se* ciklus.

Definicija 1.1.3. *Graf* $G = (V, E)$ *je* povezan *ako za svaka dva vrha* $v_0, v_1 \in V$ *postoji put kojem je* v_0 *početni vrh, a* v_1 *završni.*

Definicija 1.1.4. Stablo *reda* $k \in \mathbb{N}$ *je povezani graf u kojem ne postoji ciklus.* .Uniciklični graf *je graf u kojem postoji točno jedan ciklus.*

Definicija 1.1.5. Kompleksnost *povezanoga grafa* $G = (V, E)$ *sa* v *vrhova i* e *bridova je* $e - v + 1$.

Definicija 1.1.6. Sparivanje *u grafu* $G = (V, E)$ *je podskup bridova* $M \subseteq E$ *takav da niti jedan par različitih bridova* $e, f \in M$ *nije susjedan. Sparivanje je* savršeno *ako za svaki vrh* $v \in V$ *postoji brid* $e \in M$ *koji je incidentan s* v*.*

Sljedeće vrlo poznate tvrdnje koristiti će se u ovom radu.

Propozicija 1.1.7. *Stablo reda* k *ima točno* $k - 1$ *brid. Unicikličan povezan graf na* k *vrhova ima točno* k *bridova.*

Propozicija 1.1.8 (Cayleyeva formula)**.** *Na fiksnih* k *vrhova postoji* k^{k-2} *različitih stabala reda* k*.*

U prvom poglavlju baviti ćemo se postojanjem fiksnog podgrafa H u slučajnom grafu. Definirajmo podgraf.

Definicija 1.1.9. Podgraf *grafa* $G = (V, E)$ *je je bilo koji graf* $H = (V', E')$ *takav da* $V' \subseteq V$ *i* $E' \subseteq E$*. Oznaka je* $H \sqsubseteq G$*. Graf* $H = (V', E')$ *je* inducirani podgraf *grafa* $G = (V, E)$ *ako je* $V \subseteq V$*, a* $E' = E \cap (V' \times V')$*, također kažemo da je* H *podgraf induciran skupom vrhova* V'*.*

U proučavanju grafova rijetko je kada bitno na kojem skupu vrhova V je graf definiran, možemo poistovjetiti grafove između čijih vrhova postoji bijekcija koja zadržava bridove. Za grafove $G_1 = (V_1, E_1)$ i $G_2 = (V_2, E_2)$ kažemo da su *istovjetni* ako postoji bijekcija $f : V_1 \to V_2$ takva da:

$$\{v_1, v_2\} \in E_1 \Leftrightarrow \{f(v_1), f(v_2)\} \in E_2, \qquad \forall v_1, v_1 \in V'.$$

Dakle, u širem smislu, možemo reći i da je $H = (V', E')$ podgraf $G = (V, E)$ ako postoji bijekcija $f : V' \to V$ takva da

$$\{v_1, v_2\} \in E' \Leftrightarrow \{f(v_1), f(v_2)\} \in E, \qquad \forall v_1, v_1 \in V'.$$

U tom slučaju H je istovjetan s nekim podgrafom od G. Za $H \sqsubseteq G$ kažemo da G *sadrži* H.

Definicija 1.1.10. Komponenta *grafa* $G = (V, E)$ *je njegov inducirani podgraf* $H = (V', E')$ *koji je povezan i takav da ne postoje vrhovi* $i \in V'$ *i* $j \in V \setminus V'$ *takvi da* $\{i, j\} \in E$.

Graf može imati neke simetrije koje zovemo automorfizmi. To se događa kada je graf na nekoliko načina, pod raznim permutacijama, istovjetan sam sebi.

Definicija 1.1.11. *Bijekcija* $a : V \to V$ *je* automorfizam *grafa* $G = (V, E)$ *ako:*

$$\{v_1, v_2\} \in E' \Leftrightarrow \{a(v_1), a(v_2)\} \in E, \qquad \forall v_1, v_1 \in V'.$$

Proučavaju se svojstva grafova, na primjer: graf je povezan, graf sadrži trokut, graf ima put duljine n, graf je stablo, između svaka dva vrha postoji put duljine manje ili jednake dva itd. Formalno, *svojstvo* grafa je neki skup grafova, no u tekstu se ipak kaže da je svojstvo "povezanost" umjesto "skup svih grafova koji su povezani". Tako se događaj da slučajni graf G ima svojstvo $\mathcal{A}$ označava

$$G \in \mathcal{A} \text{ ili } G \models \mathcal{A}.$$

Definicija 1.1.12. *Svojstvo grafova* $\mathcal{A}$ *je* monotono rastuće *ako iz* $G = (V, E') \in \mathcal{A}$ *slijedi da i za bilo koji* $E \supset E'$ *i graf* $H = (V, E)$ *vrijedi* $H \in \mathcal{A}$. *Svojstvo grafova* $\mathcal{A}$ *je* monotono padajuće *ako* $G = (V, E) \notin \mathcal{A}$ *povlači da niti za jedan* $E' \subset E$, $H = (V, E') \notin \mathcal{A}$.

Primjeri monotono rastućih svojstava su: povezanost, sadrži trokut, postoji put duljine n. Komplementi, odnosno negacije monotono rastućih svojstva su monotono padajuća svojstva. Postoje i svojstva koja nisu niti monotono padajuća niti monotono rastuća, na primjer svojstvo grafovi koji su stabla.

Dodavanjem vjerojatnosti u teoriju grafova razvija se pojam slučajnog grafa, na dva vrlo bliska načina.

Definicija 1.1.13. *Vjerojatnosni prostor* $\mathbb{G}_{n,p} = (\Omega, \mathcal{F}, \mathbb{P})$ *slučajnih grafova je konačni vjerojatnosni prostor čije komponente su:*

- Ω = *skup svih grafova s vrhovima* $[n] = \{1, 2, \ldots, n\}$,
- $\mathcal{F} = \mathcal{P}(\Omega)$,
- $\mathbb{P}$ = *vjerojatnost takva da su događaji* $\{\{i, j\} \in E : \ i, j \in [n], \ i \neq j\}$ *nezavisni i:*

$$\mathbb{P}(\{i, j\} \in E) = p, \qquad \forall i, j \in [n], \ i \neq j.$$

Definicija 1.1.14. *Vjerojatnosni prostor* $\mathbb{G}_{n,m} = (\Omega, \mathcal{F}, \mathbb{P})$ *slučajnih grafova je konačni vjerojatnosni prostor u kojemu su*

- Ω = *skup svih grafova s vrhovima* $[n] = \{1, 2, \ldots, n\}$ *i* m *bridova,*
- $\mathcal{F} = \mathcal{P}(\Omega)$,
- $\mathbb{P}$ = *vjerojatnost takva da su svi elementarni događaji jednako vjerojatni.*

Slučajne elemente ovih vjerojatnosnih prostora zovemo slučajni grafovi, a uobičajeno se označavaju istim oznakama $\mathbb{G}_{n,p}$, odnosno $\mathbb{G}_{n,m}$. Najčešće se promatraju svojstva slučajnog grafa kada n teži u beskonačno. Tada fiksni m ili p daju trivijalne rezultate te se uzima da su p i m funkcije od n.

Postoje naravno i drugi vjerojatnosni prostori slučajnih grafova. U modelu $\mathbb{G}_{n,m}$ svoje su tvrdnje dokazivali Erdősa i Rényi, pa po njima i većina ranih radova. Definiranje vjerojatnosnog prostora $\mathbb{G}_{n,p}$ pripisuje se Edgaru Nelsonu Gilbertu u radu [6] iz 1959. Ovaj model se pokazao jednostavniji za korištenje. Slijedi teorem koji ih povezuje.

Teorem 1.1.15. *Neka je* $0 \leq p \leq 1$, $s(n) = n\sqrt{p(1-p)} \to \infty$, *i* $\omega(n) \to \infty$ *po volji sporo kada* $n \to \infty$.

(i) Neka je $\mathcal{P}$ *svojstvo grafova za koje vrijedi* $\mathbb{P}\left(\mathbb{G}_{n,m} \in \mathcal{P}\right) \to p_0$ *za sve*

$$m \in \left[\binom{n}{2}p - \omega(n)s(n), \binom{n}{2}p - \omega(n)s(n)\right].$$

Onda $\mathbb{P}\left(\mathbb{G}_{n,p} \in \mathcal{P}\right) \to p_0$ *kada* $n \to \infty$.

(ii) Neka za $p_- = p - \omega(n)s(n)/n^2$ *i* $p_+ = p + \omega(n)s(n)/n^2$ *i monotono svojstvo grafova* $\mathcal{P}$ *vrijedi* $\mathbb{P}\left(\mathbb{G}_{n,p_-} \in \mathcal{P}\right) \to p_0$ *i* $\mathbb{P}\left(\mathbb{G}_{n,p_+} \in \mathcal{P}\right) \to p_0$. *Onda* $\mathbb{P}\left(\mathbb{G}_{n,m} \in \mathcal{P}\right) \to p_0$ *kada* $n \to \infty$, *za* $m = \lfloor\binom{n}{2}p\rfloor$.

Iskaz ovog teorema pronađen je u knjizi Friezea i Karonskog [5], te se tamo dokaz pripisuje T. Łuczaku.

U proučavanju slučajnih grafova najčešće se radi o ponašanju slučajnih grafova u beskonačnosti, to jest za veliki broj vrhova $n \in \mathbb{N}$. Neka je $\mathcal{A} = \mathcal{A}_n$ svojstvo grafa s $n \in \mathbb{N}$ vrhova i $\mathbb{P} = \mathbb{P}_n$ vjerojatnost definirana na prostoru $\mathbb{G}_{n,p}$. Istražujemo

$$\lim_{n\to\infty} \mathbb{P}(\mathbb{G}_{n,p} \models \mathcal{A}).$$

Ako je ovaj limes jednak jedan kažemo da tipični graf $\mathbb{G}_{n,p}$ zadovoljava svojstvo $\mathcal{A}$, da $\mathbb{G}_{n,p}$ zadovoljava $\mathcal{A}$ s visokom vjerojatnošću (w.h.p. u literaturi) ili ga $\mathbb{G}_{n,p}$ zadovoljava $\mathcal{A}$ gotovo uvijek. Ovaj izraz se jednako definira i za $\mathbf{G}_{n,m}$ ili bilo koji indeksirani vjerojatnosni prostor i indeksirane događaje $\mathcal{A}_n$, što se koristi u Lemi 3.3.1.

1.2 Alati

Koristiti će se sljedeće asimptotske oznake:

$$
\begin{aligned}
a(n) \in o_n(b(n)) \Leftrightarrow &\qquad \lim_{n\to\infty} \frac{a(n)}{b(n)} = 0 \\
a(n) \ll b(n) \Leftrightarrow &\qquad a(n) \in o_n(b(n)) \\
a(n) \in O_n(b(n)) \Leftrightarrow &\qquad \limsup_{n\to\infty} \frac{a(n)}{b(n)} < \infty \\
a(n) \sim b(n) \Leftrightarrow &\qquad \lim_{n\to\infty} \frac{a(n)}{b(n)} = 1 \\
a(n) \in \Theta_n(b(n)) \Leftrightarrow &\qquad a(n) \in O_n(b_n),\ b(n) \in O_n(a(n)).
\end{aligned}
$$

Kod ovih oznaka n se često ne piše kada je jasan iz konteksta. Također, umjesto "$\in$" može stajati "=". Sva pravila računanja sa ovim oznakama neće se posebno obrazlagati.

Za mnoga pojavljivanja binomnih koeficijenata i faktorijela koristiti će se Stirlingova aproksimacija kojoj je jedan iskaz

$$n! \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n. \tag{1.1}$$

Iz gornje asimptotske tvrdnje slijedi da za dovoljno velike n

$$n! \geq \sqrt{n}\left(\frac{n}{e}\right)^n \geq \left(\frac{n}{e}\right)^n, \tag{1.2}$$

što također povlači sljedeću gornju ogradu na binomni koeficijent, koja vrijedi za dovoljno velike k

$$\binom{n}{k} \leq \left(\frac{ne}{k}\right)^k \tag{1.3}$$

Također vrijedi, po razvoju e^x u red potencija,

$$e^{\pm x} = 1 \pm x + O(x^2) \geq 1 \pm x, \qquad \forall x \in \langle 0, 1\rangle. \tag{1.4}$$

Koristiti ćemo i

$$\binom{n}{k} \leq \frac{n^k}{k!}\left(1 - \frac{k}{2n}\right)^{k-1} \leq \frac{n^k}{k!} e^{-\frac{k(k-1)}{2n}} \tag{1.5}$$

Sume indikatorskih varijabli

Neka je $(\Omega, \mathcal{F}, \mathbb{P})$ vjerojatnosni prostor, I skup indeksa i $(B_i)_{i\in I}$ skup događaja na tom vjerojatnosnom prostoru s vjerojatnostima $\mathbb{P}(B_i) = p_i$.

U daljnjem promatramo sljedeće slučajne varijable:

$$X_i := \mathbb{1}_{B_i}, \qquad i \in I,$$
$$X := \sum_{i\in I} X_i.$$

Primjenom linearnosti matematičkog očekivanja vidi se

$$\mathbb{E}X = \sum_{i\in I} p_i,$$

dok se za varijancu dobiva

$$\mathrm{Var}X = \sum_{i\in I} \mathrm{Var}(X_i) + \sum_{i,j\in I} \mathrm{Cov}(X_i, X_j). \tag{1.6}$$

Zato što su $(X_i)_{i\in I}$ indikatorske varijable, vrijedi i

$$\mathrm{Var}(X_i) = p_i(1 - p_i) \leq p_i \Rightarrow \sum_{i\in I} \mathrm{Var}(X_i) \leq \mathbb{E}X, \tag{1.7}$$

te

$$\mathrm{Cov}(X_i, X_j) = \mathbb{E}(X_i \cdot X_j) - \mathbb{E}(X_i) \cdot \mathbb{E}(X_j) \leq \mathbb{E}(X_i \cdot X_j) = \mathbb{P}(B_i \cap B_j).$$

Uzimamo u obzir još i nezavisnost. Označimo sa "$\sim$" relaciju na I gdje $i \sim j$ ako $i \neq j$ i B_i, B_j odnosno X_i, X_j su zavisni.

$$\sum_{i,j\in I} \mathrm{Cov}(X_i, X_j) = \sum_{\substack{i,j\in I\\ i\nsim j}} \underbrace{\mathrm{Cov}(X_i, X_j)}_{\substack{=0,\\ \text{nezavisnost}}} + \sum_{\substack{i,j\in I\\ i\sim j}} \mathrm{Cov}(X_i, X_j) \leq \sum_{\substack{i,j\in I\\ i\sim j}} \mathbb{P}(B_i \cap B_j). \tag{1.8}$$

Sada uvrštavanjem (1.7) i (1.8) u (1.6) dobivamo

$$\mathrm{Var}X \leq \mathbb{E}X + \sum_{\substack{i,j\in I\\ i\sim j}} \mathbb{P}(B_i \cap B_j),$$

pri čemu zadnja suma dolazi iz raspisa u (1.6), dakle ide po uređenim parovima $(i, j) \in I^2$ za koje vrijedi $i \sim j$, svaki neuređeni skup se pojavljuje točno dva puta (jer $i \sim j$ povlači $j \sim i$). Označimo li

$$\Delta := \sum_{\substack{i,j\in I\\ i\sim j}} \mathbb{P}(B_i \cap B_j) \tag{1.9}$$

dobivamo

$$\mathrm{Var}X \leq \mathbb{E}X + \Delta. \tag{1.10}$$

Ova opažanja koriste se primjerice kada je X broj pojavljivanja trokuta u $\mathbb{G}_{n,p}$. Tada je I skup svih mogućih neuređenih trojki vrhova, a B_i je događaj da na vrhovima u i postoji trokut u $\mathbb{G}_{n,p}$. U ovom slučaju postoji i simetrija, jer je svaka trojka vrhova u $\mathbb{G}_{n,p}$ ekvivalentna.

Simetrija

Ako vrijedi da je

$$\Delta^* := \sum_{j\sim i} \mathbb{P}(B_j|B_j) \tag{1.11}$$

neovisna o i, slijedi da je:

$$\Delta = \sum_{\substack{i,j\in I\\ i\sim j}} \mathbb{P}(B_i \cap B_j) = \sum_{\substack{i,j\in I\\ i\sim j}} \mathbb{P}(B_i)\mathbb{P}(B_j|B_j) = \sum_{i\in I} \mathbb{P}(B_i) \sum_{j\sim i} \mathbb{P}(B_j|B_j) = \mathbb{E}X \cdot \Delta^*,$$

što uvrštavanjem u (1.10) daje

$$\mathrm{Var}X \leq \mathbb{E}X(1 + \Delta^*). \tag{1.12}$$

Metode momenata

Navedimo prvo Markovljevu nejednakost koja vrijedi za sve nenegativne slučajne varijable X i sve $t > 0$:

$$\mathbb{P}(X \geq t) \leq \frac{\mathbb{E}X}{t}. \tag{1.13}$$

Za slučajnu varijablu X koja poprima vrijednosti u $\mathbb{N}_0$ uvrštavanjem $t = 1$ dobiva se

$$\mathbb{E}X \geq \mathbb{P}(X > 0). \tag{1.14}$$

Pa za niz takvih varijabli $(X_n)_{n\in\mathbb{N}}$ zbog monotonosti limesa slijedi

$$\lim_{n\to\infty} \mathbb{E}X_n = 0 \qquad \Rightarrow \qquad \lim_{n\to\infty} \mathbb{P}(X_n > 0) = 0. \tag{1.15}$$

Međutim uočimo da čak i kada $\mathbb{E}X_n \to \infty$ ne možemo na isti način zaključiti da

$$\lim_{n\to\infty} \mathbb{P}(X_n = 0) = 0.$$

U takvim slučajevima mogu se primijeniti metode drugog momenta.

Teorem 1.2.1. Čebiševljeva nejednakost: *Neka je X realna slučajna varijabla na* $(\Omega, \mathcal{F}, \mathbb{P})$ *s očekivanjem μ i standardnom devijacijom σ. Za svaki realni $\lambda > 0$ vrijedi:*

$$\mathbb{P}\left(|X - \mu| \geq \lambda\sigma\right) \leq \frac{1}{\lambda^2}.$$

Premećući tvrdnju teorema dobivamo i da vrijedi

$$\mathbb{P}(|X - \mathbb{E}X| \geq \epsilon|\mathbb{E}X|) \leq \frac{\mathrm{Var}(X)}{\epsilon^2(\mathbb{E}X)^2}, \qquad \forall\epsilon > 0.$$

Sada slijedi:

Korolar 1.2.2. *Neka je X_n niz realnih slučajnih varijabli i vrijedi* $\mathrm{Var}(X_n) \in o_n[(\mathbb{E}X_n)^2]$. *Tada* $(X_n/\mathbb{E}X_n) \xrightarrow{\mathbb{P}} 1$*, konkretno za bilo koji $\epsilon > 0$:*

$$\exists n_0 \in \mathbb{N} \qquad n > n_0 \Rightarrow \mathbb{P}\left(\left|\frac{X_n}{\mathbb{E}X_n} - 1\right| \geq \epsilon\right) \leq \epsilon.$$

Promotrimo slučaj kada je X_n suma indikatorskih varijabli i Δ definirana u (1.9), tada iz (1.5) dobivamo sljedeći korolar.

Korolar 1.2.3. *Ako $\mathbb{E}X_n \to \infty$ i $\Delta_n \in o_n[(\mathbb{E}X_n)^2]$, onda $(X_n/\mathbb{E}X_n) \xrightarrow{\mathbb{P}} 1$. Također, $\mathbb{P}(X_n > 0) \to 1$, kada $n \to \infty$.*

Ako dodatno vrijedi i simetrija, te je Δ^* kao u (1.11), tada iz (1.7) dobivamo i naredni korolar.

Korolar 1.2.4. *Ako $\mathbb{E}X_n \to \infty$ i $\Delta_n^* \in o_n(\mathbb{E}X_n)$, onda $(X_n/\mathbb{E}X_n) \xrightarrow{\mathbb{P}} 1$. Također, $\mathbb{P}(X_n > 0) \to 1$, kada $n \to \infty$.*

Jansonove nejednakosti

Neka su $(B_i)_{i\in I}$ događaji malih vjerojatnosti i "skoro nezavisni", idelano vrijedi:

$$\mathbb{P}\left(\bigcap_{i\in I} \overline{B_i}\right) \approx \prod_{i\in I} \mathbb{P}(\overline{B_i}) \approx \exp\left(-\sum_{i\in I} \mathbb{P}(B_i)\right).$$

Ako s X označimo slučajnu varijablu jednaku broju događaja B_i, $i \in I$ koji se ostvario, gornja izjava u terminima X postaje

$$\mathbb{P}(X = 0) \approx \exp\left(-\mathbb{E}X\right).$$

Dakle, X se, barem što se tiče događaja $X = 0$, ponaša kao Poissonova varijabla. Općenito bismo voljeli naći uvjete kada se suma indikatorskih varijabli rijetkih i nezavisnih događaja može aproksimirati Poissonovom distribucijom - tu aproksimaciju nazivaju Poissonovom paradigmom. U produktnim vjerojatnosnim prostorima gore navedene aproksimacije konkretizirane su u Jansonovim nejednakostima izloženim u dva teorema u nastavku. Izlaganjne ovih teorema je kao u [1].

Dani su konačan skup S, skup indeksa I i skup podskupova $A_i \subseteq S,\ i \in I$. Pomoću tih elemenata definira se konačni vjerojatnosni prostor $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$, gdje su elementarni događaji skupovi $R \subseteq S$ (dakle $\Omega := \mathcal{P}(S)$), a vjerojatnost je definirana sa

$$\mathbb{P}(s \in R) = p_s, \quad \forall s \in S,$$

i događaji $\{s \in R\}_{s \in S}$ su međusobno nezavisni. Dakle, elementarni događaj je slučajni skup $R \subseteq S$ u kojem se $s \in S$, nezavisno od ostalih $z \in S$ nalazi s vjerojatnošću p_s.

Sada neka su:

$$B_i := \{R \in \Omega; A_i \subseteq R\},\ \forall i \in I,$$

znači B_i je događaj da slučajno odabrani skup $R \subseteq S$ sadrži $A_i \subseteq S$. Kao i ranije, kažemo da $i \sim j$ ako $i \neq j$ te su B_i i B_j su zavisni.

Označimo:

$$\begin{aligned}
\mu &:= \sum_{i \in I} \mathbb{P}(B_i), \\
M &:= \prod_{i \in I} (1 - \mathbb{P}(B_i)) \approx \exp(-\mu), \\
\Delta &:= \sum_{j \sim i} \mathbb{P}(B_j \cap B_i).
\end{aligned}$$

Približna jednakost u drugom redu dolazi iz prikaza eksponencijalne funkcije kao reda ($e^x = 1 + x + O(x^2)$). Ovaj Δ je isti izraz koji se omeđuje u metodama drugog momenta i ide po svim uređenim parovima $i \sim j$. Primjećuje se da u ovom modelu vrijedi da su događaji B_i pozitivno zavisni jedni sa drugima, tj. vjerojatnost B_j se povećava ako znamo da se dogodila neka kombinacija ostalih B_j-ova. To je istina iz konstrukcije događaja $(B_i)_{i \in I}$, jer se B_i dogodio ako slučajni skup R sadrži A_i, i ako $A_i \cap A_j \neq \emptyset$ iz B_i znamo da R već sadržava neke elemente iz A_j.

Teorem 1.2.5. *Ako su $(B_i)_{i \in I}, M, \mu$ i Δ definirani kao gore i $\epsilon \geq \mathbb{P}(B_i),\ \forall i \in I$, onda vrijedi:*

$$M \leq \mathbb{P}\left(\bigcap_{i \in I} \overline{B_i}\right) \leq M \exp\left(\frac{1}{1-\epsilon} \cdot \frac{\Delta}{2}\right)$$

i

$$\mathbb{P}\left(\bigcap_{i\in I}\overline{B_i}\right) \leq \exp\left(-\mu + \frac{\Delta}{2}\right).$$

Teorem 1.2.6. *Ako su $(B_i)_{i\in I}, M, \mu$ i Δ definirani kao gore i $\Delta > \mu$, onda vrijedi:*

$$\mathbb{P}\left(\bigcap_{i\in I}\overline{B_i}\right) \leq \exp\left(-\mu^2/2\Delta\right).$$

Ograda iz drugog teorema je jača kada je se može upotrebljavati, ta snaga se dobiva korištenjem vjerojatnosnih metoda na ograde u prvom teoremu. Oba teorema uglavnom daju bolje ograde nego Čebiševljeva nejednakost.

Poglavlje 2

Podgrafovi

Promatramo proizvoljni fiksni graf $H = ([v], E_H)$ s v poredanih vrhova i $|E_H| = e$ bridova. Označimo sa $\mathcal{A}$ svojstvo da $\mathbf{G}_{n,p}$ sadrži H kao podgraf. Od interesa je

$$\lim_{n\to\infty} \mathbb{P}(\mathbb{G}_{n,p} \models \mathcal{A}).$$

Ako je ovaj limes jednak nula kažemo da se H gotovo sigurno ne pojavljuje u $\mathbb{G}_{n,p}$, a ako je jednak jedan kažemo da se H gotovo sigurno pojavljuje u $\mathbb{G}_{n,p}$.

Precizno govoreći, H je podgraf grafa $G = ([n], E)$ na uređenoj v-torki $(x_1, \dots x_v)$ različitih vrhova iz $[n]$ ako:

$$\{a, b\} \in E_H \Rightarrow \{x_a, x_b\} \in E, \qquad \forall a, b \in [v].$$

Iz ove definicije se vidi da dodavanje brida u G koji već ima H kao podgraf neće dovesti do toga da H više nije podgraf u G. To upravo po definiciji znači da je $\mathcal{A}$ monotono rastuće svojstvo grafova. Za takvo svojstvo definiramo funkciju praga.

Definicija 2.0.1. *Za monotono rastuće svojstvo grafova $\mathcal{A}$, $p^* = p^*(n)$ je funkcija praga ako:*

$$\begin{cases} \lim_{n\to\infty} \mathbb{P}(\mathbb{G}_{n,p} \models \mathcal{A}) = 0, & p \ll p^*(n) \\ \lim_{n\to\infty} \mathbb{P}(\mathbb{G}_{n,p} \models \mathcal{A}) = 1, & p \gg p^*(n). \end{cases}$$

Pored ove definicije valja napomenuti da funkcija praga nikada nije jedinstvena, već da ako je p^* funkcija praga, onda je to i svaka funkcija $\Theta(p^*)$.

U ovom poglavlju ćemo prvo pronaći funkciju praga $p^* = p^*(n)$ za $\mathcal{A}$. Odnosno da se za $p \ll p^*$ H gotovo sigurno ne pojavljuje u $\mathbb{G}_{n,p}$, a za $p \gg p^*$ se H gotovo sigurno pojavljuje u $\mathbb{G}_{n,p}$. Prvo će se pokazati da je za određenu klasu grafova H

$$p^* = n^{-v/e},$$

a potom da je za ostale grafove

$$p^* = n^{-v'/e'},$$

gdje su v' i e' redom broj vrhova i bridova "najgušćeg podgrafa".

Zatim će se u Potpoglavlju 2.2 pokazati da za $p = cn^{-v/e}$ i H iz određene klase grafova, s a automorfizama, preciznije vrijedi

$$\lim_{n\to\infty} \mathbb{P}(\mathbb{G}_{n,p} \not\models \mathcal{A}) = \exp(-c^e/a).$$

Konačno, u Potpoglavlju 2.3 usporediti će se ova dva teorema, dokazi i rezultati.

Pristup za proučavanje ovog svojstva biti će promatrati varijablu X:

$$X = \text{broj pojavljivanja grafa } H \text{ u } \mathrm{G}_{n,p}.$$

Veza svojstva $\mathcal{A}$ i varijable X je

$$\mathcal{A} \Leftrightarrow X > 0. \tag{2.1}$$

X je suma indikatorskih varijabli X_s, gdje svaka varijabla broji jednu moguću realizaciju H u $\mathbb{G}_{n,p}$, to jest

$$X = \sum_{s\in S} X_s.$$

Sada je pitanje po kojim indeksima $s \in S$ brojimo različite realizacije H u $\mathbb{G}_{n,p}$. Svakako, svaka realizacija H kao podgrafa od $\mathbb{G}_{n,p}$ veže jednu v-torku različitih vrhova u $\mathbb{G}_{n,p}$. No tu se događaju prebrojavanja. Naime, ako se H dogodi na v-torci x i f je automorfizam od H, H se realizirao i na v-torci $f(x)$. Dakle, svako pojavljivanje se broji a puta, gdje je

$$a = \text{broj automorfizama od } H.$$

Neka je:

$$x \diamond y \overset{def}{\Leftrightarrow} \exists f \text{ automorfizam od H takav da } y = f(x).$$

Ovo je relacija ekvivalencije na skupu svih v-torki različitih vrhova $\mathbb{G}_{n,p}$, koje nazivamo i "reprezentativnim v-torkama". Neka je S skup klasa ekvivalencije ove relacije. Neka je B_s, $s \in S$ događaj da je H podgraf od $\mathbb{G}_{n,p}$ na v-torci $x \in s$. Događaj B_s je dobro definiran kojeg god predstavnika $x \in s$ se uzelo, i neka je

$$X_s := \mathbb{1}_{B_s}, \quad s \in S.$$

Postoji $\binom{n}{v} \cdot v!$ v-torki različitih vrhova u $\mathbb{G}_{n,p}$, i po a njih je u istoj klasi ekvivalencije. Tako da je broj klasa ekvivalencije jednak

$$|S| = \frac{1}{a} \cdot \binom{n}{v} \cdot v!.$$

Računamo očekivanje od X. Po linearnosti očekivanja slijedi

$$\mathbb{E}X = \sum_{s \in S} \mathbb{E}X_s,$$

S ima $\frac{1}{a}\binom{n}{v} \cdot v!$ elemenata i za svaki $s \in S$, $\mathrm{E}X_s = \mathrm{P}(X_s = 1) = p^e$, dakle:

$$\mathrm{E}X = \frac{1}{a} \cdot \binom{n}{v} \cdot v! \cdot p^e = \Theta(n^v p^e). \tag{2.2}$$

Iz ovoga je vidljivo da ako p držimo fiksnim, a $n \to \infty$, očekivani broj pojavljivanja bilo kojeg grafa isto teži u beskonačnost. To nažalost nije dovoljno da se kaže da ako je p fiksan, svaki graf H se gotovo sigurno pojavljuje u $\mathbb{G}_{n,p}$. No čak i ako je $p = p(n)$ padajuća funkcija od n do granice $p = n^{-v/e}$, i dalje se očekuje više od 0 pojavljivanja H kada $n \to \infty$.

Definicija 2.0.2. Gustoća grafa *G s e bridova i v vrhova je* $\rho(G) = \dfrac{e}{v}$.

Jednadžba (2.2) također znači da veća gustoća grafa H smanjuje očekivani broj pojavljivanja grafa H u $\mathbb{G}_{n,p}$. Dakle, gusti podgrafovi će biti "usko grlo" pojavljivanja cijelog grafa u $\mathbb{G}_{n,p}$. Zato definiramo i grafove koji su gušći od svih svojih podgrafova.

Definicija 2.0.3. Graf G je balansiran *ako za svaki podgraf* $G' \sqsubseteq G$ *vrijedi* $\rho(G') \leq \rho(G)$. Graf je strogo balansiran *ako za svaki pravi podgraf* $G' \sqsubset G$ *vrijedi* $\rho(G') < \rho(G)$.

Ako graf H nije balansiran, recimo postoji $H' \sqsubseteq H$ s e' bridova i v' vrhova gdje $\rho(H') > \rho(H)$, tada

$$n^{-v/e} < n^{-v'/e'}.$$

Odnosno kada $n \to \infty$,

$$n^{-v/e} \ll n^{-v'/e'}.$$

Iz ovoga se vidi da iako bi za $p = cn^{-v/e}$ iz (2.2) očekivani broj pojavljivanja H bio neki realni broj, očekivani broj pojavljivanja H' bi težio u 0, pa se (detaljnije pokazano u Teoremu 2.1.2) H' neće pojavljivati gotovo nikada, kao onda i H. Tu se još jednom vidi da očekivanje nije dovoljna mjera slučajne varijable.

2.1 Funkcija praga

Teorem 2.1.1. *Neka je H fiksan, balansiran graf i neka je $\mathcal{A}$ svojstvo da $\mathbf{G}_{n,p}$ sadrži H kao podgraf. Tada je*

$$p = n^{-1/\rho(H)}$$

funkcija praga za $\mathcal{A}$.

Dokaz. Neka je kao i prije broj vrhova od H jednak v i broj bridova jednak e. Također, koristimo i slučajnu varijablu X jednaku broju pojavljivanja grafa H u $\mathbb{G}_{n,p}$. Imajući na umu vezu između X i $\mathcal{A}$ iz (2.1), da $n^{-1/\rho(H)} = n^{-v/e}$ bude funkcija praga za $\mathcal{A}$ treba dokazati[1]:

$$\lim_{n\to\infty} \mathbb{P}(X > 0) = 0 \qquad p \ll n^{-v/e}, \tag{2.3}$$

$$\lim_{n\to\infty} \mathbb{P}(X > 0) = 1 \qquad p \gg n^{-v/e}. \tag{2.4}$$

Neka je prvo $p \ll n^{-v/e}$, iz toga direktno dobivamo

$$n^v p^e = o(1).$$

Iz prijašnje diskusije te formule (2.2), za očekivanje od X vrijedi

$$\mathbb{E}X = \Theta(n^v p^e).$$

Te dvije činjenice daju

$$\mathbb{E}X = o(1),$$

odnosno ekvivalentno $\lim_{n\to\infty} \mathbb{E}X = 0$.

Kako $\mathbb{E}X \to 0$ i X poprima vrijednosti u $\mathbb{N}_0$ iz tvrdnje (1.15) slijedi

$$\lim_{n\to\infty} \mathbb{P}(X > 0) = 0.$$

Dakle, dokazano je (2.3).

Sada neka je $p \gg n^{-v/e}$. To znači da je

$$n^v p^e \to \infty,$$

što zajedno s (2.2) daje

$$\mathbb{E}X \to \infty.$$

Uočimo sada da je X suma simetričnih indikatorskih varijabli. Simetrija dolazi iz toga što je svaka uređena v-torka vrhova u prostoru $\mathbb{G}_{n,p}$ ekvivalentna po definiciji. Pa tako i vjerojatnosti $\mathbb{P}(\cdot|X_s = 1)$ tvore identične vjerojatnosne prostore.

Neka je $i \in S$ fiksan. Dobivamo

$$\Delta^* = \sum_{j\sim i} \mathbb{P}(X_j = 1|X_i = 1).$$

[1]Tumačenje danih limesa je: i vjerojatnosti $\mathbb{P}$ i varijable X definirane su prostorom slučajnih grafova $\mathbb{G}_{n,p}$, ovisnom o n.

Za neki $j \sim i$, neka je broj zajedničkih vrhova od i i j jednak k. Da bi vrijedilo $i \sim j$, po definiciji da su X_i i X_j korelirane, nužno je da $2 \leq k \leq v$. Broj reprezentativnih v-torki $j \sim i$ koji imaju k zajedničkih točaka sa i je:

$$\Theta(n^{v-k}),$$

asimptotski s obzirom na n.

Neka je l broj bridova koji je zajednički realizaciji H na i i realizaciji H na j. Vrijedi

$$\mathbb{P}(X_j = 1|X_i = 1) = p^{v-l}.$$

Kako je H balansiran, a tih l bridova i k vrhova su dio induciranog podgrafa od H, slijedi

$$l \leq \frac{ek}{v}.$$

Sada imamo

$$\mathbb{P}(X_j = 1|X_i = 1) \leq p^{e-\frac{ek}{v}} = p^{e(1-k/v)}.$$

Sve dobiveno uvršteno u definicijsku formulu za Δ^* daje

$$\begin{aligned}
\Delta^* &= \sum_{k=2}^{v} \sum_{\substack{j\sim i \\ |i\cap j|=k}} \mathbb{P}(X_j = 1|X_i = 1) \leq \sum_{k=2}^{v} \Theta(n^{v-k}) p^{e(1-k/v)} \\
&= \sum_{k=2}^{v} \Theta((n^v p^e)^{1-k/v}) = \sum_{k=2}^{v} o(n^v p^e) = o(n^v p^e).
\end{aligned}$$

Kako je $\mathbb{E}X = \Theta(n^v p^e)$ slijedi

$$\Delta^* = o(\mathbb{E}X).$$

Pa po Korolaru 1.2.4,

$$\lim_{n\to\infty} \mathbb{P}(X > 0) = 1,$$

čime je dobiveno i (2.4).

□

Teorem 2.1.2. *Neka je $H' \sqsubseteq H$ podgraf grafa H s najvećom gustoćom. Tada je*

$$p^* = n^{-1/\rho(H')}$$

funkcija praga za svojstvo $\mathcal{A}$ - H je podgraf $\mathbb{G}_{n,p}$.

Dokaz. Za bilo koji graf K neka X_K označava slučajnu varijablu koja je jednaka broju pojavljivanja K u $\mathbb{G}_{n,p}$ i $\mathcal{A}_K$ događaj da je K podgraf od $\mathbb{G}_{n,p}$.

Neka je $p \ll n^{-1/\rho(H')}$.

H' je balansiran. Naime, zato što je H' podgraf od H s najvećom gustoćom, dakle H' ima veću gustoću od svih svojih podgrafova jer su oni podgrafovi od H. Po prethodnom teoremu, $n^{-1/\rho(H')}$ je funkcija praga za $\mathcal{A}_{H'}$. Dakle,

$$\lim_{n\to\infty} \mathbb{P}(\mathbb{G}_{n,p} \models \mathcal{A}_{H'}) = 0.$$

Također, jer $H' \sqsubset H$ vrijedi $\mathbb{G}_{n,p} \models \mathcal{A} \Rightarrow \mathbb{G}_{n,p} \models \mathcal{A}_{H'}$ zbog čega

$$\mathbb{P}(\mathbb{G}_{n,p} \models \mathcal{A}_{H'}) \geq \mathbb{P}(\mathbb{G}_{n,p} \models \mathcal{A}).$$

Konačno,

$$\lim_{n\to\infty} \mathbb{P}(\mathbb{G}_{n,p} \models \mathcal{A}) = 0.$$

Neka je nadalje $p \gg n^{-1/\rho(H')}$.

Svaki graf $H'' \sqsubseteq H$ s v vrhova i e bridova je manje gust od H' to jest $\rho(H') \geq \frac{e}{v}$. Očekivanje varijable $X_{H''}$ je po formuli s početka poglavlja:

$$\mathbb{E}X_{H''} = \Theta(n^v p^e) = \Theta[(np^{e/v})^v] \gg (n \cdot n^{(-1/\rho(H'))\cdot(e/v)})^v = (n^{1-(e/v)/\rho(H')})^v \to \infty.$$

Ovo uključuje i slučaj $H'' = H$.

Ograničimo Δ za H, kao u prethodnom teoremu. Za reprezentativne v-torke i i j na kojima se H može pojaviti, ako $i \sim j$ znači da je pojavljivanje H na i i j korelirano, tj. H na i i H na j imaju $l > 0$ bridova zajedničkih. Neka je k broj vrhova koji se pojavljuju i u i i u j. Tih l bridova i k vrhova čine podgraf od H (i od H na i i od H na j), što znači da $\Theta(n^k p^l) \to \infty$.

$$\Delta = \sum_{k,l} \sum_{\substack{i\sim j \\ |i\cap j|=k}} \mathbb{P}(X_i \cap X_j) = \sum_{k,l} \Theta(n^{2v-k}) p^{2e-l} = \Theta\left(\frac{(n^v p^e)^2}{n^k p^l}\right) = o[(\mathbb{E}X)^2].$$

Sada po Korolaru 1.2.3 iz metode drugog momenta slijedi da

$$\lim_{n\to\infty} \mathbb{P}(X > 0) = \lim_{n\to infty} \mathbb{P}(\mathbb{G}_{n,p} \models \mathcal{A}) = 1.$$

$\square$

2.2 Prijelazno područje

Teorem 2.2.1. *Neka je H strogo balansiran graf sa v vrhova, e bridova i a automorfizama i neka je $p = cn^{-v/e}$, $c > 0$, konstanta. Neka je $\mathcal{A}$ svojstvo da $\mathbb{G}_{n,p}$ sadrži H. Tada vrijedi*

$$\lim_{n\to\infty} \mathbb{P}(\mathbb{G}_{n,p} \not\models \mathcal{A}) = \exp\left(-c^e/a\right).$$

Dokaz. Neka je X kao i dosada, broj pojavljivanja grafa H u $\mathbb{G}_{n,p}$, X_i indikatorska varijabla za pojavljivanje H na $i \in S$ i S skup različitih mogućih pojavljivanja grafa H od ranije. Po Jansonovoj nejednakosti, tj. Teoremu 1.2.5:

$$M \leq \mathbb{P}(X = 0) \leq M \exp\left(\frac{1}{1-\epsilon} \cdot \frac{\Delta}{2}\right), \tag{2.5}$$

pri čemu je

$$M := \exp(-\mathbb{E}X) = \exp\left[-\frac{1}{a}\binom{n}{v} \cdot v! \cdot p^e\right] \to \exp\left(-\frac{c^e}{a}\right).$$

Prva jednakost je uvrštavanje otprije izračunatoga očekivanja $\mathbb{E}X$, a druga uvrštavanje $p = cn^{-v/e}$ iz pretpostavke.

Definicija Δ je po neuređenim parovima $i \sim j$

$$\Delta = \sum_{i,j,\ i\sim j} \mathbb{P}(\{X_i = 1\} \cap \{X_j = 1\}).$$

Ovu sumu može se dodatno grupirati s obzirom na to koliko je vrhova u presjeku reprezentativnih v-torki i i j. Zato što $i \sim j$ slijedi da $2 \leq k \leq v$

$$\Delta = \sum_{k=2}^{v} \sum_{\substack{j\sim i \\ |i\cap j|=k}} \mathbb{P}(\{X_j = 1\} \cap \{X_i = 1\}).$$

Neka je j odabrana tako da maksimizira l - broj zajedničkih bridova realizaciji H na i i H na j. Tih l bridova samo su dio induciranog podgrafa od H na zajedničkih k vrhova (recimo podgraf u H na i). Graf H je strogo balansiran, pa slijedi

$$l < \frac{e}{v} \cdot k \qquad \left(\Leftrightarrow k > \frac{v}{e} \cdot l\right). \tag{2.6}$$

A da se dogode i $\{X_i = 1\}$ i $\{X_j = 1\}$ mora se realizirati svih $2e$ bridova, no njih l je zajedničko, dakle:

$$\mathbb{P}(\{X_j = 1\} \cap \{X_i = 1\}) = p^{2e-l}.$$

Ako $p = O(n^{-v/e})$, onda

$$\mathbb{P}(\{X_j = 1\} \cap \{X_i = 1\}) = O(n^{\frac{-v}{e}(2e-l)}) = O(n^{-2v+\frac{v}{e}l}).$$

Za sve ostale j, gdje broj bridova u presjeku i nije najveći mogući, i dalje vrijedi ova asimptotska tvrdnja.

Izbora za i i j s k zajedničkih vrhova ima $O(n^{2v-k})$, jer kada se svih $2v-k$ vrhova u $i \cup j$ odaberu, broj preraspodjela je konačan. Konačno:

$$\Delta = \sum_{k=2}^{v} O(n^{2v-k}) \cdot O(n^{-2v+\frac{v}{e}l}) = \sum_{k=2}^{v} O(n^{\frac{v}{e}l-k}).$$

Uz (2.6):

$$O(n^{\frac{v}{e}l-k}) = o(1).$$

Budući da je suma u Δ konačna, slijedi

$$\Delta = o(1).$$

Za ϵ se može uzeti $\epsilon = p^e = c^e n^{-v}$, dakle $1-\epsilon$ je omeđen odozdo te

$$\frac{1}{1-\epsilon} \cdot \frac{\Delta}{2} = o(1).$$

Jansonova nejednakost s početka dokaza i teorem o sendviču povlače

$$\lim_{n\to\infty} \mathbb{P}(\mathbb{G}_{n,p} \not\models \mathcal{A}) = M = \exp\left(-\frac{c^e}{a}\right).$$

□

U ovom teoremu objekt interesa je bilo svojstvo da $\mathbb{G}_{n,p}$ sadrži fiksni graf H. I pokazano je da je vjerojatnost da je broj pojavljivanja strogo balansiranog grafa H nula upravo vjerojatnost da Poissonova varijabla s očekivanjem $\lambda = c^e/a$ poprimi vrijednost nula. Općenito je točno da je broj pojavljivanja strogo balansiranog grafa H Poissonova varijabla. Slijedi formulacija te tvrdnje kao teorema čiji je dokaz dostupan u [2].

Teorem 2.2.2. *Neka je H fiksan strogo balansiran graf s v vrhova, $e \geq 2$ bridova i a automorfizama. Neka je c pozitivna konstanta i $p = cn^{-v/e}$. Neka je X broj pojavljivanja grafa H u $\mathbb{G}_{n,p}$. Vrijedi*

$$X \xrightarrow{d} Poiss(c^e/a).$$

2.3 Zaključak

U oba teorema promatra se vjerojatnost

$$\mathbb{P}(X = 0)$$

pri čemu je X varijabla koja broji sva pojavljivanja grafa H u $\mathbb{G}_{n,p}$. Teorem 2.1.1 koristi samo metode drugog momenta i uspijeva naći funkciju praga za balansirane grafove. Detaljnije, Teorem 2.2.1 aproksimira tu konkretnu vjerojatnost za dovoljno velike n.

Ključna veličina Δ je mjera interakcije između varijabla $X_i,\ i \in S$ to jest zavisnosti realiziranja grafa H na različitim $i \sim j$. Iz dokaza Teorema 2.2.1 vidi se da je $\Delta = o(1)$ kada je graf strogo balansiran i $p = p^* = cn^{-1/\rho(H)}$, a to je naravno slučaj i kada je $p \ll p^*$ čak i kada graf nije strogo balansiran. Također se u dokazu Teorema 2.1.1 vidi da $\Delta = o(n^v p^e)$ kada $p \gg n^{-v/e}$.

Stroga balansiranost je bitna za zaključak $O(n^{vl/e-k}) = o(1)$, iz dokaza Teorema 2.2.1. Taj zaključak ne bi vrijedio da je graf samo balansiran, pa može vrijediti $k = \dfrac{v}{e}l$. Odgovarajući dio u slabijemu Teoremu 2.1.1 je kada se dobije $\Delta^* \leq o(n^v p^e)$, ali ovdje nije bitno je li jednakost u pitanju stroga, jer u svakom slučaju $n^v p^e \to \infty$.

Neka je $p = cn^{-v/e}$. Brzina konvergencije $\mathbb{P}(X = 0)$ u Teoremu 2.1.1 je ograničena na sljedeći način

$$\begin{cases} c \to 0 & P(X > 0) \leq \mathbb{E}X = \Theta(n^v p^e) = \Theta(c^e), \\ c \to \infty & P(X = 0) \leq \dfrac{1 + \Delta^*}{\mathbb{E}X} \leq \dfrac{1 + \Theta((n^v p^e)^{1-2/v})}{\Theta(n^v p^e)} = \dfrac{\Theta(c^{e(1-2/v)})}{\Theta(c^e)} = \Theta(c^{-2e/v}). \end{cases}$$

U Teoremu 2.2.1 brzina konvergencije ograničena je sa:

$$\exp\left(\frac{1}{1-\epsilon} \cdot \frac{\Delta}{2}\right) \leq \exp(\Delta) \leq \exp[O(n^{vl/e-k})],$$

ako su vjerojatnosti da se graf realizirao na nekoj v-torki manje od $1/2$. Dakle, što je manja gustoća idućeg najgušćeg podgrafa u H, to je bliža asimptotska vjerojatnost iz tvrdnje teorema. Uzmimo za primjer da je H trokut. Trokut je strogo balansiran graf s $a = 3$ automorfizma i gustoća mu je $\rho(H) = 1$. Neka je $c \in \langle 0, n\rangle$ i

$$p(n) = \frac{c}{n}.$$

Iz Teorema 2.1.1 za dovoljno velike n proizlazi:

$$\begin{cases} c \to 0: & \mathbb{P}(\mathbb{G}_{n,p} \models \mathcal{A}) \approx 0, \\ c \to \infty: & \mathbb{P}(\mathbb{G}_{n,p} \models \mathcal{A}) \approx 1. \end{cases}$$

Teorem 2.2.1 daje za konstantan c i dovoljno velike n

$$\mathbb{P}(\mathbb{G}_{n,p} \models \mathcal{A}) \approx 1 - \exp\left(-\frac{c^3}{3}\right).$$

Još jedan primjer mogao bi biti kada je H stablo. Stabla su balansirani podgrafovi iako imaju najmanju gustoću od svih povezanoga grafova. Budući da je H stablo s v vrhova $\rho(H) = \frac{v-1}{v}$. Funkcija praga za postojanje stabla H u $\mathbb{G}_{n,p}$ je:

$$p^*(n) = n^{-v/(v-1)} \ll n^{-1}.$$

Dakle, graf će prije (ako krenemo od manjih p) sadržavati stablo veličine v za bilo koji v nego trokut. Prethodna tvrdnja već je bliže ideji faznog prijelaza slučajnog grafa. Fazni prijelaz naime proučava kada se s obzirom na p neka svojstva pojavljuju.

Proučavanjem faznog prijelaza, proširiti će se pitanje ovog poglavlja tako da će graf $H = H(n)$ biti ovisan o n. Graf H bi mogao biti na primjer: put duljine n, put duljine $\log(n)$, stablo veličine n.

Poglavlje 3

Evolucija slučajnih grafova

Evolucija slučajnog grafa je promjena strukture tipičnog grafa kako se p mijenja od brže opadajućih funkcija npr. n^{-2} do sporije opadajućih npr. $p = \ln n/n$.

Ovdje spomenuti rezultati odnose se na veličinu i kompleksnost komponenti, a posebno na razlike u veličinama komponenti tipičnog grafa. U ovom će se poglavlju također dokazivati dio tvrdnji o glasovitom dvostrukom skoku u veličini divovske komponente.

Označimo sa L_i broj vrhova u i-toj komponenti po veličini. Također, podsjetimo se da stabla imaju minimalan broj bridova ($k-1$ ako je k vrhova), a ako se doda jedan brid dobiva se graf sa jednim ciklusom koji nazivamo unicikličnim grafom. Te dvije klase grafova zvati ćemo *jednostavnima*, dok grafove sa dva i više ciklusa zovemo *kompleksnima*.

U ovom poglavlju cilj je opisati evoluciju slučajnog grafa $\mathbb{G}_{n,p}$ i predstaviti nekoliko iz mnoštva rezultata koji se odnose na ovu temu. Prvo potpoglavlje pokušava dočarati kako se slučajni graf razvija te potom iznosi pregled režima kako je opisano u [1]. Drugo potpoglavlje dokazuje rezultate o veličini najveće komponente za $p = c/n$, $c < 1$ i $c > 1$ odnosno u vrlo subkritičnom i vrlo superkritičnom režimu. Dokazi u tom poglavlju po uzoru su na [5]. Naredno potpoglavlje prikazuje drugačiji i jednostavniji dokaz sličnih rezultata, po uzoru na nedavni članak [7] autora Krivelevich i Sudakov iz 2012. godine.

3.1 Pregled

Odredimo prije svega relevantni raspon parametra p. Uočimo da je brid grafa fiksni podgraf s dva vrha i jednim bridom, dakle gustoćom $\rho = 1/2$. Po Teoremu 2.1.1 iz prethodnog poglavlja, prvi brid pojavljuje se za $p = cn^{-2}$. S druge strane neka je $p \in \langle 0, 1\rangle$ konstanta i H proizvoljni fiksni graf. Svaki graf ima gustoću veću od nule, stoga $p \gg n^{-1/\rho(H)}$ i po Teoremu 2.1.1 iz prethodnog poglavlja $\mathbb{G}_{n,p}$ sadržava s visokom vjerojatnošću proizvoljno odabran graf H. Promatra se struktura $\mathbb{G}_{n,p}$ kada p pomičemo od vrlo male (preciznije brzo opadajuće) razine n^{-2} do konstante n^0.

Krenimo sada postupno sužavati taj raspon. Kada je $p \in o(n^{-1})$, $\mathbb{G}_{n,p}$ je šuma. Kako bismo to argumentirali, pretpostavimo suprotno: u $\mathbb{G}_{n,p}$ postoji ciklus C_k veličine $k \in \mathbb{N}$. Očekivani broj k-ciklusa manji je od $(np)^k \in o(1)$, vrijedi da $\mathbb{G}_{n,p}$ s visokom vjerojatnošću ne sadrži ciklus C_k. Slijedi da $\mathbb{G}_{n,p}$ s visokom vjerojatnošću ne sadrži ciklus.

Sljedeći teorem iz [2] opisuje ponašanje komponenata koje su stabla veličine k. Ilustrativan je i zanimljiv te se iznosi ovdje bez dokaza.

Primijetimo prije iskaza teorema da je različito promatranje nekog stabla kao podgrafa, kao što se radilo u prethodnom poglavlju, od promatranja stabla koje je komponenta. Fiksno stablo na k vrhova je podgraf u $\mathbb{G}_{n,p}$ sa vjerojatnošću p^{k-1}, dok je to stablo izolirana komponenta sa vjerojatnošću $p^{k-1}(1-p)^{nk+\binom{k}{2}-k+1}$.

Teorem 3.1.1. *Označimo sa T_k broj komponenti u $\mathbb{G}_{n,p}$ koje su stabla veličine $k \geq 2$.*

(i) Ako $p \in o\left(n^{-k/(k-1)}\right)$, onda s visokom vjerojatnošću u $\mathbb{G}_{n,p}$ je $T_k = 0$.

(ii) Ako $p = cn^{-k/(k-1)}$ za neku konstantu $c > 0$, onda u $\mathbb{G}_{n,p}$ je $T_k \xrightarrow{d} Poiss\left(\dfrac{c^{k-1}k^{k-2}}{k!}\right)$.

(iii) Ako $pn^{k/(k-1)} \to \infty$ i $(pkn - \ln n - (k-1)\ln\ln n) \to -\infty$, onda za svaki $L \in \mathbb{R}$ vrijedi $\mathbb{P}(T_k > L) \to 1$.

(iv) Ako $pn^{k/(k-1)} \to \infty$ i $(pkn - \ln n - (k-1)\ln\ln n) \to x \in \mathbb{R}$, onda je $T_k \xrightarrow{d} Poiss\,(e^{-x}/(k \cdot k!))$.

(v) Ako $pn^{k/(k-1)} \to \infty$ i $(pkn - \ln n - (k-1)\ln\ln n) \to \infty$, onda s visokom vjerojatnošću u $\mathbb{G}_{n,p}$ je $T_k = 0$.

Jedno tumačenje ovog teorema je da očekivani broj komponenata koje su stabla veličine k u tipičnom grafu $\mathbb{G}_{n,p}$ raste do razine opisane u (*iii*) te nakon toga opet opada. Primijetimo da ta razina uključuje $p \in \Theta(n^{-1})$.

Intuitivno, činilo bi nam se da to što se broj komponenata koje su stabla veličine k smanjuje što je je p veći proizlazi iz toga da su komponente veće i kompleksnije što je p veći. U nastavku ćemo vidjeti kako veličina i kompleksnost komponenata ne rastu svugdje u $\mathbb{G}_{n,p}$. Umjesto toga izdvaja se jedna komponenta čija se kompleksnost i veličina povećavaju, i koja "guta" ostale komponente prije nego što one same postanu kompleksnije.

Najzanimljivije promjene se događaju za $p \in \Theta(1/n)$. Gruba parametrizacija za p je

$$p = \frac{c}{n}, \quad c > 0 \text{ konstanta.}$$

Razlikujući vrijednosti konstante c razdvajaju se tri faze, odnosno režima.

Subkritičan režim

Kažemo da se $\mathbb{G}_{n,p}$ nalazi u subkritičnom režimu kada je $p = c/n$, $c < 1$. Takav graf $\mathbb{G}_{n,p}$ tipično ne sadrži kompleksne komponente. Vrijedi i:

$$
\begin{aligned}
&L_1 \in \Theta(\ln n),\\
&L_k \sim L_1,\ \forall k \in \mathbb{N}.
\end{aligned}
$$

Kritičan režim

Kažemo da se $\mathbb{G}_{n,p}$ nalazi u subkritičnom režimu kada je $p = 1/n$. Najveća komponenta je veličine $L_1 \in \Theta(n^{2/3})$.

Superkritičan režim

Kažemo da se $\mathbb{G}_{n,p}$ nalazi u superkritičnom režimu kada je $p = c/n$, $c > 1$. Takav graf $\mathbb{G}_{n,p}$ tipično sadrži jednu dominantnu komponentu veličine $L_1 \sim yn$, gdje je $y \in \langle 0, \infty\rangle$ rješenje jednadžbe $e^{-cy} = 1 - y$ i ostale komponente koje su jednostavne i veličine $O(\ln n)$.

Detaljnije za $p \sim 1/n$

Kada želimo detaljnije promatrati kako se $\mathbb{G}_{n,p}$ ponaša za $p \sim 1/n$ uvode se još sljedeće ekvivalentne parametrizacije

$$
\begin{aligned}
p &= \frac{1}{n} \pm \lambda n^{-4/3},\\
p &= \frac{1 \pm \epsilon}{n},\ \epsilon = \lambda n^{-1/3}.
\end{aligned}
$$

Ove parametrizacije zovemo finim parametrizacijama. Sada se može razlikovati jedva subkritičan režim, kritični prozor i jedva superkritičan režim. U tom kontekstu se subkritičan i superkritičan režim nazivaju i vrlo subkritičan odnosno vrlo superkritičan.

Jedva subkritičan režim

Kažemo da se $\mathbb{G}_{n,p}$ nalazi u jedva subkritičnom režimu kada $p = \frac{1-\epsilon}{n} = \frac{1}{n} - \lambda n^{-4/3}$ i to tako da $\epsilon \to 0$, $\lambda \to \infty$. U tipičnom grafu $\mathbb{G}_{n,p}$ ne postoje kompleksne komponente, a za veličinu komponenata vrijedi

$$
\begin{aligned}
&L_1 = \theta(\epsilon^{-2} \ln \lambda) = \Theta(n^{2/3}\lambda^2 \ln \lambda),\\
&L_k \sim L_1 \text{ za svaki fiksni } k \in \mathbb{N}.
\end{aligned}
$$

Kritičan prozor

Kažemo da se $\mathbb{G}_{n,p}$ nalazi u subkritičnom režimu kada $p = \frac{1}{n} - \lambda n^{-4/3}$, a $\lambda \in \mathbb{R}$. Postoji $k \in \mathbb{N}$ najvećih komponenata i sve imaju veličinu $\Theta(n^{2/3})$. Može se odrediti zajednička distribucija konkretnih veličina i kompleksnosti tih komponenata. Ta distribucija je nedegenerirana, što znači da ne možemo konkretno odrediti veličinu i kompleksnost k najvećih komponenti, što je razlika s obzirom na ostale režime.

Jedva superkritičan režim

Kažemo da se $\mathbb{G}_{n,p}$ nalazi u jedva superkritičnom režimu kada $p = \frac{1+\epsilon}{n} = \frac{1}{n} + \lambda n^{-4/3}$ i to tako da $\epsilon \to 0$, $\lambda \to \infty$. U tipičnom grafu $\mathbb{G}_{n,p}$ postoji jedna divovska komponenta veličine $L_1 \sim 2\epsilon n = 2\lambda n^{2/3}$ koja ima po volji veliku kompleksnost, dok su ostale komponente jednostavne. Druga najveća komponenta tipično je veličine $L_2 \Theta(\epsilon^{-2} \ln(\lambda)) = \Theta(n^{2/3}\lambda^{-2} \ln(\lambda))$

3.2 Veličine najveće komponente u subkritičnom i superkritičnom režimu

Teorem 3.2.1. *Neka je $p = c/n$, gdje je $c \in \langle 0, 1\rangle$ konstanta. Onda u $\mathbb{G}_{n,p}$ s visokom vjerojatnošću vrijedi*

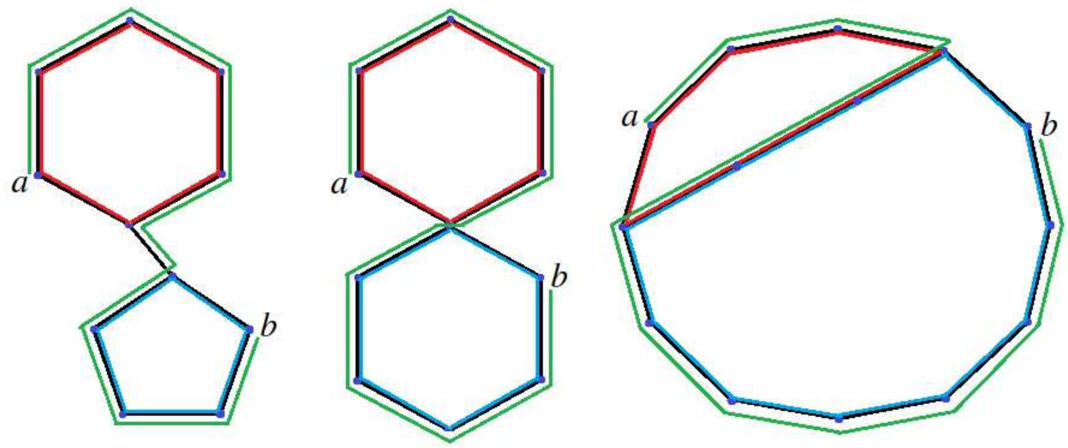
$$L_1 \in O(\ln n).$$

Ovaj rezultat slijedi iz tri leme. Prva lema govori da su sve komponente u subkritičnom režimu jednostavne. Druga govori da je od tih jednostavnih komponenti dovoljno malo vrhova u unicikličnim komponentama da niti jedna takva nije najveća. Zaključno, treća lema govori o veličini najvećeg stabla.

Lema 3.2.2. *Neka je $p < \frac{1}{n} - \frac{\omega}{n^{4/3}}$, gdje je $\omega = \omega(n) \to \infty$. Onda u $\mathbb{G}_{n,p}$ s visokom vjerojatnošću vrijedi da svaka komponenta sadrži najviše jedan ciklus.*

Dokaz. Ograničimo prvo broj mogućih realizacija dva povezana ciklusa. Neka je K_k graf na k vrhova koji sadrži dva ciklusa, povezan je i ne sadrži dodatne bridove. Na Slici 3.1 prikazana su tri moguća slučaja izgleda grafa K_k. Prvi slučaj je kada ciklusi nemaju niti jedan vrh zajednički, drugi kada imaju točno jedan zajednički vrh i treći je kada imaju više od jednog zajedničkog brida.

Zovimo cikluse A i B, redom crveni i plavi na Slici 3.1. Vidimo da u svim slučajevima možemo odabrati vrh $a \in A \setminus B$ koji je susjed u smjeru kazaljke na satu nekog vrha stupnja barem 3 i $b \in B \setminus A$ koji je susjed u smjeru kazaljke na satu nekog vrha stupnja 3. Tada bridovi koji nisu oni koji spajaju a i b s vrhovima stupnja tri tvore put kroz sve vrhove grafa K (taj put je označen zeleno na Slici 3.1). Također, graf definira izbor još jednog susjeda za a i još jednog za b. To jest, izbor permutacije od $k!$ permutacija i dva izbora vrha koja se mogu napraviti na manje od k^2 načina definiraju svaki i točno jedan graf K.



Slika 3.1: Tri mogućnosti za dva povezana ciklusa

Graf K_k se tako na k odabranih vrhova može realizirati na maksimalno $k! \cdot k^2$ različitih načina. Svaki od njih se događa kada postoji $k+1$ bridova iz K_k, s vjerojatnošću p^{k+1}.

Neka je D broj pojavljivanja svih grafova K_k ukupno. Tada računamo

$$\mathbb{E}D = \sum_{k=4}^{n} \binom{n}{k} k! k^2 p^{k+1} \leq \sum_{k=4}^{n} n^k \cdot k^2 \left[\frac{1}{n} \left(1 - \frac{\omega}{n^{1/3}} \right) \right]^{k+1} \leq \sum_{k=4}^{n} \frac{1}{n} \cdot k^2 \left(1 - \frac{\omega}{n^{1/3}} \right)^k$$
$$\leq \sum_{k=4}^{n} \frac{1}{n} \cdot k^2 \exp\left(-\frac{k\omega}{n^{1/3}} \right) \leq \int_0^\infty \frac{x^2}{n} \exp\left(-\frac{x\omega}{n^{1/3}} \right) = \frac{2}{\omega^3} \in o(1).$$

Sada po metodi prvog momenta (1.15) s visokom vjerojatnošću vrijedi $D = 0$. Uvijek kada $D = 0$ nijedna komponenta nema dva ciklusa. Time je tvrdnja teorema dokazana.

□

Primijetimo da ova lema vrijedi i u jedva subkritičnom režimu što znači da smo i dokazali da i u tom režimu postoje samo jednostavne komponente.

Lema 3.2.3. *Neka je $p = \frac{c}{n}$, gdje je $c < 1$ konstanta i $\omega = \omega(n)$, $\omega \to \infty$. Onda u $\mathbb{G}_{n,p}$ s visokom vjerojatnošću vrijedi da je broj vrhova u unicikličnim komponentama $O(\omega)$.*

Dokaz. Neka je X_k broj unicikličnih komponenti s k vrhova. Broj mogućih takvih komponenti ograničen je brojem mogućih izbora k vrhova, nekog stabla na tih k vrhova te dva vrha koji su povezani dodatnim bridom. Podsjetimo se da je po Cayleyevoj formuli broj mogućih stabla na k vrhova k^{k-2}. Dakle,

$$|X_k| \leq \binom{n}{k} k^{k-2} \binom{k}{2}.$$

Događaj da se neka uniciklična komponenta realizirala sastoji se od toga da se realiziralo njenih k bridova te da se nije realizirao niti jedan od $\binom{k}{2} - k$ preostalih bridova među tih k vrhova, kao ni $k(n-k)$ bridova prema ostalim vrhovima. Po linearnosti matematičkog očekivanja vrijedi

$$\mathbb{E}X_k \leq \binom{n}{k} k^{k-2} \binom{k}{2} p^k (1-p)^{k(n-k)+\binom{k}{2}-k}. \tag{3.1}$$

Koristeći se sljedećim formulama tehničke prirode (1.2), (1.4) i (1.5) sada možemo poboljšati ogradu u (3.1) kako slijedi

$$\begin{aligned}
\mathbb{E}X_k &\leq \frac{n^k}{k!} e^{-\frac{k(k-1)}{2n}} k^{k-2} \binom{k}{2} p^k (1-p)^{k(n-k)+\binom{k}{2}-k} \\
&\leq \frac{e^k}{k^k} k^{k-2} \binom{k}{2} c^k \exp\left[-\frac{k(k-1)}{2n} - \frac{c}{n}\left(k(n-k) + \binom{k}{2} - k\right)\right] \\
&\leq \frac{1}{2} c^k \exp\left[(1-c)\left(k - \frac{k(k-1)}{2n}\right) + \frac{2kc}{n}\right] \\
&\leq \frac{1}{2} (ce^{1-c})^k e^{2c} \exp\left[(1-c)\left(-\frac{k(k-1)}{2n}\right)\right].
\end{aligned} \tag{*}$$

Za $c < 1$ zadnji faktor je manji od jedan pa vrijedi

$$\mathbb{E}X_k \leq \frac{1}{2} (ce^{1-c})^k e^{2c}.$$

Neka je X broj vrhova na unicikličnim komponentama tada

$$\mathbb{E}X = \sum_{k=3}^{n} k \mathbb{E}X_k \leq \sum_{k=3}^{n} \frac{1}{2} k (ce^{1-c})^k e^{2c} \in O(1).$$

Po Markovljevoj nejednakosti za $\omega \to \infty$ vrijedi

$$\mathbb{P}(X \geq \omega) \in O\left(\frac{1}{\omega}\right) \subset o(1).$$

□

Lema 3.2.4. *Neka je $p = \frac{c}{n}$, gdje je $c \neq 1$ konstanta, $\alpha = \alpha(c) := c - 1 - \ln c$ i $\omega = \omega(n)$, $\omega \to \infty, \omega \in o(\ln \ln n)$.*

(i) U $\mathbb{G}_{n,p}$ s visokom vjerojatnošću postoji izolirano stablo veličine k_- gdje je

$$k_- := \frac{1}{\alpha}\left[\ln n - \frac{5}{2}\ln \ln n\right] - \omega$$

(ii) U $\mathbb{G}_{n,p}$ s visokom vjerojatnošću ne postoji izolirano stablo veličine k_+ ili veće gdje je

$$k_+ := \frac{1}{\alpha}\left[\ln n - \frac{5}{2}\ln \ln n\right] + \omega$$

Dokaz. Primijetimo prvo da je

$$\alpha = -\ln(ce^{1-c}).$$

Funkcija $f(x) = xe^{1-x}$ za $x \neq 1$ je manja od jedan, dakle pod uvjetima teorema vrijedi $\alpha > 0$.

Neka je X_k broj komponenata koje su stabla veličine k. Znamo da je

$$X_k = \sum_{i \in I} \mathbb{1}_{B_i},$$

gdje je I skup mogućih stabala na skupu vrhova $[n]$, a B_i događaj da se realiziralo stablo $i \in I$ kao komponenta. Postoji $\binom{n}{k}k^{k-2}$ mogućih stabala na $[n]$, te se svako od njih događa ako postoji $k-1$ brid za to stablo te ne postoji niti jedan od preostalih $\binom{k}{2} - k + 1$ bridova između tih k vrhova, kao ni $k(n-k)$ bridova između tih vrhova i preostalih vrhova. Dakle, po linearnosti matematičkog očekivanja vrijedi

$$\mathbb{E}X_k = \binom{n}{k}k^{k-2}p^{k-1}(1-p)^{k(n-k)+\binom{k-1}{2}-k+1}. \tag{3.2}$$

Za dokaz tvrdnje (i) neka je $k \in O(\ln n)$, što vrijedi za k_-. Tada vrijede sljedeće lako dokazive tvrdnje:

$$\begin{aligned}
\left(1 - \frac{c}{n}\right)^{k(n-k)+\binom{k-1}{2}-k+1} &= e^{-ck}(1 + o(1)), \\
n(n-1)\cdot(n-k+1) &= n^k(1 + o(1)).
\end{aligned}$$

Koristeći ove tvrdnje i Stirlingovu formulu (1.1), jednakost (3.2) može se asimptotski napisati kao

$$\begin{aligned}
\mathbb{E}X_k &= \frac{1 + o(1)}{\sqrt{2\pi}} \cdot \frac{n}{ck^{5/2}}(ce^{1-c})^k \\
&= \frac{1 + o(1)}{\sqrt{2\pi}} \cdot \frac{n}{ck^{5/2}}e^{-\alpha k}, \qquad k \in O(\ln n).
\end{aligned}$$

Uvrstimo li $k = k_- \in O(\ln n)$, dobivamo

$$\mathbb{E}X_k = \frac{1 + o(1)}{\sqrt{2\pi}} \cdot \frac{n}{ck^{5/2}} \cdot \frac{e^{\omega\alpha}(\ln n)^{5/2}}{n} \geq Ae^{\omega\alpha},$$

za neku konstantu $A > 0$. Dodatno su $\omega \to \infty$ i $\alpha > 0$ pa vrijedi

$$\mathbb{E}X_{k_-} \to \infty.$$

Sada koristimo metodu drugog momenta. Varijabla $X_{k_-} = \sum_{i \in I} \mathbb{1}_{B_i}$ je suma simetričnih indikatorskih varijabli. Za ograničavanje varijance ograničavamo prvo $\mathbb{E}[X_k^2]$ kako slijedi

$$\begin{aligned}
\mathbb{E}[X_k^2] &= \mathbb{E}\left[\left(\sum_{i \in I} \mathbb{1}_{B_i}\right)^2\right] \\
&= \mathbb{E}\left[\sum_{i,j \in I} \mathbb{1}_{B_i} \cdot \mathbb{1}_{B_j}\right] \\
&= \sum_{i,j \in I} \mathbb{P}\left[B_i \cap B_j\right] \\
&= \sum_{i,j \in I} \mathbb{P}\left[B_j | B_i\right] \mathbb{P}\left[B_i\right] && \text{(neka je } i_0 \in I \text{ proizvoljan)} \\
&= \mathbb{E}X_k \sum_{j \in I} \mathbb{P}\left[B_j | B_{i_0}\right].
\end{aligned}$$

Simetričnost je bila ključna za zadnji redak, da se može zaključiti da je suma po j jednaka za svaki $i \in I$.

Događaj B_i je da se realiziralo izolirano stablo i. Stoga, ako $i = j$, $P(B_j|B_1) = 1$; ako stabla $i \neq j$ imaju zajedničkih vrhova, $\mathbb{P}(B_j|B_i) = 0$, a ako i i j nemaju zajedničkih vrhova i postoji izolirano stablo i, to mijenja $P(B_j)$ jedino utoliko što uvjet daje da ne postoje bridovi između vrhova od i i vrhova od j, odnosno

$$\mathbb{P}(B_j|B_i) = (1-p)^{-k^2}\mathbb{P}(B_j).$$

Uvrstimo li te uvjetne vjerojatnosti u dobiveni izraz za $\mathbb{E}[X_k^2]$ možemo odrediti

$$\mathbb{E}[X_k^2] \leq \mathbb{E}X_k \left[1 + (1-p)^{-k^2} \sum_{j \in I} \mathbb{P}(B_j)\right] = \mathbb{E}X_k \left[1 + (1-p)^{-k^2}\mathbb{E}X_k\right]. \tag{3.3}$$

Kada je $k = k_- \in O(\ln n)$ vrijedi $\mathbb{E}X_{k_-} \to \infty$, i po prethodnom,

$$\mathrm{Var}(X_{k_-}) = \mathbb{E}[X_{k_-}^2] - (\mathbb{E}X_{k_-})^2 \leq \mathbb{E}X_{k_-} + (\mathbb{E}X_{k_-})^2[(1-p)^{-k_-^2} - 1] \in o(\mathbb{E}X_{k_-}^2).$$

Po Korolaru 1.2.2, $\mathbb{P}(X_{k_-} = 0) \to 0$ odnosno u $\mathbb{G}_{n,p}$ s visokom vjerojatnošću postoji izolirano stablo veličine k_-.

Prijeđimo sada na dokaz tvrdnje (*ii*). Uvrštavanjem nejednakosti (1.5) i (1.2) u (3.2) dobivamo

$$\begin{aligned}\mathbb{E}X_k &\leq \frac{B}{\sqrt{k}}\left(\frac{ne}{k}\right)^k k^{k-2}\left(1-\frac{k}{2n}\right)^{k-1}\left(\frac{c}{n}\right)^{k-1} c^{-ck+\frac{ck^2}{2n}} \\ &\leq \frac{Bn}{\hat{c}_k k^{5/2}}\left(\hat{c}_k e^{1-\hat{c}_k}\right)^k,\end{aligned}$$

gdje je B neka konstanta i $\hat{c}_k := c(1-\frac{k}{2n})$.

U slučaju kada je $c < 1$ vrijedi $\hat{c}_k e^{1-\hat{c}_k} \leq ce^{1-c}$ i $\hat{c}_k \geq 2c$ te iz prošle nejednakosti slijedi

$$\begin{aligned}\sum_{k=k_+}^{n} \mathbb{E}X_k &\leq \frac{2Bn}{c}\sum_{k=k_+}^{n}\frac{\left(ce^{1-c}\right)^k}{k^{5/2}} \\ &\leq \frac{2Bn}{ck_+^{5/2}}\sum_{k=k_+}^{n} e^{-\alpha k} \\ &= \frac{2Bne^{-\alpha k_+}}{ck_+^{5/2}(1-e^{-\alpha})} \\ &= \frac{(2B+o(1))\alpha^{5/2}e^{-\alpha\omega}}{c(1-e^{-\alpha})} \in o(1).\end{aligned}$$

U slučaju kada je $c > 1$ za $k \leq \frac{n}{\ln n}$ vrijedi

$$\hat{c}_k e^{1-\hat{c}_k} = c\left(1-\frac{k}{2n}\right)e^{1-c\left(1-\frac{k}{2n}\right)} = e^{-\alpha-O(1/\ln n)},$$

$$n\sum_{k=k_+}^{n/\ln n}\left(e^{-\alpha-O(1/\ln n)}\right)^k \leq n^2\cdot\frac{e^{-(\alpha+O(1/\ln n))k_+}}{1-e^{-\alpha-O(1/\ln n)}} \in o(1),$$

a za $k > \frac{n}{\ln n}$ vrijedi $\hat{c}_k e^{1-\hat{c}_k} \leq 1$ i

$$n\sum_{k=n/\ln n}^{n} k^{-5/2} \leq n^2\cdot\left(\frac{\ln n}{n}\right)^{5/2} = \sqrt{\frac{(\ln n)^5}{n}} \in o(1).$$

Stoga

$$\mathbb{E}X_k \leq \frac{2An}{ck_+^{5/2}}\sum_{k=k_+}^{n/\ln n}\left(e^{-\alpha-O(1/\ln n)}\right)^k + \frac{2An}{c}\sum_{n/\ln n}^{n} k^{-5/2} \in o(1).$$

I u slučaju $c > 1$ i u slučaju $c < 1$ vrijedi $\sum_{k=k_+}^{n} \mathbb{E}X_k \in o(1)$ pa po metodi prvog momenta (1.14) vrijedi da je $\sum_{k=k_+}^{n} X_k$, broj izoliranih stabla s k_+ i više vrhova, s visokom vjerojatnošću 0. □

Kroz ove tri leme dokazan je Teorem 3.2.1 i mnogo više. Prvo je u Lemi 3.2.2 dokazano da tipičan graf sadrži samo stabla i unicikličtne komponente. Lema 3.2.3 pokazuje da je broj vrhova u unicikličnim komponentama po volji malen, iz čega slijedi da najveća komponenta ne može biti uniciklična. Lema 3.2.4 govori precizno kojeg reda je najveće izolirano stablo koje ćemo pronaći, ono je s visokom vjerojatnošću reda između k_- i k_+, dakle $O(\ln n)$.

Teorem 3.2.5. *Ako je $p = \frac{c}{n}$, $c > 1$ konstanta, onda $\mathbb{G}_{n,p}$ s visokom vjerojatnošću sadrži jedinstvenu divovsku komponentu u kojoj je $(1-\frac{x}{c}+o(1))n$ vrhova, gdje je $x \in \langle 0, 1\rangle$ rješenje jednadžbe $xe^x = ce^c$.*

Može se pokazati da se za svaki $c > 0$, $c \neq 1$ vrijednost $x = x(c)$ definirana kao jedinstveno rješenje jednadžbe $xe^x = ce^c$ u $\langle 0, 1\rangle$ može razviti u red. Vrijedi

$$x = \sum_{k=1}^{\infty} \frac{k^{k-1}}{k!}(ce^{-c})^k.$$

Dokaz. Neka je $p = \frac{c}{n}$, $c > 1$ konstanta. Neka su X_k broj komponenata koje su stabla reda k, Y_k broj komponenata koje nisu stabla veličine k te Z_k ukupni broj komponenata veličine k u $\mathbb{G}_{n,p}$.

Dokažimo prvo da postoje $\beta_0, \beta_1 > 0$ takvi da ne postoji komponenta veličine k za $k \in [\beta_0 \ln n, \beta_1 n]$. Ograničimo li broj mogućih realizacija komponenata veličine k mogućim brojem realizacija stabla veličine k, po linearnosti matematičkog očekivanja vrijedi

$$\begin{aligned}
\mathbb{E}Z_k &\leq \binom{n}{k} k^{k-2} p^{k-1} (1-p)^{k(n-k)} && \text{(uvrštavanje (1.3) i } p = \frac{c}{n}) \\
&\leq \frac{ne^k}{k^2} c^{k-1} e^{-\frac{c}{n}k(n-k)} \\
&= \frac{n}{k^2}\left(ce^{1-c+c\frac{k}{n}}\right)^k .
\end{aligned}$$

Budući da je $c > 1$, postoji $\beta_1 > 0$ dovoljno malen da vrijedi

$$ce^{1-c+c\beta_1} < 1,$$

i postoji $\beta_0 > 0$ dovoljno velik da vrijedi

$$\left(ce^{1-c+o(1)}\right)^{\beta_0 \ln n} \leq \frac{1}{n^2}.$$

Sada za $k \in [\beta_0 \ln n, \beta_1 n]$ vrijedi

$$k \in o(n) \overset{k>\beta_0 \ln n}{\Rightarrow} \mathbb{E}Z_k \leq \frac{n}{k^2} \cdot \frac{1}{n^2} \to 0,$$
$$k \in [\alpha n, \beta_1 n] \Rightarrow \varphi := ce^{1-c+c\frac{k}{n}} < 1 \Rightarrow \mathbb{E}Z_k \leq \frac{n}{k^2} \cdot \varphi^{\alpha n} \to 0.$$

Po metodi prvog momenta (1.14) vrijedi da s visokom vjerojatnošću ne postoje komponente veličine između $\beta_0 \ln n$ i $\beta_1 n$.

Pokažimo sada da je broj vrhova u komponentama s manje od $\beta_0 \ln n$ vrhova $\frac{x}{c}n + o(n)$.

Računamo prvo očekivanje broja vrhova u komponentama koje nisu stabla, a imaju manje od $\beta_0 \ln n$ vrhova. Istim načinom kao u računu $\mathbb{E}Z_k$:

$$\begin{aligned}\mathbb{E}\left[\sum_{k=1}^{\beta_0 \ln n} kY_k\right] &\leq \sum_{k=1}^{\beta_0 \ln n} k\binom{n}{k}k^{k-2}\binom{k}{2}p^k(1-p)^{k(n-k)} \\ &\leq \sum_{k=1}^{\beta_0 \ln n} k\left(ce^{1-c+ck/n}\right)^k \in O(1).\end{aligned}$$

Po Markovljevoj nejednakosti (1.13) je s visokom vjerojatnošću broj vrhova u komponentama s manje od $\beta_0 \ln n$ vrhova koje nisu stabla [1]

$$\sum_{k=1}^{\beta_0 \ln n} kY_k \in o(n).$$

Na sličan način sada pokazujemo da je broj vrhova u komponentama koje su stabla reda između $k_0 = \frac{1}{2\alpha} \ln n$ i $\beta_0 \ln n$ također $o(n)$. Ovdje je $\alpha = c - 1 - \ln c = -\ln(ce^{1-c})$ kao u prethodnoj lemi.

$$\begin{aligned}\mathbb{E}\left[\sum_{k=\frac{1}{2\alpha}\ln n}^{\beta_0 \ln n} kX_k\right] &\leq \sum_{k=k_0}^{\beta_0 \ln n} k\binom{n}{k}k^{k-2}p^{k-1}(1-p)^{k(n-k)} \\ &\leq \frac{n}{c}\sum_{k=k_0}^{\beta_0 \ln n}\left(ce^{1-c+ck/n}\right)^k \\ &\leq O\left[n\left(ce^{1-c+ck_0/n}\right)^{k_0}\right] \\ &\leq O\left[ne^{-\alpha k_0}e^{o(1)}\right] \\ &\leq O\left[n^{1/2+o(1)}\right].\end{aligned}$$

[1] Vrijedi da za svaki $\epsilon > 0$

$$\lim_{n\to\infty} \mathbb{P}\left(\sum_{k=1}^{\beta_0 \ln n} kY_k \geq \epsilon n\right) = 0$$

Po Markovljevoj nejednakosti (1.13) s visokom vjerojatnošću vrijedi

$$\sum_{k=\frac{1}{2\alpha}\ln n}^{\beta_0 \ln n} kX_k \in o(n). \tag{3.4}$$

Dokazujemo još da je broj vrhova u komponentama koje su stabla veličine $k < k_0$ s visokom vjerojatnošću $\sim xn/c$. Računamo očekivanje kao u Lemi 3.2.4 te uvrštavamo (3.4)

$$\begin{aligned}\mathbb{E}\left[\sum_{k=1}^{k_0-1} kX_k\right] &= \sum_{k=1}^{k_0-1} k\frac{n^k(1+o(1))}{k!}k^{k-2}\left(\frac{c}{n}\right)^{k-1} e^{-ck}(1+o(1))\\ &\sim \frac{n}{c}\sum_{k=1}^{k_0-1}\frac{k^{k-1}}{k!}\left(ce^{-c}\right)^k \sim \frac{nx}{c}.\end{aligned}$$

Prisjetimo se sada da je iz (3.3) varijanca X_k ograničena s

$$\begin{aligned}\mathrm{Var}X_k &= \mathbb{E}[X_k^2] - (\mathbb{E}X_k)^2\\ &\le \mathbb{E}X_k + (\mathbb{E}X_k)^2\left[(1-p)^{-k^2} - 1\right]\\ &\le \mathbb{E}X_k + \frac{2ck^2}{n}(\mathbb{E}X_k)^2.\end{aligned}$$

Iz Čebiševljeve nejednakosti (Teorem 1.2.1) slijedi

$$\mathbb{P}(|X_k - \mathbb{E}X_k| \ge \epsilon\mathbb{E}X_k) \le \frac{\mathrm{Var}X_k}{\epsilon^2(\mathbb{E}X_k)^2} \le \frac{1}{\epsilon^2\mathbb{E}X_k} + \frac{2ck^2}{\epsilon^2 n}. \tag{3.5}$$

U Lemi 3.2.4 pokazano je da za $k \in O(\ln n)$ vrijedi

$$\mathbb{E}X_k = \frac{1+o(1)}{\sqrt{2\pi}}\frac{n}{ck^{5/2}}e^{-\alpha k},$$

što je padajuća funkcija u ovisnosti o veličini komponenti k, stoga za $k \le k_0$

$$\mathbb{E}X_k \ge \mathbb{E}X_{k_0} = \frac{1+o(1)}{\sqrt{2\pi}}\frac{n}{c(\ln n/2\alpha)^{5/2}}e^{-\frac{1}{2}\ln n} = n^{1/2+o(1)}.$$

Koristeći ovu donju ogradu na očekivanje i (3.5) za $\epsilon = (\ln n)^{-1}$

$$\mathbb{P}(|X_k - \mathbb{E}X_k| \ge \epsilon\mathbb{E}X_k) \le \frac{(\ln n)^2}{n^{1/2+o(1)}} + O\left(\frac{(\ln n)^4}{n}\right),\ \forall k \le k_0.$$

Sada je

$$\mathbb{P}\left(\exists k \text{ takav da } |kX_k - \mathbb{E}(kX_k)| \geq \epsilon \mathbb{E}(kX_k)\right) \leq \sum_{k=1}^{k_0-1} \left[\frac{(\ln n)^2}{n^{1/2+o(1)}} + O\left(\frac{(\ln n)^4}{n}\right) \right] \in o(1),$$

odnosno vrijedi s visokom vjerojatnošću

$$\sum_{k=1}^{k_0-1} kX_k \sim \mathbb{E}\left[\sum_{k=1}^{k_0-1} kX_k\right] \sim \frac{nx}{c}.$$

Broj vrhova u komponentama manjima od $\beta_0 \ln n$ je

$$\sum_{k=1}^{\beta_0 \ln n} (kX_k + kY_k) \sim \frac{nx}{c} + o(n) + o(n) = \left(\frac{x}{c} + o(1)\right) n.$$

Dokažimo sada jedinstvenost divovske komponente u $\mathbb{G}_{n,p}$. Definirajmo

$$c_1 := c - \frac{\ln n}{n}, \quad p_1 := \frac{c_1}{n}$$

te $p_2 \in \langle 0, 1 \rangle$ takav da

$$1 - p = (1 - p_1)(1 - p_2).$$

Tvrdio da je slučajni graf koji je dobiven kao unija (po bridovima) nezavisnih grafova $\mathbb{G}_{n,p_1}$ i $\mathbb{G}_{n,p_2}$ jednak slučajnom elementu $\mathbb{G}_{n,p}$. Naime, u takvoj uniji brid ne bi postojao sa vjerojatnošću $(1 - p_1)(1 - p_2) = 1 - p$, te su ti događaji za različite bridove nezavisni jer odgovaraju različitim bridovima u $\mathbb{G}_{n,p_1}$ i $\mathbb{G}_{n,p_2}$. Pišemo

$$\mathbb{G}_{n,p} = \mathbb{G}_{n,p_1} \cup \mathbb{G}_{n,p_2}.$$

Nadalje vrijedi

$$p_2 \geq \frac{\ln n}{n^2}, \qquad x(c) \sim x(c_1), \qquad \beta_1(c) \sim \beta_1(c_1).$$

Neka je $\mathbb{G}_{n,p_1}$ slučajan graf i $\mathbb{G}_{n,p_2}$ s njime nezavisan slučajan graf. Pretpostavimo da $\mathbb{G}_{n,p_1}$ ima $l \in \mathbb{N}$ velikih komponenti $G_1, \ldots G_l$, veličine veće od $\beta_1(c)n$. Po prethodno dokazanom vrijedi $\sum_{i=1}^{l} |G_i| = (1 - \frac{x(c_1)}{c} + o(1))n$. Svaki G_i, $i = 1, 2, \ldots l$ ima barem $\beta_1(c_1)n$ vrhova stoga je vjerojatnost da za neke $i \neq j \in [l]$, G_i i G_j nisu povezani niti sa jednim bridom u $\mathbb{G}_{n,p_2}$ je

$$\begin{aligned}
\mathbb{P}(G_i \text{ i } G_j \text{ nisu povezani u } \mathbb{G}_{n,p_2}) &\leq (1 - p_2)^{\beta_1(c_1)^2 n^2} \\
&\leq \left(1 - \frac{\ln n}{n^2}\right)^{\beta_1(c_1)^2 n^2} \\
&\leq e^{-\beta_1(c_1) \ln n}.
\end{aligned}$$

Dakle, vrijedi

$$\mathbb{P}(\exists i, j \in [l];\ G_i \text{ i } G_j \text{ nisu povezani u } \mathbb{G}_{n,p_2}) \leq \binom{j}{2} e^{-\beta_1(c_1)\ln n} \overset{n\to\infty}{\longrightarrow} 0.$$

Dakle, s visokom vjerojatnošću su svi G_i, $i = 1, 2, \ldots, l$ povezani u $\mathbb{G}_{n,p} = \mathbb{G}_{n,p_1} \cup \mathbb{G}_{n,p_2}$. Oni čine jednu komponentu veličine $(1 - \frac{x(c_1)}{c_1} + o(1))n \sim (1 - \frac{x}{c} + o(1))n$, što znači da je to jedina komponenta u $\mathbb{G}_{n,p}$ sa više ili jednako $\beta_1 n$ vrhova. □

3.3 Jednostavniji dokaz pomoću pretraživanja u dubinu

U članku [7] iz 2012. godine dokazuju se rezultati da u vrlo subkritičnom režimu ne postoji komponenta veća od $O(\ln n)$, a da u vrlo superkritičnom režimu postoji komponenta veličine $O(n)$. Te dokaze ćemo prikazati u ovom potpoglavlju.

Pretpostavimo da je zadan fiksan graf G sa skupom vrhova V koji su svi usporedivi. Pretraživanje u dubinu ili DFS (*depth-first search*) je dobro poznati algoritam kojim se može otkriti struktura grafa G. Provodi se tako da počevši od jednog vrha krenemo niz neki put do kraja, kada nema više neistraženih vrhova koji su susjedi zadnjeg istraživanog vrha. Kada se do takvog kraja dolazi, taj zadnji vrh se sprema u istražene vrhove te se vraćamo za jedan vrh na putu istražujući opet u dubinu.

Definirajmo algoritam. Varijable su: S = skup istraženih vrhova grafa G, U = skup vrhova grafa G koji se trenutno istražuju, T = skup neistraženih vrhova grafa G. Skup U ponaša se kao stog (*stack*) to jest pristupa se samo zadnjem dodanom elementu. Po definiciji, ova tri skupa uvijek oblikuju particiju skupa vrhova V. Na početku algoritma postavlja se $S = U = \emptyset$, $T = V$.

Korak algoritma je sljedeći

- ako $U = T = \emptyset$ algoritam staje,
- ako $U = \emptyset, T \neq \emptyset$ u U se dodaje najmanji element iz T,
- inače se za zadnji element u iz U traži najmanji susjedni element v iz T
  - ako v postoji u T, vrh v se briše iz T i dodaje na kraj stoga U,
  - ako ne postoji niti jedan susjed $v \in T$ od u element u se briše iz U i dodaje u S.

Po završetku ovih koraka, kada je $T = \emptyset$ otkrivene su sve komponente u grafu. Algoritam završava tako da provjerava postoje li svi bridovi koji još nisu ispitani. Time je u nekom trenutku provjereno svih $N := \binom{n}{2}$ bridova između n vrhova.

Nekoliko činjenica o ovom algoritmu biti će bitno za dokaze u nastavku. U ovom algoritmu struktura grafa se otkriva komponentu po komponentu, u svakom trenutku su svi vrhovi u U u istoj komponenti, štoviše tvore put. Kada je $U = \emptyset$ to znači da su otkriveni svi vrhovi u nekoj komponenti. Kada novi vrh dolazi u prazan U počinje se otkrivati nova komponenta. Vrijeme između uzastopnih ispražnjavanja U zovemo epoha. Svi osim prvog vrha svake komponente stavljeni su u U jer je postojao neki brid. Također u svakom trenutku, provjereno, ne postoje bridovi između vrhova u T i vrhova u S.

Slučajan graf $\mathbb{G}_{n,p}$ se može poistovjetiti sa skupom N nezavisnih Bernoullijevih varijabli s parametrom p na način da svaka varijabla odgovara jednom bridu te ako ona poprimi vrijednost 1 brid postoji, a ako poprimi vrijednost 0 odgovarajući brid ne postoji. Poanta ovog pristupa je da se $\mathbb{G}_{n,p}$ poistovjeti s nizom nezavisnih Bernoullijevih varijabli $\tilde{X} = (X_i)_{i=1}^N$ gdje X_i karakterizira postojanje i-tog brida po redu kojem DFS algoritam treba pristupiti.

Za neku epohu neka je $I \subseteq [N]$ segment takav da su $(X_i)_{i\in I}$ bridovi provjeravani u toj epohi, s obzirom da je svaki vrh osim prvog u nekoj komponenti dodan u U jer postoji brid imamo

$$|U| \leq 1 + \sum_{i\in I} X_i \leq 1 + \sum_{i=1}^{t} X_i, \ \forall t \geq \max(I). \tag{3.6}$$

Nejednakost je najčešće stroga jer se vrhovi miču iz U u S. Za svaki $t \in [N]$ vrijedi

$$|S \cup U| \geq \sum_{i=1}^{t} X_i. \tag{3.7}$$

Za dokaze teorema biti će potrebna sljedeća tehnička lema. U njoj i nadalje su oznake najmanjeg, najvećeg i najbližeg cijelog izostavljene zbog čitkosti.

Lema 3.3.1. *Neka je $\epsilon > 0$ dovoljno mala konstanta i $\tilde{X} = (X_i; i \in [N])$ niz nezavisnih jednako distribuiranih Bernoullijevih varijabli s parametrom p.*

(i) Neka je $p = \frac{1-\epsilon}{n}$, a $k = \frac{7}{\epsilon^2}\ln n$. Tada sa velikom vjerojatnošću ne postoji interval I duljine kn u $[N]$ takav da za barem k indeksa $i \in I$ vrijedi $X_i = 1$.

(ii) Neka je $p = \frac{1+\epsilon}{n}$, a $N_0 = \frac{\epsilon n^2}{2}$. Tada sa velikom vjerojatnošću za svaki $n^{7/4} \leq t \leq N_0$

$$\Big|\sum_{i=1}^{t} X_i - (1+\epsilon)\frac{t}{n}.\Big| \leq n^{2/3}.$$

(iii) Neka je $p = \frac{1+\epsilon}{n}$. Tada sa velikom vjerojatnošću

$$\sum_{t=1}^{n^{7/4}} X_t \leq n^{5/6}.$$

Slijedi teorem za vrlo subkritičan režim.

Teorem 3.3.2. *Neka je $\epsilon > 0$ dovoljno mala konstanta. Tada za $p = \frac{1-\epsilon}{n}$ s visokom vjerojatnošću u $\mathbb{G}_{n,p}$ ne postoji komponenta veličine veće od $\frac{7}{\epsilon^2} \ln n$.*

Dokaz. Pretpostavimo suprotno. Neka je C komponenta sa više od $k := \frac{7}{\epsilon^2} \ln n$ elemenata. Neka je $\tilde{X} = (X_i; i \in [N])$ niz nezavisnih jednako distribuiranih Bernoullijevih varijabli s parametrom p koji prikazuju po redu bridove kojima pristupa DFS u pretraživanju grafa $\mathbb{G}_{n,p}$.

Neka je t_0 trenutak kada prvi vrh iz C bude premješten iz T u U, neka je trenutak t takav da $X_t = 1$ dodaje $k+1$. element komponente C iz T u U. Promotrimo segment varijabli $(X_i)_{t_0}^t$. S obzirom da je u U tokom cijelog tog razdoblja bilo samo k do sada otkrivenih vrhova iz C provjeravali su se samo bridovi između tih k vrhova i između tih k i preostalih $n-k$ vrhova. Dakle, $(X_i)_{t_0}^t$ ima maksimalno

$$t - t_0 \leq \binom{k}{2} + k(n-k) < nk$$

elemenata.

U $(X_i)_{t_0}^t$ je bilo više od k jedinica. Tako je zato što je svaki od $k+1$ elemenata komponente C koji su do sada pronađeni (premješteni u U i onda eventualno u S) pronađen zbog toga što su susjedni nekom vrhu u C to jest jer je neki X_i za $t_0 \leq i \leq t$ bio jednak jedan.

Dakle, $(X_i)_{t_0}^t$ bi bio segment u $\tilde{X}$ duljine manje od nk, a sa više od k jedinica što je u suprotnosti s Lemom 3.3.1. □

Teorem 3.3.3. *Neka je $\epsilon > 0$ po volji mala konstanta. Tada za $p = \frac{1+\epsilon}{n}$ s visokom vjerojatnošću u $\mathbb{G}_{n,p}$ postoji komponenta veličine barem $\frac{\epsilon n}{2}$.*

Dokaz. Neka je $N_0 = \frac{\epsilon n^2}{2}$ i neka je $\tilde{X} = (X_i; i \in [N])$ niz nezavisnih Bernoullijevih varijabli koji je po DFS pridružen $\mathbb{G}_{n,p}$. Komponenta koja će se otkrivati u trenutku N_0 imati će s visokom vjerojatnošću barem $\frac{\epsilon n}{2}$ elemenata.

Za $n^{7/4} \leq t \leq N_0$, po tvrdnji (*ii*) iz Leme 3.3.1 i (3.7) s visokom vjerojatnošću vrijedi

$$|S \cup U| \geq \frac{(1+\epsilon)t}{n} - n^{2/3}. \tag{3.8}$$

Pokažimo kao pomoćnu tvrdnju da je $|S| < \frac{n}{3}$ u trenutku N_0, a onda i za $t \leq N_0$. Kada bi $|S| \geq \frac{n}{3}$, postojao bi trenutak t kada je $|S| = \frac{n}{3}$. Za taj trenutak po svojstvu (*ii*) Leme 3.3.1 s visokom vjerojatnošću vrijedi $\sum^t X_i \leq \dfrac{n}{3}$. Uvrstimo li to u (3.6) vrijedi $|U| \leq 1 + \frac{n}{3}$. Tako imamo da $|T| = n - |U| - |S| \geq \frac{n}{3}$. Kako su uvijek provjereni svi bridovi između S i T vrijedi

$$t \geq |S| \cdot |T| \geq \frac{n}{3} \cdot \frac{n}{3} = \frac{n^2}{9}.$$

Ovo je za dovoljno male ϵ kontradikcija sa $t \leq N_0 = \frac{\epsilon n^2}{2}$. Dakle, $|S| < \frac{n}{3}$ u vremenima $t \leq N_0$.

Pretpostavimo da je U u nekom trenutku $n^{7/4} \leq t \leq N_0$ prazan. Provjereni su svi bridovi između S i T, dakle

$$t \geq |S| \cdot |T| = |S| \cdot (n - |S|).$$

Kako je $|S| < \frac{n}{3}$, zbog oblika funkcije $x \to x(n-x)$ i (3.8) uz $|U| = 0$

$$t \geq \left(\frac{(1+\epsilon)t}{n} - n^{2/3}\right)\left(n - \frac{(1+\epsilon)t}{n} + n^{2/3}\right).$$

Pojednostavljivanje vodi do

$$t \geq t + \epsilon t - \frac{\epsilon}{2}(1+\epsilon)^2 t - 2n^{5/3} > t + \frac{\epsilon}{3}t - 2n^{5/3}.$$

Zadnja nejednakost vrijedi za dovoljno male ϵ. Kako je $t \gg n^{5/3}$ desna strana je za dovoljno velike n veća od t. To je kontradikcija, dakle U nikada nije prazan za $n^{7/4} \leq t \leq N_0$.

Kako U nije prazan za $n^{7/4} \leq t \leq N_0$, svi vrhovi dodani u $|S \cup U|$ u tom intervalu su u istoj komponenti. Svaki dodani vrh odgovara jednom postojećem bridu odnosno po (3.7) ima ih najmanje

$$\sum_{i=n^{7/4}}^{N_0} X_i,$$

koristeći tvrdnje (*ii*) i (*iii*) iz Leme 3.3.1 s visokom vjerojatnošću

$$\sum_{i=n^{7/4}}^{N_0} X_i \geq \frac{(1+\epsilon)N_0}{n} - n^{2/3} - n^{5/6} \geq \frac{\epsilon n}{2} + \frac{\epsilon^2 n}{2} - n^{2/3} - n^{5/6} \overset{n\to\infty}{\geq} \frac{\epsilon n}{2}.$$

Dakle, vrhova u ovoj komponenti ima najmanje $\frac{\epsilon n}{2}$. □

Zaključak

Primijetimo da teoremi dokazani pretragom u dubinu već u iskazu govore mnogo manje nego teoremi ranijeg potpoglavlja. U vrlo subkritičnom režimu ne pokazuje se da doista postoji komponenta veličine $O(\ln n)$, niti se govori o kompleksnosti komponenata, što je dostupno u klasičnim dokazima. U vrlo superkritičnom režimu ne pokazuje se jedinstvenost, niti je red divovske komponente vrlo usko određen kao u Teoremu 3.2.5.

S druge strane dokazi pretragom u dubinu su kraći i jasniji. Povezuju slučajni graf $\mathbb{G}_{n,p}$ sa vrlo razumljivom strukturom niza N nezavisnih Bernoullijevih varijabli s parametrom p. Jedan od glavnih rezultata ovog pristupa je ipak sljedeća tvrdnja iz [7].

Teorem 3.3.4. *Neka je $\epsilon > 0$ dovoljno malen, i neka je $p = \dfrac{1+\epsilon}{n}$. Tada s visokom vjerojatnošću u $\mathbb{G}_{n,p}$ postoji put duljine $\frac{\epsilon^2 n}{5}$.*

Ovo se lako pokazuje iz činjenice da su vrhovi u U po konstrukciji uvijek povezani u put.

Ako usporedimo alate, teoremi iz Potpoglavlja 3.2 baziraju se na jednostavnim, ali ne elegantnim metodama prvog i drugog reda, Čebiševljevim i Markovljevim nejednakostima. S druge strane pristup metodom pretraživanja u dubinu koristi Lemu 3.3.1 kojoj su u pozadini grublje ograde iz Čebiševljeve nejednakosti, ali i Chernoffove ograde. Chernoffove ograde su malo jači alat jer daju eksponencijalnu ogradu da rep binomne distribucije.

Srodan pristup pristupu pretrage u dubinu je promatranje komponente stabla kao procesa grananja. Pristup pomoću procesa grananja mogao bi se smatrati pristupom pretrage u širinu. Ovo je jednako zanimljivo, a također daje i sljedeću intuiciju. U grafu $\mathbb{G}_{n,c/n}$ komponente se poistovjećuju sa procesom grananja gdje je broj potomaka nekog vrha Poissonova varijabla sa očekivanjem c. Za taj proces znamo da je konačan ako je $c < 1$. A ako je $c > 1$ vjerojatnost da je konačan je z koji zadovoljava jednadžbu

$$z = \sum_{i=k}^{\infty} \frac{e^{-c} c^k}{k!} z^k = e^{c(z-1)},$$

odnosno proces preživljava s vjerojatnošću $y = 1 - z$ koja zadovoljava jednadžbu $1 - y = e^{-cy}$.

Poglavlje 4

Razapinjuće strukture u slučajnom grafu

U ovome poglavlju bavimo se događajima u slučajnom grafu $\mathbb{G}_{n,p}$ koji se tiču svih n vrhova. Glavni objekt interesa su *savršena sparivanja*, skupovi koji se sastoje od $n/2$ međusobno disjunktnih bridova. Hamiltonov ciklus u grafu G može se promatrati kao podgraf tog grafa na istom skupu vrhova. Ukoliko graf ima paran broj vrhova, svaki drugi brid Hamiltonovog ciklusa tvori savršeno sparivanje. Također, ako u grafu postoji Hamiltonov ciklus, taj graf je povezan.

Povezanost slučajnog grafa bilo je jedno od prvih pitanja koje se proučavalo u području slučajnih grafova. Rezultata vezanih za to pitanje ima toliko mnogo da se čak u opsežnim djelima poput [2] navodi da je predstavljen samo mali dio rezultata.

Najznačajnija činjenica je da se i povezanost i savršeno sparivanje pojavljuju u istom trenutku kada nestaje zadnji izolirani vrh. Za graf $G = (V, E)$ definiramo najmanji stupanj grafa

$$\delta(G) := \min_{v \in V} \{k \in \mathbb{N}_0;\ v \text{ je povezan sa točno } k \text{ vrhova}\}.$$

U prvom potpoglavlju iskazuje se teorem o funkciji praga za svojstvo postojanja savršenog sparivanja. Izlaganje je kao u [5], dok se prvo pojavljivanje ovakvog dokaza pripisuje Bollobásu i i Friezeu u radu iz 1985. godine. Napomenimo ovdje da će teorem biti iskazan u većem opsegu nego što je dokazan, ali se u [5] navodi da se dokaz može prilagoditi da se dobije tvrdnja punog iskaza.

Drugo potpoglavlje je zaključno i tamo se iznosi i komentira još nekoliko važnih rezultata: o funkciji praga za svojstvo postojanja Hamiltonovog ciklusa i o funkciji praga za svojstvo povezanosti.

4.1 Teorem o savršenom sparivanju

Napomenimo prije teorema nekoliko detalja o izoliranim vrhovima u $\mathbb{G}_{n,p}$. Neka je $p = \frac{\ln n+\omega}{n}$, gdje je ω neka dovoljno sporo rastuća funkcija i neka je X_0 broj izoliranih vrhova u $\mathbb{G}_{n,p}$. Po linearnosti matematičkog očekivanja

$$\mathbb{E}X_0 = n(1-p)^{n-1} \sim n \cdot e^{-(\ln n+\omega)} = e^{-\omega} \in o(1).$$

Dapače, može se pokazati da ako promijenimo da $\omega \to c \in \mathbb{R}$ vrijedi

$$\mathbb{P}(\textit{Postoji izolirani vrh u } \mathbb{G}_{n,p}) \to e^{-e^{-c}}.$$

Kada $\omega \to -\infty$, s visokom vjerojatnošću u $\mathbb{G}_{n,p}$ ne postoji izolirani vrh.

Prijeđimo na definiranje pojmova vezanih za sparivanja. Za graf $G = (V, E)$ stupanj maksimalnog sparivanja je

$$\mu(G) = \max\{|M|; \textit{M je sparivanje u G}\}.$$

Sparivanje M za koje se ovaj maksimum postiže naziva se *maksimalno sparivanje*. Ako $v \in V$ nije incidentan niti sa jednim bridom u sparivanju M kažemo da *M izolira v*. Neka je $v \in V$ te neka je M maksimalno sparivanje koje izolira $v \in V$ tada definiramo

$$S_0(v, M) := \{u \in V;\ u \neq v,\ \textit{M izolira u}\}.$$

Sada neka je $u \in S_0(v, M)$ takav da postoji $e = (u, x) \in E$ i $f = (x, y) \in M$ tada za maksimalno sparivanje $M' = M - f + e$ kažemo da je dobiveno "flippingom" sparivanja M. U tom slučaju vrijedi $y \in S_0(v, M)$, a za brid f kažemo da je *izbačen*. Definiramo nadalje

$$A(v, M) = \bigcup_{M'} S_0(v, M'),$$

pri čemu je unija indeksirana po sparivanjima M' koja se mogu dobiti iz M kroz niz "flippinga". Definirajmo još oznaku za skup svih susjeda u grafu $G = (V, E)$ nekog skupa vrhova $A \subset V$,

$$N_G(A) := \{w \in V \setminus A;\ \exists v \in A \textit{ takav da } \{v, w\} \in E\}.$$

Sada smo spremni iskazati kombinatornu lemu koja će biti ključna u dokazu Teorema 4.1.3. Izjava i dokaz su slicni poznatom Hallovom teoremu o savrsenim sparivanjima u bipartitnim grafovima.

Lema 4.1.1. *Neka je G graf bez savršenog sparivanja, M maksimalno sparivanje te v vrh koji M izolira. Tada vrijedi*

$$|N_G(A(v, M))| < |A(v, M)|.$$

Od sada nadalje govori se kao u [5] o skupu $A(v)$, umjesto $A(v, M)$, podrazumijevajući da postoji neko maksimalno sparivanje M koje izolira v. Ako je $u \in A(v)$, onda je dobro definirano i $A(u)$. Također, vidimo da ako je v kao i do sada te je $u \in A(v)$

$$\mu(G + \{u, v\}) = \mu(G) + 1,$$

jer sparivanju M' koje izolira u možemo dodati brid $\{u, v\}$.

Glavni tehnički rezultat za dokaz Teorema 4.1.3 je sljedeća lema. Iznesena tvrdnja slabija je od odgovarajuće leme u [5] jer je tamo korištena i dokazivanju teorema o Hamiltonovom ciklusu

Lema 4.1.2. *Neka je $\omega \in \ln\ln n$, $\omega \to \infty$ te $p_1 := \frac{\ln n + \omega/2}{n}$. Tada s visokom vjerojatnošću za skup $S \subseteq [n]$ za koji vrijedi $|S| \leq \frac{n}{2e \cdot 5 \cdot 7 \cdot 100}$, vrijedi i $N_{\mathbb{G}_{n,p_1}}(S) \geq |S|$.*

Dokaz ove leme odvija se promatrajući posebno skupove vrhova sa velikim i sa malim stupnjem, sa granicom $\lambda = \ln n/100$.

Ovaj teorem o funkciji praga za postojanje savršenog sparivanja iskazan je i dokazan kao u [5].

Teorem 4.1.3. *Neka je $\omega = \omega(n)$ funkcija, $c > 0$ konstanta te $p = \dfrac{\ln n + \omega}{n}$. Tada vrijedi*

$$\lim_{\substack{n\to\infty \\ n \text{ paran}}} \mathbb{P}(U\ \mathbb{G}_{n,p} \text{ postoji savršeno sparivanje}) = \begin{cases} 0 & \text{za } \omega \to -\infty, \\ e^{-e^{-c}} & \text{za } \omega \to c, \\ 1 & \text{za } \omega \to \infty. \end{cases}$$

Također

$$\lim_{\substack{n\to\infty \\ n \text{ paran}}} \mathbb{P}(U\ \mathbb{G}_{n,p} \text{ postoji savršeno sparivanje}) = \lim_{\substack{n\to\infty \\ n \text{ paran}}} \mathbb{P}(\delta(\mathbb{G}_{n,p}) \geq 1).$$

Dokaz. Kada $\omega \to -\infty$ u $\mathbb{G}_{n,p}$ s visokom vjerojatnošću postoji izolirani vrh te zbog toga ne postoji savršeno sparivanje.

Neka je $\omega \to \infty$ te $\omega \in o(\ln\ln n)$.

Neka je $p_1 := \dfrac{\ln n + \omega/2}{n}$, a p_2 takav da $1 - p = (1 - p_1)(1 - p_2)$. Iz toga slijedi

$$\mathbb{G}_{n,p} = \mathbb{G}_{n,p_1} \cup \mathbb{G}_{n,p_2}$$

te

$$p_2 \sim \frac{\omega}{2n}.$$

Primijetimo da, obzirom da je $\ln n/n$ funkcija praga za ne postojanje izoliranog vrha, u $\mathbb{G}_{n,p_1}$ s visokom vjerojatnošću vrijedi $\delta \geq 1$.

Neka su $\{f_1, f_2, \ldots f_s\}$ bridovi iz $\mathbb{G}_{n,p_2}$ u nasumičnom redoslijedu. Vrijedi $s \sim n\omega/4$. Definirajmo $\mathbb{G}_0 = \mathbb{G}_{n,p_1}$ te $\mathbb{G}_i = \mathbb{G}_0 + \{f_1, f_2, \ldots, f_i\}$ za $i \geq 1$.

Fiksiramo neki trenutak $1 \leq i \leq s$ u kojemu $\mathbb{G}_i$ nema savršeno sparivanje. Ako je $f_{i+1} = (x, y)$ takav da $y \in A(x)$, onda vrijedi $\mu(\mathbb{G}_{i+1}) \geq \mu(\mathbb{G}_i) + 1$ odnosno red sparivanja strogo raste. Vidjeti ćemo da vrijedi $|A(x)| \geq \alpha n$, za neki $\alpha > 0$. Brid f_{i+1} bira se uniformno iz $\binom{n}{2}$ mogućih bridova. Ako za fiksan $x \in V$ odaberemo $x' \in A(x)$ i $y \in A(x')$ na maksimalno $\binom{\alpha n}{2}$ načina vrijedi da je $f_{i+1} = (x', y')$ brid za koji $y' \in A(x')$. Ako uračunamo da f_{i+1} ne ponavlja jedan od prethodnih $i \in O(n)$ bridova imamo ogradu

$$\begin{aligned} \mathbb{P}\left[\mu(\mathbb{G}_{i+1}) \geq \mu(\mathbb{G}_i) + 1 | f_1, f_2, \ldots f_i\right] &\geq \frac{\binom{\alpha n}{2} - i}{\binom{n}{2}} \\ &\geq \frac{\alpha^2}{2}. \end{aligned} \tag{4.1}$$

Primijetimo da po Lemi 4.1.2 svaki skup S sa manje od $n/(10 \cdot e \cdot 700)$ vrijedi $N_{\mathbb{G}_i}(S) \geq |S|$. Sada po Lemi 4.1.1 vrijedi da za svaki v za koji je $A(v)$ definiran vrijedi $|A(v)| \geq n/(10 \cdot e \cdot 700)$, odnosno $\alpha = (10e \cdot 700)^{-1}$.

Iz (4.1) slijedi da $Y := \mu(\mathbb{G}_s) - \mu(\mathbb{G}_0)$ stohastički dominira $X \sim B(s, \alpha^2/2)$, stoga

$$\mathbb{P}(U\ \mathbb{G}_{n,p}\ \textit{ne postoji savršeno sparivanje}) \leq o(1) + \mathbb{P}(B(s, \alpha^2/2) < n/2) = o(1).$$

□

4.2 Povezani rezultati

Za razliku od funkcija praga u Poglavlju 2 , funkcija $p_0 = \ln n/n$ je po Teoremu 4.1.3 *oštra funkcija praga*. To znači da za svaki $\epsilon > 0$ vrijedi

$$\mathbb{P}(\text{U } \mathbb{G}_{n,p} \text{ postoji savršeno sparivanje}) = \begin{cases} 0 & \text{za } (p/p_0) \to 1 - \epsilon, \\ 1 & \text{za } (p/p_0) \to 1 + \epsilon. \end{cases}$$

Iz rezultata u Potpoglavlju 2.2 vidi se da funkcije praga za postojanje strogo balansiranog podgrafa nisu stroge. Iz Teorema 4.1.3 se vidi da je $\ln n/n$ također stroga funkcija praga za svojstvo ne postojanja izoliranog vrha.

Istu oštru funkciju praga ima i svojstvo povezanosti. Sljedeći rezultat nalazi se u [2].

Teorem 4.2.1. *Neka je $c \in \mathbb{R}$ fiksan i neka je $p = \dfrac{\ln n + c + o(1)}{n}$. Tada*

$$\lim_{n\to\infty} (\mathbb{G}_{n,p}\ \textit{je povezan}) = e^{-e^{-c}}.$$

Dakle, "u isto vrijeme" kada nestaje zadnji izolirani vrh, odmah se pojavljuje povezanost i savršeno sparivanje. To znači da se svi drugi preduvjeti za postojanje povezanosti i savršenog sparivanja ostvaruju prije nego što nestaje zadnji izolirani vrh. Iz perspektive evolucije slučajnog grafa, divovska komponenta guta preostale komponente od većih prema manjima te se povezanost ostvaruje gutanjem zadnjeg izoliranog vrha.

Povezanost i najmanji stupanj slučajnog grafa povezani su i općenitije - k-povezanost se pojavljuje u isto vrijeme kada i najmanji stupanj u grafu postaje k. S $\kappa(G)$ označimo vršnu povezanost grafa G (više detalja u [9]). Proces slučajnog grafa je $\tilde{G}_n := \left(G_t; t \in \left[\binom{n}{2}\right]\right)$, gdje je G_t dobiven iz G_{t-1} dodavanjem jednog brida uniformnim biranjem, a G_0 je prazan graf na skupu vrhova $[n]$. Za taj proces intuitivno definiramo vrijeme pogađanja monotono rastućeg svojstva $\mathcal{A}$, koje označimo s $\tau(\mathcal{A})$. Sada smo u mogućnosti iskazati sljedeći teorem iz [2].

Teorem 4.2.2. *Za svaki $k \in \mathbb{N}$, za gotovo svaki proces slučajnog grafa vrijedi*

$$\tau(\delta(G) \geq k) = \tau(\kappa(G) \geq k).$$

Ovaj teorem opravdava izraz da se nestajanje zadnjeg izoliranog vrha i povezanost pojavljuju "u isto vrijeme". To je upravo tvrdnja gornjeg teorema kada je $k = 1$.

Nužan uvjet za postojanje Hamiltonovog ciklusa je da je minimalni stupanj u grafu barem 2. Kao i u slučaju povezanosti i savršenog sparivanja, uvjet na minimalni stupanj je i dovoljan. Rezultat iz [5] u nastavku formulira tu tvrdnju na sličan način kao i Teorem 4.1.3.

Teorem 4.2.3. *Neka je $\omega = \omega(n)$ funkcija, $c \in \mathbb{R}$ konstanta te $p = \dfrac{\ln n + \ln\ln n + \omega}{n}$. Tada vrijedi*

$$\lim_{n\to\infty} \mathbb{P}(U\ \mathbb{G}_{n,p} \text{ postoji Hamiltonov ciklus}) = \begin{cases} 0 & \text{za } \omega \to -\infty, \\ e^{-e^{-c}} & \text{za } \omega \to c, \\ 1 & \text{za } \omega \to \infty. \end{cases}$$

Također

$$\lim_{n\to\infty} \mathbb{P}(U\ \mathbb{G}_{n,p} \text{ postoji Hamiltonov ciklus}) = \lim_{n\to\infty} \mathbb{P}(\delta(\mathbb{G}_{n,p}) \geq 2).$$

Bibliografija

[1] Noga Alon i Joel H. Spencer, *The probabilistic method*, third ed.., JOHN WILEY & SONS, INC., 2007.

[2] Béla Bollobás, *Random Graphs*, second ed.., Cambridge University Press, 2001.

[3] Jordana Cepelewicz, *Elegant Six-Page Proof Reveals the Emergence of Random Structure*, Quanta Magazine (2022).

[4] Paul Erdős, Alfréd Rényi et al., *On the evolution of random graphs*, Publ. math. inst. hung. acad. sci **5** (1960), br. 1, 17–60.

[5] Alan Frieze i Michał Karoński, *Introduction to Random Graphs*, first ed.., Cambridge University Press, 2015.

[6] Edgar N Gilbert, *Random graphs*, The Annals of Mathematical Statistics **30** (1959), br. 4, 1141–1144.

[7] Michael Krivelevich i Benny Sudakov, *The phase transition in random graphs: A simple proof*, Random Structures & Algorithms **43** (2013), br. 2, 131–138.

[8] Jinyoung Park i Huy Pham, *A proof of the Kahn–Kalai conjecture*, Journal of the American Mathematical Society (2023), `https://arxiv.org/pdf/2203.17207.pdf`, posjećena 2023-08-24.

[9] Matea Petohleb, *Povezanost grafova*, 2022, `https://urn.nsk.hr/urn:nbn:hr:196:624816`, posjećena 2023-08-21.

[10] Erik Thörnblad, *Another proof of Moon's theorem on generalised tournament score sequences*, 2016, `https://arxiv.org/pdf/1605.06407`, posjećena 2023-08-24.

Sažetak

U ovom radu izlažu se neke osnovne spoznaje o slučajnim grafovima, koristeći se modelom $\mathbb{G}_{n,p}$. U tom modelu slučajan graf je element vjerojatnosnog prostora na skupu svih grafova sa vrhovima $[n]$, a gdje je vjerojatnost grafa sa k bridova $p^k(1-p)^{\binom{n}{2}-k}$, $p \in \langle 0, 1\rangle$. To jest, svaki brid nezavisno postoji s vjerojatnošću $p \in \langle 0, 1\rangle$ koja je uglavnom funkcija veličine grafa n.

U uvodnom poglavlju definiraju se osnovni pojmovi vezani uz grafove koji se spominju i u nastavku rada: stabla, komponente, Hamiltonovi putovi itd. Također se definiraju osnovni modeli slučajnih grafova. U uvodnom poglavlju dodatno najavljujemo alate koji su korisni u proučavanju slučajnih grafova kao što su metode momenata.

U drugom poglavlju dokazuju se dva rezultata vezana za pojavljivanje fiksnog grafa H u grafu $\mathbb{G}_{n,p}$. Pokazuje se postojanje granične funkcije $p^* = p^*(n)$ za to svojstvo te vjerojatnost da $\mathbb{G}_{n,p}$ sadrži H baš u tom prijelaznom području. Teoremi su prvenstveno dokazani za balansirane, odnosno strogo balansirane grafove, ali se mogu proširiti na širu klasu grafova.

U trećem poglavlju dokazano je nekoliko tvrdnji koje se tiču evolucije grafova, odnosno promjene tipične strukture $\mathbb{G}_{n,p}$ s obzirom na promjenu funkcije $p = p(n)$. Objašnjava se kako se definiraju faze, odnosno režimi i iznose se neke tvrdnje o $\mathbb{G}_{n,p}$ u pojedinom režimu. Potom se na dva načina, standardno i pomoću pretrage u dubinu kao u članku [7] dokazuju tvrdnje o veličini najveće komponente u $\mathbb{G}_{n,p}$.

U posljednjem poglavlju dokazuje se teorem o postojanju savršenog sparivanja u $\mathbb{G}_{n,p}$. Povezuje se taj rezultat s Hamiltonovim ciklusima, postojanjem izoliranog vrha te povezanošću u $\mathbb{G}_{n,p}$.

Summary

This thesis is concerned with fundamental properties of random graphs. The model we use is $\mathbb{G}_{n,p}$. A random graph $\mathbb{G}_{n,p}$ is a random element of the space of all graphs on n vertices, where each k-edge graph occurs with probability $p^k(1-p)^{\binom{n}{2}-k}$. That is: all edges are independent and the probability of an edge existing in $\mathbb{G}_{n,p}$ is $p \in \langle 0, 1 \rangle$. This probability p is usually a function of the graph size n, since the properties of $\mathbb{G}_{n,p}$ are explored as n approaches infinity.

The introductory chapter focuses on the definitions of the basic structures and properties concerning graphs, such as trees, components, Hamilton cycles, etc. It also defines random graphs and related notions. This chapter also introduces the tools often used in proving statements about random graphs, with the focus on the method of moments.

The second chapter focuses on two theorems which relate to the property of random graph having a fixed graph H as a subgraph. We find a threshold function p^* for this property and calculate the probability of this property when $p \in \Theta(p^*)$. These theorems are proven for balanced and strictly balanced graphs, but can be extended.

The third chapter contains proofs of several statements concerning the evolution of random graphs. The evolution of random graphs means observing the change in a typical structure of $\mathbb{G}_{n,p}$ as $p = p(n)$ changes. We explain how phases or regimes of p are defined and state a few facts about random graph $\mathbb{G}_{n,p}$ in each phase. We proceed to prove theorems about the size of the largest component in very superctirical and very subcritical regimes. Overlapping proofs are presented, a standard proof, and a simpler proof of a weaker statement presented in the paper [7].

The final chapter is concerned with the topic of perfect matchings in a random graph. We explain how this connects to the property of having a Hamilton cycle, being connected and having an isolated vertex.

Životopis

Rođena sam u Zagrebu 17. srpnja 1996. godine. Osnovnoškolsko obrazovanje završila sam u Osnovnoj školi Miroslava Krleže. Potom sam upisala XV. Gimnaziju u Zagrebu po programu pojačane matematike. Godine 2015./16. upisujem inženjerski smjer matematike na Prirodoslovno-Matematičkom fakultetu Sveučilišta u Zagrebu. Diplomski studij financijske i poslovne matematike upisujem 2019./20. te u svrhu završavanja tog studija pišem ovaj diplomski rad.