Математички факултет
Универзитет у Београду

Мастер рад

Вјероватносне методе у теорији графова

Андријана Марјановић, 1043/2018

Ментор: **Др Јелена Јоцковић**

Септембар 2019.

Садржај

Поглавље 1

Увод

У раду су представљене централне теме у области случајних графова које се ослањају на теорију вјероватноћа. Рад садржи преглед најчешћих модела случајних графова, као и њихове особине.

Прије почетка проучавања случајних графова, наведене су основне вјероватносне методе и општи појмови теорије графова који се јављају у раду. С обзиром на то да се на случајне графове гледа из угла теорије вјероватноћа наглашено је право значење тог појма и дефинисан је вјероватносни простор случајних графова. Након приказа најпознатијих модела случајних графова и њиховог поређења, наведени су познати резултати и неки докази. Структура случајних графова је анализирана кроз представљање њихових вјероватносних својстава и додатних услова како би неко својство важило за случајан граф. Дати су примјери графова који садрже нека од тих својстава, а за одређена својства је показано да она заиста важе за случајне графове. На крају су наведене области примјене, као и примјери неких свакодневних и реалних проблема у којима се користе случајни графови.

Оно по чему се случајни графови разликују од графова јесте начин на који они настају, тј. егзистенција грана у случајном графу може да се разумије као један догађај који настаје са одређеном вјероватноћом, док у типичном графу гране прецизно задате скупом грана увијек постоје. Како је теорија случајних графова заправо комбинација теорије графова и теорије вјероватноћа, случајан граф се интуитивно може разумјети као структура настала примјеном једне од варијанти повезивања чворова гранама. Најједноставније речено, ако се под појмом потпун граф сматра граф који има n чворова и максималан број грана $N = \binom{n}{2}$, тј. све могуће гране, онда случајан граф настаје у складу са једним од могућих подскупова скупа грана потпуног графа. Ако се фиксира вриједност $p \in [0, 1]$ таква да се у складу са Бернулијевом расподјелом i-та грана бира са вјероватноћом p независно до других грана, добија се Ердеш-Ренијев ($Erdös - Rényi$) [3] граф.

Осим два најзначајнија математичара у овој области по којима модел случајног графа носи назив, треба поменути и друге математичаре који су дали велики допринос у изучавању случајних графова, као што су Гилберт (*Gilbert*), Болобас (*Bollobás*) [2], Томасон (*Thomasson*).

Како се суочавамо са изазовом разумијевања структуре и развоја великих система, теорија случајних графова постала је битна област која се још увијек истражује и шири. Повећање интересовања за ову веома обимну област приписује се спознаји да су мреже свуда, од сложених интеракција ћелија човјековог тијела до друштвених мрежа. Комплексне мреже почивају на случајним графовима који у многим случајевима обезбјеђују почетну тачку моделирања нетривијалних структура. Велике мреже у природи које расту непредвидиво брзо погодне су за моделирање случајним графовима.

Теорија графова је нашла свој пут у друге науке као богат извор модела. Случајни графови као објекти са веома занимљивом структуром примјењиви су у многим областима, при чему је једна од популарнијих њихова примјена у рачунарству, односно програмирању. Допринос интересовању за графове, такође, дала је и могућност реализације и графичког приказа сложених графова (мрежа) помоћу различитих програмских језика.

Поглавље 2

Основни појмови из теорије вјероватноћа

Ово поглавље пружа кратак подсјетник основних принципа теорије вјероватноћа који се користе у даљем раду и потребни су за разумијевање случајних графова.

2.1 Конвергенција Биномне расподјеле ка Пуасоновој расподјели

2.1.1 Конвергенција у расподјели

Конвергенција у расподјели је најслабији тип конвергенције. Све што је потребно је да функција расподјеле случајне величине X_n конвергира функцији расподјеле случајне величине X када $n \longrightarrow \infty$, не захтјевајући зависност случајних величина X_n и X. Конвергенција у расподјели се записује као $X_n \xrightarrow{D} X$.

Дефиниција 2.1. *Низ случајних величина $(X_n)_{n \in \mathbb{N}}$, са функцијом расподјеле F_n, конвергира у расподјели случајној величини X чија је функција расподјеле F ако за све тачке x у којима је функција F непрекидна важи*

$$\lim_{n \longrightarrow \infty} F_n(x) = F(x).$$

2.1.2 Пуасонова апроксимација Биномне расподјеле

Иако се на први поглед чини да су неповезане, детаљније разматрање открива занимљиву релацију између Пуасонове и Биномне расподјеле. Пуасонова расподјела је заправо специјалан случај Биномне расподјеле када је број понављања експеримента велики тј. $n \longrightarrow \infty$, а вјероватноћа успјеха јако мала тј. $p \longrightarrow 0$.

Нека су параметри Биномне расподјеле број експеримената n и вјероватноћа успјеха p. Нека је λ константа дефинисана као производ броја покушаја и вјероватноће успјеха, тј. са $\lambda = np$. Одавде је $p = \lambda/n$ што ће бити коришћено у наставку.
Вјероватноћа да случајна величина X која има Биномну расподјелу са параметрима n и p, узима вриједност $k = 0, 1 \ldots n$ је

$$P\{X = k\} = \binom{n}{k} p^k (1-p)^{n-k}.$$

Када у овај израз замјенимо p изражено преко λ, а затим посматрамо граничну вриједност вјероватноће добија се

$$\lim_{n \longrightarrow \infty} P\{X = k\} = \lim_{n \longrightarrow \infty} \binom{n}{k} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k} =$$

$$= \frac{\lambda^k}{k!} \lim_{n \longrightarrow \infty} \frac{n!}{(n-k)!n^k} \left(1 - \frac{\lambda}{n}\right)^{n-k} =$$

$$= \frac{\lambda^k}{k!} e^{-\lambda}.$$

За добијање апроксимације коришћено је

$$\lim_{n \longrightarrow \infty} \left(1 - \frac{\lambda}{n}\right)^n = e^{-\lambda},$$

$$\lim_{n \longrightarrow \infty} \left(1 - \frac{\lambda}{n}\right)^{-k} = 1 \text{ и}$$

$$\binom{n}{k} \sim \frac{n^k}{k!}, \text{ када } n \longrightarrow \infty.$$

Добијени израз је еквивалентан вјероватноћи k успјеха из закона Пуасонове расподјеле са параметром λ који се реализују по одређеном временском периоду.

О Пуасоновом процесу као граничном за биномни процес говори сљедећа теорема позната под називом Закон малих (ријетких) бројева.

Теорема 2.1. ***(Закон малих бројева)*** *Нека је $(X_n)_{n \in \mathbb{N}}$ низ случајних величина које имају Биномну расподјелу са параметрима n и p_n за $n \in \mathbb{N}$ при чему $p_n \longrightarrow 0$, $n \longrightarrow \infty$, тако да важи $np_n \longrightarrow \lambda$, када $n \longrightarrow \infty$, за $\lambda > 0$, $\lambda = const$. Тада важи да низ случајних величина X_n када $n \longrightarrow \infty$ у расподјели конвергира случајној величини X која има Пуасонову расподјелу са параметром λ.*

2.2 Метода првог и другог момента

Марковљева неједнакости и Чебишевљева неједнакост пружају два елементарна пробабилистичка алата корисна за доказивање нетривијалних резултата. Марковљева неједнакост одређује горњу границу вјероватноће да је ненегативна случајна величина већа или једнака од неке позитивне константе. Чебишевљева неједнакост даје горњу границу концентрације случајне величине око њеног очекивања.

Теорема 2.2. ***(Марковљева неједнакост)*** *За сваку ненегативну случајну величину X и за $\forall k > 0$ важи*

$$P\{X \geq k\} \leq \frac{E[X]}{k}.$$

Доказ. *Нека је I_k индикаторска случајна величина догађаја $X \geq k$, тј.*

$$I_k = \begin{cases} 1, & \text{ако је } X \geq k \\ 0, & \text{иначе.} \end{cases}$$

Према томе, овај догађај се може записати као $X \geq k \cdot I_k$, из чега слиједи да је

$$E[X] \geq E[k \cdot I_k] = k \cdot E[I_k] = k \cdot P\{I_k = 1\} = k \cdot P\{X \geq k\}.$$

□

Посљедица 2.1. *Ако је X ненегативна случајна величина, онда за $\forall c \geq 1$ важи*

$$P\{X \geq c \cdot E[X]\} \leq \frac{1}{c}.$$

Доказ. *Нека је $k = c \cdot E[X]$. Користећи Марковљеву неједнакост добија се*

$$P\{X \geq c \cdot E[X]\} \leq \frac{E[X]}{c \cdot E[X]} = \frac{1}{c}.$$

□

Марковљева неједнакост је основа методе првог момента. У даљем раду среће се доказивање да је $P\{X > 0\} = o(1)$, гдје је X ненегативна случајна величина од интереса. Израчунавање очекивања случајне величине X може бити лакше од директног израчунавања вјероватноће $P\{X > 0\}$. Ако је

$$E[X] = o(1),$$

онда на основу Марковљеве неједнакости важи

$$P\{X > 0\} \leq E[X] = o(1).$$

Из овога се закључује да је $X = 0$ са великом вјероватноћом. Овај начин процјене вјероватноће је познат као **метода првог момента**. Даље се на основу Марковљеве неједнакости могу добити вјероватносне неједнакости за више редове. Ако је ϕ строго растућа функција, онда важи

$$P\{X \geq t\} = P\{\phi(X) \geq \phi(t)\} \leq \frac{E[\phi(X)]}{\phi(t)}.$$

За $\phi(t) = t^2$, $t > 0$ добија се Чебишевљева неједнакост.

Теорема 2.3. ***(Чебишевљева неједнакост)*** *Нека је X случајна величина са коначним очекивањем и коначном дисперзијом. Тада за $\forall t > 0$ важи*

$$P\{|X - E[X]| \geq t\} \leq \frac{D[X]}{t^2}.$$

Доказ. *Нека је случајна величина Y дефинисана са $Y = (X - E[X])^2$. Y је ненегативна и $|X - E[X]| \geq t$ ако и само ако $Y \geq t^2$. Осим тога, важи и $E[Y] = E[(X - E[X])^2] = D[X]$. Користећи Марковљеву неједнакост добија се*

$$P\{|X - E[X]| \geq t\} = P\{(X - E[X])^2 \geq t^2\} = P\{Y \geq t^2\} \leq \tfrac{E[Y]}{t^2} = \tfrac{D[X]}{t^2}.$$

□

Из Чебишевљеве неједнакости слиједе једноставне посљедице познатије под називом **метода другог момента**.

Посљедица 2.2. *Нека је X ненегативна случајна величина која узима цијеле вриједности, онда важи*

$$P\{X = 0\} \leq \frac{D[X]}{(E[X])^2}.$$

Доказ.

$$P\{X = 0\} \leq P\{|X - E[X]| \geq E[X]\} \leq \frac{D[X]}{(E[X])^2}.$$

□

Посљедица 2.3. *Нека је X случајна величина таква да важи $D[X] = o((E[X])^2)$, онда X скоро сигурно узима позитивне вриједности.*

Посљедица 2.4. *Нека је X случајна величина са позитивним очекивањем $E[X] > 0$ и ако*

$$E[X^2] \leq (E[X])^2(1 + o(1)),$$

онда X скоро сигурно узима позитивне вриједности.

Нека је случајна величина X сума m индикаторских случајних величина $X_1, ..., X_m$, гдје је $P\{X_i = 1\} = p_i$, тј.

$$X = X_1 + ... + X_m.$$

У даљем раду потребно је доказати да је $X > 0$ са великом вјероватноћом. Чак ако $E[X] \longrightarrow \infty$, то не значи да је $X > 0$ скоро сигурно. За доказивање ове тврдње користи се метода другог момента. Како из посљедице 2.2 важи да је $P\{X = 0\} \leq \dfrac{D[X]}{(E[X])^2}$ довољно је показати да је $\dfrac{D[X]}{(E[X])^2} = o(1)$. Према томе, проблем је поједностављен на израчунавање горње границе дисперзије.

$$D[X] = \textstyle\sum_{i=1}^{m} D[X_i] + \sum_{i\neq j} Cov[X_i, X_j] \leq E[X] + \sum_{i\neq j} Cov[X_i, X_j].$$

Треба примјетити да за дисперзију индикаторске случајне промјенљиве важи

$$D[X_i] = p_i(1 - p_i) \leq p_i = E[X_i].$$

Због линеарности очекивања важи да је

$$\sum_i D[X_i] \leq \sum_i E[X_i] = E\left[\sum_i X_i\right] = E[X].$$

Коваријација двије случајне величине A, B дефинисана је са

$$Cov[A, B] = E[AB] - E[A] \cdot E[B].$$

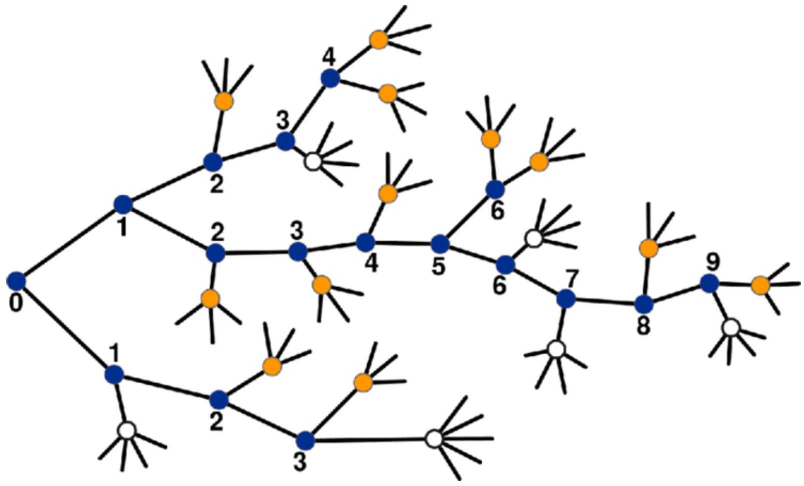
У случају индикаторских случајних величина добија се сљедећи израз

$$Cov[X_i, X_j] = P\{X_i = X_j = 1\} - P\{X_i = 1\}P\{X_j = 1\}.$$

Илустрација употребе ових метода садржана је у даљем раду.

2.3 Процес гранања

Модел популације која расте на случајан начин уз одсуство ограничења, може се представити процесом гранања. Неформално, процес се може описати на сљедећи начин. Популацију у почетном тренутку $n = 0$ чини само једна јединка, која се у наредном тренутку распада и даје фамилију јединки исте врсте, што се наставља кроз генерације. Вриједност процеса у тренутку n, односно у генерацији n, је број јединки рођених у n-тој генерацији. На слици 2.1 приказан је начин на који се нумеришу генерације. Претпоставка је да је расподјела броја потомака било које јединке иста.



Слика 2.1: Процес гранања са назначеним генерацијама јединки које су исте врсте

Ненегативна случајна величина X_n која узима вриједности из скупа $\mathbb{N}_0$ означава број јединки у популацији у тренутку $n \geq 0$, што се често интерпретира као величина n-те генерације. Нека је $\{Y_{n,j}, n \in \mathbb{N}, j \in \mathbb{N}\}$ фамилија случајних величина које су независне и једнако расподијељене и узимају вриједности у скупу $\mathbb{N}_0$. Нека је $p_k = P\{Y_{n,j} = k\}$ за $k \in \mathbb{N}_0$. Случајна величина $Y_{n,j}$ означава број јединки у n-тој генерацији које су потомци j-те јединке из $n-1$-ве генерације, гдје је $n \in \mathbb{N}_0$ и $j \in \{1, 2..., X_n\}$. Нека је $\{X_n, n \in \mathbb{N}_0\}$ процес дефинисан на сљедећи начин:

$$X_0 = 1$$

$$X_1 = Y_{1,1} + Y_{1,2} + ... + Y_{1,X_0}$$

$$X_2 = Y_{2,1} + Y_{2,2} + ... + Y_{2,X_1}$$

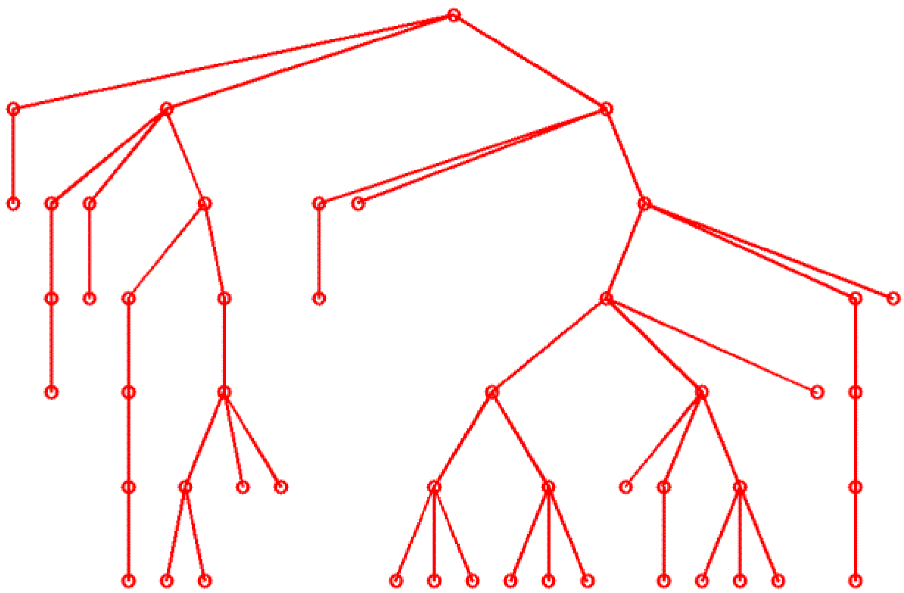
$$\vdots$$

$$X_n = Y_{n,1} + Y_{n,2} + ... + Y_{n,X_{n-1}}$$

Процес $\{X_n, n \in \mathbb{N}_0\}$ се назива **прост процес гранања** или $Galton - Watson$-ов процес.

Прост процес гранања је Марковљев процес јер X_n зависи само од X_{n-1}, не и од претходних стања.

Приликом изучавања процеса гранања јављају се вјероватноће од интереса као што су вјероватноћа опстанка популације до n-те генерације, вјероватноћа опстанка кроз бесконачно много генерација, вјероватноћа изумирања популације, вјероватноћа изумирања популације најкасније у n-тој генерацији и слично.



Слика 2.2: Процес гранања

На слици 2.2 дат је шематски приказ једног процеса гранања. У почетном тренутку постоји једна јединка, што се означава са $X_0 = 1$. У наредном тренутку $n = 1$, та јединка даје 3 потомка, па је $X_1 = 3$, а они редом имају 1, 3 и 3 потомка па је $Y_{1,1} = 1$, $Y_{1,2} = 3$ и $Y_{1,3} = 3$. Такође, важи да је $X_2 = Y_{1,1} + Y_{1,2} + Y_{1,3} = 7$, што се може провјерити на слици. Поступак се наставља све док у популацији постоји јединка која ће дати потомке за наредну генерацију.

2.4 Генераторна функција

Нека је $(a_n)_{n\in\mathbb{N}_0}$ низ реалних бројева. Ако функција

$$\varphi(s) = a_0 + a_1 s + a_2 s^2 \ldots = \sum_{i=0}^{\infty} a_i s^i$$

конвергира у неком интервалу $|s| < s_0$, онда се φ назива **генераторна функција** низа $(a_n)_{n\in\mathbb{N}_0}$. Промјенљива s се овдје назива **скривена промјенљива** (енг. *dummy*). Ако је низ $(a_n)_{n\in\mathbb{N}_0}$ ограничен, $\varphi(s)$ конвергира бар за неко s које је различито од 0.

Нека је X случајна величина која узима вриједности из скупа $\mathbb{N}_0$ са одговарајућом вјероватноћом
$$P\{X=k\}=p_k.$$

Генераторна функција низа $(p_n)_{n\in\mathbb{N}_0}$ је

$$\varphi(s)=p_0+p_1s+p_2s^2...=\sum_{i=0}^{\infty}p_is^i.$$

За $s=1$ добија се
$$\varphi(1)=p_0+p_1+...=1.$$

Стога, генераторна функција конвергира у неком интервалу који садржи 1. Нека је $\varphi(s)$ генераторна функција случајне величине X која узима вриједности из скупа $\mathbb{N}_0$. Под условом да очекивање постоји, тада важи

$$\varphi'(1)=\sum_{i=0}^{\infty}ip_i=E[X].$$

Дакле, ако је позната генераторна функција лако се може наћи очекивана вриједнсот. Слично,
$$E[X(X-1)]=\sum_{i=1}^{\infty}i(i-1)p_i=\varphi''(1)$$
или уопштено
$$E[X(X-1)\cdots(X-r+1)]=\varphi^{(r)}(1).$$

Дисперзија се може израчунати помоћу генераторне функције на сљедећи начин

$$D[X]=\varphi''(1)+\varphi'(1)-\varphi'(1)^2.$$

Нека је X случајна величина која узима вриједности из скупа $\mathbb{N}_0$. За неко s израз s^X је добро дефинисана случајна величина са очекивањем

$$E[s^X]=\sum_{i=0}^{\infty}p_i\cdot s^i=\varphi(s).$$

Ако су X и Y независне случајне величине, онда су такве и s^X и s^Y и важи

$$E[s^{X+Y}]=E[s^X]\cdot E[s^Y],$$

односно исто то може бити записано помоћу генераторне функције као

$$\varphi_{X+Y}(s)=\varphi_X(s)\varphi_Y(s).$$

Овај израз се може уопштити за суму n независних једнако расподјељених случајних величини $X_1, ..., X_n$ са генераторном функцијом $\varphi_X(s)$, на сљедећи начин:

$$S_n = X_1 + ... + X_n$$

$$\varphi_{S_n}(s) = \varphi_X(s)^n.$$

2.5 Асимптотске ознаке

Приликом апроксимације сложених израза често се користе краћи записи који садрже само доминантан члан развоја израза. Неке од сљедећих нотација користе се у раду.

Запис	Асимптотско поређење	Дефиниција
$f \in o(g)$	”$<$”	$\lim_{x \longrightarrow \infty} \dfrac{f(x)}{g(x)} = 0$
$f \in \mathcal{O}(g)$	”$\leq$”	$\lim_{x \longrightarrow \infty} \dfrac{f(x)}{g(x)} < \infty$
$f \in \Theta(g)$	”$=$”	$f \in \mathcal{O}(g)$ и $g \in \mathcal{O}(f)$
$f \in \Omega(g)$	”$\geq$”	$g \in \mathcal{O}(f)$

Поглавље 3

Основни појмови из теорије графова

Ово поглавље даје дефиницију графа и основних појмова теорије графова. Представљени су елементарни концепти графова и њихова својства која важе и за случајне графове.

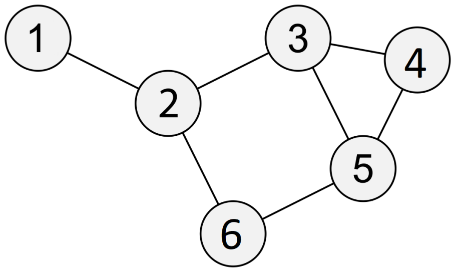
3.1 Дефиниција графа и неке особине

Дефиниција 3.1. *Граф је уређен пар $G = (V, E)$, гдје је V коначан скуп, а E подскуп скупа двочланих подскупова скупа V.*

Овако дефинисан граф у литератури се назива још и **неоријентисан граф** или **прост граф**. Елементи скупа V се називају **чворови** (тјемена) графа G, а елементи скупа E **гране** (ивице) графа G.

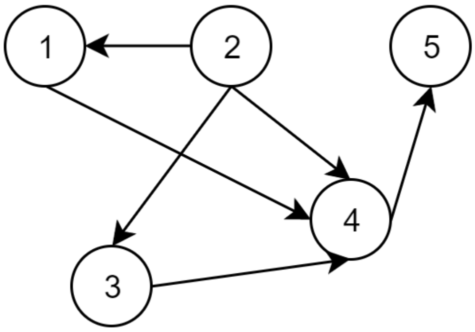
Како би се структура графова лакше разумјела, они се представљају у равни тако што се сваком чвору графа додијели једна тачка равни, а затим се тачке које одговарају чворовима x и y споје кривом ако и само ако је $\{x, y\} \in E$. Приказ једноставног графа $G = (V, E)$ са скупом чворова $V = \{1, 2, 3, 4, 5, 6\}$ и гранама $E = \{\{1, 2\}, \{2, 3\}, \{2, 6\}, \{3, 4\}, \{3, 5\}, \{4, 5\}, \{5, 6\}\}$ дат је на слици 3.1. Графички приказ је често практичнији начин задавања графа од експлицитног навођења скупова V и E, јер они могу бити нарочито обимни када је у питању сложенији граф.

Ако је $\{x, y\} \in E$, то значи да између чворова x и y постоји грана и тада се за те чворове каже да су **сусједни**, а то се може краће записати и као $xy \in E$. Ако је грана $yx \in E$ исто што и $xy \in E$, и то важи за све гране графа, односно није битан смијер релације која је успостављена између два чвора, онда се такав граф



Слика 3.1: Неоријентисан граф

назива **неоријентисан** или **симетричан граф**, какав је и граф на слици 3.1. Ако је битно нагласити смијер релације која је представљена граном, онда је граф који садржи такве гране **оријентисан** или **антисиметричан граф**. Приликом његовог графичког представљања на гранама су означене стрелице.



Слика 3.2: Оријентисан граф

Скуп сусједа чвора v означава се са $N(v)$, а **степен чвора**, који означава број његових сусједа, означава се са $d(v)$ (енг. *degree*), тј. важи

$$N(v) := \{u \in V | uv \in E\}$$

$$d(v) := |N(v)|.$$

Чвор степена 0 назива се **изолован чвор**, а чвор степена 1 назива се **лист**. За дати граф G са $\delta(G)$ и $\Delta(G)$ означавамо минимални и максимални степен његових чворова, тј.

$$\delta(G) := min\{d(v)|v \in V\}$$
$$\Delta(G) := max\{d(v)|v \in V\}.$$

Код оријентисаног графа могу се одредити **улазни** (*indegree*) и **излазни сте-**

пен чвора (*outdegree*) у зависности од смијера релације која је представљена граном.

Тврђење 3.1. *У сваком графу* $G = (V, E)$ *важи*

$$\sum_{v \in V} d(v) = 2 \cdot |E|.$$

Доказ. *Број* $d(v)$ *је број грана на којима се налази чвор* v. *Дакле, број* $\sum_{v \in V} d(v)$ *може се добити тако што се за сваки чвор* $v \in V$ *посматра број грана на којима се налази тај чвор* v *и затим се сви бројеви саберу. Како свака грана садржи два чвора, приликом овог бројања свака грана* xy *је рачуната два пута (једном као грана која садржи* x, *а други пут као грана која садржи* y*), тако да је дати збир заиста једнак* $2 \cdot |E|$.

□

Посљедица 3.1. *У сваком графу број чворова непарног степена је паран.*

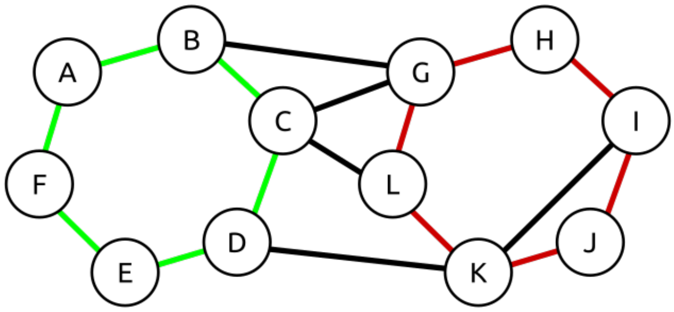
Доказ. *По претходном тврђењу број* $\sum_{v \in V} d(v)$ *је паран јер се може представити као производ броја два броја од којих је један 2. У збиру* $\sum_{v \in V} d(v)$ *неки бројеви су парни, а неки непарни. С обзиром на то да је збир (неких) бројева паран ако и само ако је у њему паран број непарних сабирака, закључује се да број оних сабирака* $d(v)$ *који су непарни мора бити паран, а то значи да број чворова који имају непаран степен мора бити паран.*

□

Дефиниција 3.2. *Низ чворова* $v_1, v_2, ..., v_n$, *такав да за свако* $i \in \{1, 2..., n-1\}$ *важи* $v_i v_{i+1} \in E$ *назива се* ***шетња*** *у графу. За гране* $v_i v_{i+1}$, $1 \leq i \leq n-1$ *се каже да су садржане у овој шетњи.*

Шетња може бити:

- **затворена** ако је $v_1 = v_n$, тј. шетња која се завршава у истом чвору у ком је и почела;
- **пут** ако су чворови $v_1, v_2, ..., v_n$ различити тј. приликом шетње кроз сваки чвор се пролази тачно једном;
- **циклус** ако је $v_1 = v_n$, а чворови $v_2, v_3, ..., v_n - 1$ су међусобно различити и различити од v_1 тј. од v_n, односно то је затворена шетња која је уједно и пут.



Слика 3.3: Неоријентисан граф који садржи циклусе

Дефиниција 3.3. *Дужина пута (циклуса) представља број различитих чворова на том путу (циклусу).*

Дакле, пут v_1, v_2, v_3 је дужине 3, јер садржи 3 чвора, док је број грана које садржи 2. Такође, циклус v_1, v_2, v_3, v_1 је дужине 3, а садржи и 3 гране.

Дефиниција 3.4. *Дужина шетње $v_1, v_2, ..., v_n$ је једнака броју чворова n ако је $v_1 \neq v_n$, а једнака $n-1$ ако је $v_1 = v_n$.*

С обзиром на то да се током шетње чворови и гране могу понављати, дужина шетње не мора бити једнака ни броју чворова ни броју грана који се током ње појављују.

Дефиниција 3.5. ***Повезан граф*** *је неоријентисан граф код кога су било која два чвора повезана путем. Ако постоје два чвора која се не могу повезати путем граф је неповезан.*

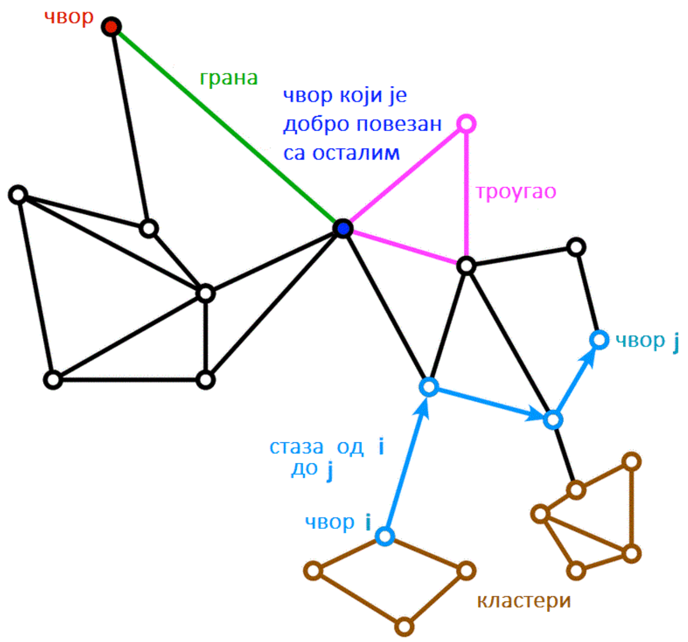
Чвор који је повезан са већим бројем чворова за разлику од осталих чворова у графу (који су повезани са мањим бројем чворова) и стога има висок степен назива се **главни чвор** (енг. *hub*). На слици 3.4 дат је приказ основних елемената графа.

Дефиниција 3.6. *Грана графа која полази из једног чвора и завршава се у истом том чвору назива се* ***петља****.*

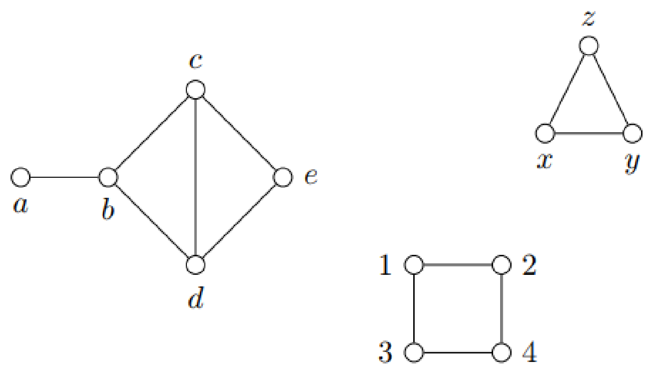
На скупу чворова V графа G може се увести релација $\sim$ на сљедећи начин

$$a \sim b \text{ акко } \textit{постоји шетња од } a \textit{ до } b \textit{ у графу } G.$$

Овако дефинисана релација је релација еквиваленције на скупу чворова V. Класе еквиваленције релације $\sim$ се називају **компоненте повезаности** графа G. Као што је већ речено, ако граф G има тачно једну компоненту повезаности каже се да је тај граф **повезан**. Ред компоненте представља број чворова који припадају тој компоненти. На слици 3.5 дат је граф чије су компоненте повезаности $\{a, b, c, d, e\}$, $\{1, 2, 3, 4\}$ и $\{x, y, z\}$.



Слика 3.4: Основни елементи графа



Слика 3.5: Граф са компонентама повезаности

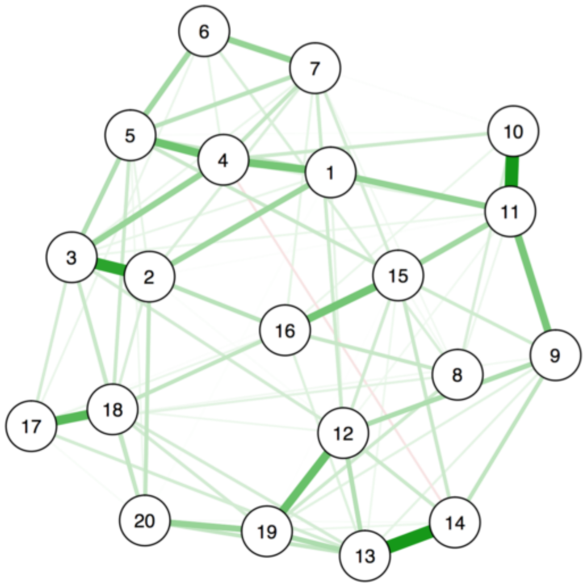
Дијаметар графа G у ознаци $diam(G)$ означава највећу удаљеност $d(x, y)$ између два чвора x и y који припадају истој компоненти графа G.

$$diam(G) = max\{d(x, y)|x, y \in V \wedge d(x, y) < \infty\}$$

Дефиниција 3.7. *Граф $G' = (V', E')$ је **подграф** графа $G = (V, E)$ ако је $V' \subseteq V$ и за све $xy \in E'$ важи $xy \in E$. Граф $G' = (V', E')$ је **индукован подграф** графа $G = (V, E)$ ако је $V' \subseteq V$ и за све $x, y \in V'$ важи $xy \in E'$ ако и само ако $xy \in E$.*

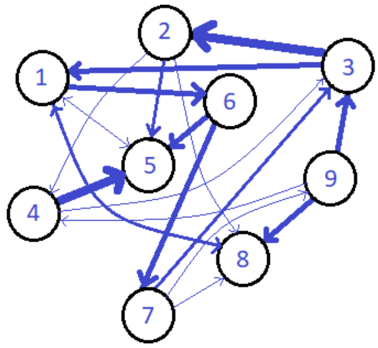
Дефиниција 3.8. *Нека су $G = (V, E)$ и $G' = (V', E')$ графови. За пресликавање $f : V \longrightarrow V'$ се каже да је **изоморфизам графова** G и G' ако је f бијекција и* *за све $x, y \in V$ важи $xy \in E$ ако и само ако $f(x)f(y) \in E'$. Уколико за графове $G = (V, E)$ и $G' = (V', E')$ постоји бар један изоморфизам каже се да су они* ***изоморфни*** *и пише се $G \cong G'$. Релација $\cong$ је релација еквиваленције.*

Дефиниција 3.9. ***Тежински граф*** *је граф чија свака грана има одређену тежину. Тежински граф који је оријентисан назива се* ***мрежа****.*

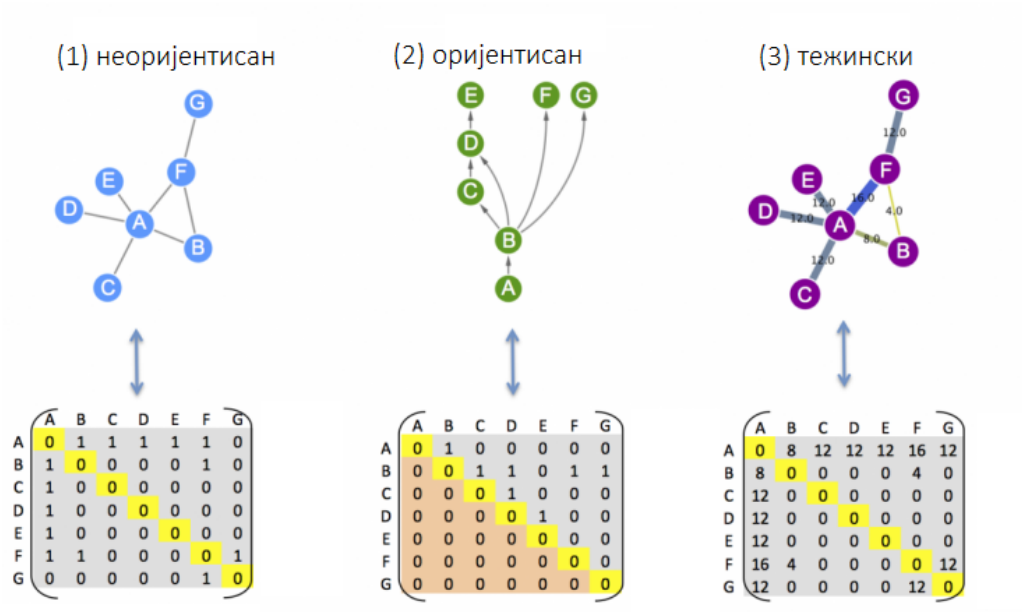


Слика 3.6: Неоријентисан тежински граф

Тежински графови се често користе за представљање реалних проблема, као што је приказ мреже путева, при чему тежина гране може да се односи на дужину пута између два чвора, односно два града. Други примјер примјене тежинских графова, који ће бити представљен касније је моделовање банкарског сектора неке државе, при чему тежина гране одговара међубанкарским трансакцијама.



Слика 3.7: Оријентисан тежински граф (мрежа)



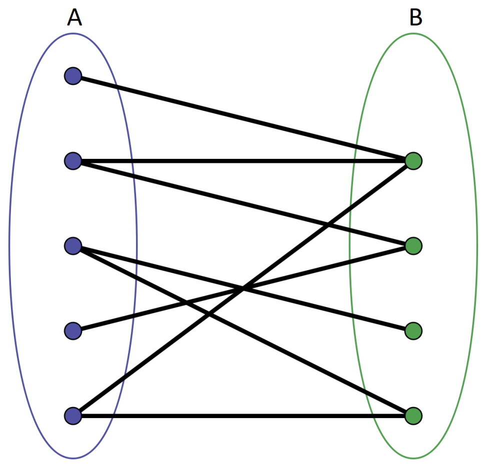
Слика 3.8: Графови различитих особина са својим матрицама сусједства

На сликама 3.6 и 3.7 су дати тежински графови. На слици 3.7 која приказује оријентисан тежински граф може се примјетити да су релације између неких чворова изражене у оба смјера и тада се на оба краја гране стављају стрелице. Овакви графови се често користе у пракси јер на једноставан начин дају прецизне односе међу чворовима у погледу смјера, дебљине гране, нијансе боје гране и слично.

Приликом представљања веза између чворова важан појам је **матрица сусједства** или **повезаности** (енг. *adjacency matrix*). Ријеч је о квадратној матрици која пружа увид у повезаност чворова, односно на одговарајућем пољу у матрици сусједства налази се ненула вриједност која индицира да веза два чвора постоји. Ако чворови i и j нису повезани, онда поље које одговара i-тој врсти и j-тој колони, као и j-тој врсти и i-тој колони садржи 0. У супротном одговарајуће поље од два наведена, садржи 1 зависно од смјера у ком се грана пружа, осим када је у питању тежински граф. Тада та вриједност представља тежину гране, односно то је неки реалан број. Примјер матрица сусједства са одговарајућим графовима дат је на слици 3.8.

Дефиниција 3.10. *Граф $G = (V, E)$ је* ***бипартитиван*** *ако постоје непразни скупови A и B такви да је $A \cap B = \emptyset$, $A \cup B = V$, и за све $xy \in E$ тачно један елемент скупа $\{x, y\}$ је у A и тачно један је у B. Тада је $(A,\ B)$* ***партиција бипартитивног графа*** *G. (слика 3.9)*

Дефиниција 3.11. ***Комплетан (потпун)*** *граф или* ***клика*** *је граф у коме су свака два чвора спојена граном.*



Слика 3.9: Бипартитиван граф

Комплетан граф који има n чворова се означава са K_n и има $\binom{n}{2}$ грана јер сваки двочлани подскуп скупа чворова означава једну грану.

Дефиниција 3.12. *Граф $G = (V, E)$ је* ***стабло (дрво)*** *ако је повезан и ако у њему не постоји циклус.*

За стабло се може рећи да је најједеноставнији повезан граф.

За писање овог поглавља кориштена је литература [1], [8].

Поглавље 4

Случајни графови

4.1 Почеци теорије случајних графова

Теорија случајних графова се бави асимптотским одликама графова при чему се ослања на расподјелу вјероватноћа. На примјер, она проучава начин на који структура компонената случајног графа еволуира када број његових грана расте.

Након што су мађарски математичари Пол Ердеш (*Paul Erdös*) и Алфред Рени (*Alfred Rényi*) [3] 1959. године дали темеље теорији случајних графова тако што су примјенили вјероватносне методе за откривања одређених својстава графова, бројни модели случајних графова су представљени и проучавани. Сматра се да су они поставили основу модерној теорији случајних графова, због чега је модел случајних графова познат под називом Ердеш-Ренијев (*Erdös-Rényi*-јев) модел.

Случајни графови су прво били кориштени за стицање увида у понашање графова и највише су примјењивани за рјешавање комбинаторних проблема. Тема се проширила на моделирање бројних мрежа које настају на случајан начин, као што су непредвидиво ширење Интернет мреже, распростирање људске популације и друштвених веза, неуронске мреже које су велики успјех доживјеле у машинском учењу и слично.

Појам **случајан граф** се први пут појавио у студијама Пола Ердеша, гдје је био употријебљен у доказу теореме која се односила на демонстрирање значаја вјероватносних метода за егзистенцију неких графова. Идеја на којој почива доказ лежи у разумјевању простора који садржи све могуће графове конструисане на случајан начин при чему они имају исте почетне карактеристике.

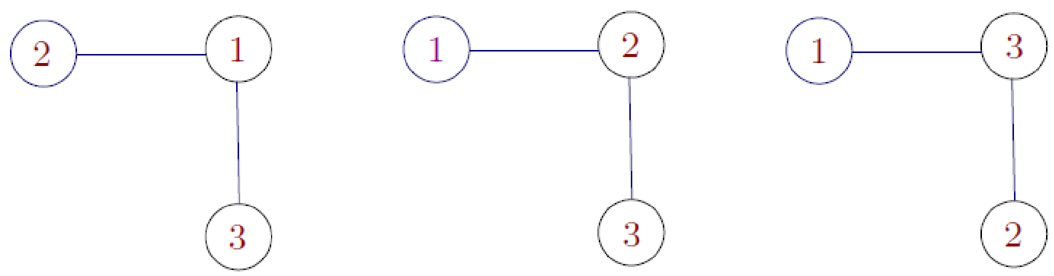
4.2 Вјероватносни простор случајних графова

Иако се назив **случајан граф** често придружује само једном конкретном графу чије су гране добијене на случајан начин, треба имати на уму да је овдје заправо ријеч о **вјероватносном простору** $\mathcal{G} = (\Omega, \mathcal{F}, P)$. Прецизније, Ω је **простор исхода** чији су елементи сви могући графови са n чворова који имају различит распоред и број грана у складу са расподјелом вјероватноћа и сваки од њих се може наћи иза уопштеног појма **случајан граф**. Број елемената скупа је $|\Omega| = 2^N$, гдје је $N = \binom{n}{2}$ максималан број грана који може имати граф са n чворова, уколико моделом није другачије дефинисано. Са P је означена **вјероватносна мјера**, а $\mathcal{F}$ означава **σ-алгебру** подскупова скупа Ω односно фамилију догађаја који се природно у овој области интерпретирају као **својства графа**. На примјер, нека A означава својство да је случајан граф повезан, онда важи

$$P(A) = \sum_{G \in A, G \in \mathcal{G}} P(G),$$

гдје се сумирају вјероватноће кроз све повезане графове у $\mathcal{G}$.

Слика 4.1 представља могуће графове вјероватносног простора графова који имају 3 чвора $V = \{1, 2, 3\}$ и 2 гране. Сваки граф се бира са вјероватноћом $\frac{1}{3}$, односно заступљена је униформна дискретна расподјела.



Слика 4.1: Графички приказ могућих случајних графова који имају 3 чвора и 2 гране

4.3 Модели случајних графова

Нека је $V = \{1, 2, ..., n\}$ скуп чворова графова који ће бити разматрани у наставку рада. Скуп свих таквих графова који имају n чворова означен је са $\mathcal{G}_n$.

Случајан граф са n чворова добија се на основу једног од подскупова скупа грана $E(K_n)$ комплетног графа K_n. Начин на који се подскуп бира зависи од модела случајног графа. Постоји више модела који се крију под појмом Ердеш-Ренијев граф. Апроксимативно они се сматрају еквивалентним, али сваки модел има неке карактеристичне особине.

4.3.1 Униформни случајан граф

Униформни случајан граф $\mathcal{G}_{n,M}$ који се често назива само Ердеш-Ренијев модел први је био проучаван. Овај модел обухвата све случајне графове са скупом чворова $V = \{1, 2, ..., n\}$ који имају тачно M грана изабраних на случајан начин. Ако се уведе ознака $N = \binom{n}{2}$, простор $\mathcal{G}_{n,M}$, $0 \leq M \leq N$ има $\binom{N}{M}$ елемената и сваки елемент односно случајни граф се појављује са истом вјероватноћом $\binom{N}{M}^{-1}$. Скоро увијек M је функција која зависи од броја чворова n, тј. $M = M(n)$.

На примјер, посматра се вјероватносни простор $\mathcal{G}_{4,3}$ свих случајних графова који имају 4 чвора и 3 гране. Нека је са $\Gamma \in \mathcal{G}_{4,3}$ означен граф са скупом грана $E(\Gamma) = \{\{1,2\},\{2,4\},\{3,4\}\}$. Онда је

$$P\{\mathcal{G}_{4,3} = \Gamma\} = \frac{1}{\binom{\binom{4}{2}}{3}} = \frac{1}{\binom{6}{3}} = \frac{1}{20}.$$

4.3.2 Биномни случајан граф

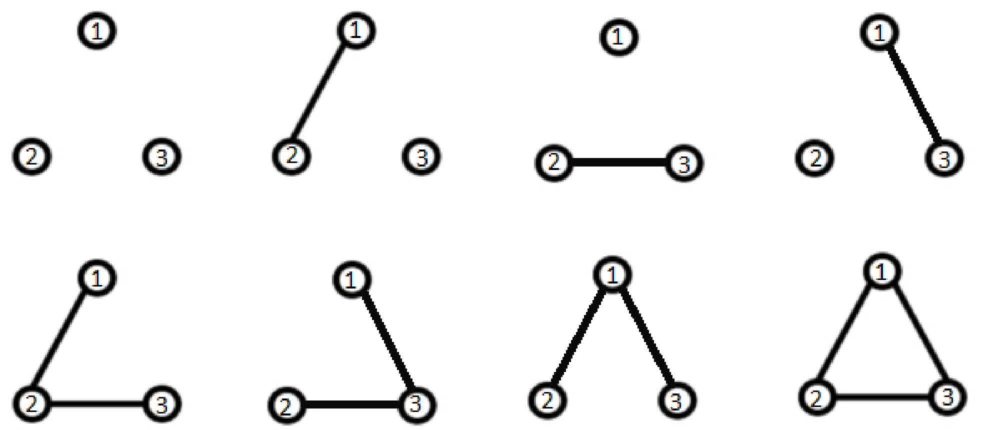
Биномни случајан граф превазилази један од проблема који се јављају у моделу униформног случајног графа, а то је да се гране не бирају независно једна од друге. Ово је омогућило лакши рад са случајним графовима, тј. случајне графове је још више приближило теорији вјероватноћа.

Биномни случајан граф познат и као Гилбертов модел $\mathcal{G}_{n,p}$, за $0 \leq p \leq 1$, садржи све графове са скупом чворова $V = \{1, 2, ..., n\}$ у коме се гране које повезују два чвора бирају независно са вјероватноћом p. Таквих графова има 2^N, при чему је $N = \binom{n}{2}$. Другим ријечима, ако је Γ_0 граф са скупом чворова V и M грана, онда важи

$$P(\{\Gamma_0\}) = P\{\mathcal{G}_{n,p} = \Gamma_0\} = p^M(1-p)^{N-M}.$$

Параметар p често упућује на **густину графа** $\mathcal{G}_{n,p}$.

На слици 4.2 приказани су сви графови са 3 чвора, чије се гране појављују на случајан начин. Таквих графова има $2^{\binom{3}{2}} = 8$.



Слика 4.2: Сви графови са 3 чвора, али са различитим распоредом и бројем грана

Тврђење 4.1. *Очекивани број грана у графу $\mathcal{G}_{n,\frac{1}{2}}$ је $\frac{N}{2}$, гдје је $N = \binom{n}{2}$.*

Доказ. *Нека су гране комплетног графа K_n са n чворова нумерисане од 1 до N. Случајна величина X_i је индикатор догађаја да је грана i, $1 \leq i \leq N$, садржана у графу $\mathcal{G}_{n,\frac{1}{2}}$, тј.*

$$X_i = \begin{cases} 1, & \text{ако граф } \mathcal{G}_{n,\frac{1}{2}} \text{ садржи грану } i \\ 0, & \text{иначе.} \end{cases}$$

Случајна величина $X = \sum_{i=1}^{N} X_i$ означава укупан број грана у случајном графу. Како је очекивање индикаторске случајне величине једнако вјероватноћи успјеха и на основу линеарности очекивања добија се

$$E[X] = \sum_{i=1}^{N} E[X_i] = \sum_{i=1}^{N} P\{X_i = 1\} = \sum_{i=1}^{N} \frac{1}{2} = \frac{N}{2}.$$

□

Природно уопштење модела $\mathcal{G}_{n,p}$ је модел $\mathcal{G}_{n,(p_{ij})}$, гдје је $0 \leq p_{ij} \leq 1$ за $1 \leq i < j \leq n$. Он садржи све графове са скупом чворова $V = \{1, 2, ..., n\}$ у којима се гране бирају независно, али не са истом вјероватноћом, већ за $1 \leq i < j \leq n$ вјероватноћа одабира гране која спаја чворове i и j је p_{ij}.

Специјалан случај $\mathcal{G}_{n,(p_{ij})}$ је $\mathcal{G}(H;p)$ гдје је H фиксиран граф и $0 < p < 1$. У овом случају бирају се гране графа H са вјероватноћом p независно једне од других, а гране које не припадају графу H не могу бити одабране. Ако је $V(H) = V$, а V је дефинисано као $V = \{1, 2, ..., n\}$, онда је вјероватносни простор $\mathcal{G}(H;p)$ идентичан простору $\mathcal{G}_{n,(p_{ij})}$ при чему је

$$p_{ij} = \begin{cases} p, & ij \in E(H) \\ 0, & \text{иначе.} \end{cases}$$

Такође, $\mathcal{G}(K_n, p)$ је исти вјероватносни простор као $\mathcal{G}_{n,p}$.

Простор $\mathcal{G}(K_{n,m}; p)$ садржи бипартитивне графове са скупом чворова U и W, такве да је $|U| = n$ и $|W| = m$, у којима су све $U - W$ гране изабране са вјероватноћом p.

Теорема 4.1. *Униформни случајан граф $\mathcal{G}_{n,M}$ је заправо биномни случајан граф $\mathcal{G}_{n,p}$ који има M грана.*

Доказ. *Нека је са $\mathcal{G}$ означен било који униформни случајан граф који има n чворова и M грана. При одређивању вјероватноће да је Гилбертов (биномни случајан) граф у ствари граф $\mathcal{G}$ са датих M грана добија се*

$$\begin{aligned} P\{\mathcal{G}_{n,p} = \mathcal{G} \,|\, |E(\mathcal{G}_{n,p})| = M\} &= \frac{P\{\mathcal{G}_{n,p} = \mathcal{G}\}}{P\{|E(\mathcal{G}_{n,p})| = M\}} \\ &= \frac{p^M (1-p)^{\binom{n}{2} - M}}{\binom{\binom{n}{2}}{M} p^M (1-p)^{\binom{n}{2} - M}} \\ &= \frac{1}{\binom{\binom{n}{2}}{M}}. \end{aligned}$$

□

У Ердеш-Ренијевом моделу број чворова и грана је фиксиран, док је у Гилбертовом моделу фиксиран само број чворова. Иако се оба модела често користе, може се рећи да Гилбертов модел преовладава у теорији случајних графова због своје веће флексибилности избора вриједности p, те и ближе везе са теоријом вјероватноћа. У литератури се неријетко и он среће под називом Ердеш-Ренијев модел. Овај модел је подложнији аналитичким истраживањима усљед независности успостављања грана у графу.

Ова два модела су у суштини еквивалентна када су параметри изабрани на одговарајући начин, тј. када је $M = p\binom{n}{2}$. Независно повезивање два чвора граном у случајаном биномном графу може се симулирати бацањем новчића, тако да је догађај добијања писма на примјер еквивалентан додавању гране у граф са вјероватноћом p. Очекивани број грана је $p\binom{n}{2}$. За $p = \frac{M}{\binom{n}{2}}$ очекиван број грана случајног биномног графа је M. Тада је интуитивно јасно да се ова два модела понашају слично, што ће касније бити размотрено.

4.3.3 Ердеш-Ренијев процес

Ердеш-Ренијев процес $\{\mathcal{G}_n(M) : M = 0, 1, ..., \binom{n}{2}\}$ или кратко $\mathcal{G}_n(M)$ почиње празним графом $\mathcal{G}_n(0)$ који има n изолованих чворова и нема гране, а завршава се комплетним графом који садржи све могуће гране. У сваком тренутку M, $1 \leq M \leq \binom{n}{2}$ у еволуирајући граф $\mathcal{G}_n(M-1)$ се додаје по једна нова грана изабрана на случајан начин и добија се нови граф $\mathcal{G}_n(M)$, који је у тренутку M расподијељен као униформни случајан граф $\mathcal{G}_{n,M}$.

Британски математичар Болобас [2] је дао идеју разматрања униформног случајног графа $\mathcal{G}_{n,M}$ као графа $\mathcal{G}_n(M)$ који настаје у тренутку M Ердеш-Ренијевог процеса.

Тврђење 4.2. *Ако је са* Γ *означен случајан граф који настаје Ердеш-Ренијевим процесом* $\{\mathcal{G}_n(M)\}$*, онда је* Γ *у тренутку* M *тог процеса еквивалентан униформном случајном графу* $\mathcal{G}_{n,M}$ *који има* M *грана.*

Доказ. *Нека је са* $\mathcal{G}_{n,M}$ *означен произвољан граф са* n *чворова и* M *грана. Потребно је израчунати вјероватноћу да је добијени граф* Γ *у* M*-том тренутку процеса изоморфан почетном графу* $\mathcal{G}_{n,M}$*. Све док првих* M *грана које су се појавиле до тог тренутка постоје и у графу* $\mathcal{G}_{n,M}$*, важи да је* $\mathcal{G}_{n,M} \cong \Gamma$*. Постоји* $M!$ *пермутација тих* M *грана и није потребно водити рачуна о уређењу преосталих грана које се појављују. Стога, постоји тачно* $M!(N-M)!$ *пермутација за које је* $\mathcal{G}_{n,M} \cong \Gamma$*. Тврђење важи јер сљедећа вјероватноћа задовољава услове униформног случајног графа*

$$P\{\mathcal{G}_{n,M} = \Gamma\} = \frac{M!(N-M)!}{N!} = \frac{1}{\binom{N}{M}}.$$

□

Ова врста процеса може се замислити као еволуција празног графа ка комплетном графу, која се дешава додавањем грана у складу са пермутацијом елемената N-точланог скупа. Његов значај је у томе што омогућава да се особине графа посматрају из корака у корак и на тај начин постигне структура која задовољава неке жељене одлике.

4.4 Расподјела вјероватноће графа $\mathcal{G}_{n,p}$

Просјечан број грана у графу $\mathcal{G}_{n,p}$ је $\binom{n}{2}p$, до кога се долази израчунавањем очекиване вриједности случајне величине X која представља број грана у графу $\mathcal{G}_{n,p}$ и има Биномну расподјелу са параметрима N и p. Вјероватноћа да граф

$\mathcal{G}_{n,p}$ има k грана дата је са

$$P\{X = k\} = \binom{N}{k} p^k (1-p)^{N-k}.$$

Степен чвора графа $\mathcal{G}_{n,p}$, такође, има Биномну расподјелу, али са параметрима $n-1$ и p. Ако је D_i случајна величина која означава степен чвора i, онда је вјероватноћа да је степен чвора i једнак k дата са

$$P\{D_i = k\} = \binom{n-1}{k} p^k (1-p)^{n-k-1},$$

јер сваки чвор може бити повезан са преосталих $n-1$ чворова, а вјероватноћа успјешног повезивања два чвора је p. Биномни коефицијент $\binom{n-1}{k}$ који броји начине на које може бити изабрано k грана ка осталим чворовима од могућих $n-1$, помножен је вјероватноћом да тих k изабраних грана постоји и вјероватноћом да преосталих $n-k-1$ грана не постоји. Када се посматра гранично понашање, често се $n-1$ замјени са n, па се узима да је очекиван степен чвора $d \approx pn$ умјесто $d \approx p(n-1)$.

За два различита чвора i и j случајне величине D_i и D_j нису потпуно независне. На примјер, ако је чвор i повезан са свим осталим чворовима тј. $D_i = n-1$, онда је очигледно $D_j \neq 0$ јер је сваки чвор осим i-тог повезан бар са i-тим па је $D_j > 0$. Међутим, за довољно велики n се може претпоставити да су оне независне и да се расподјела степена чвора случајног графа $\mathcal{G}_{n,p}$ приближава Пуасоновој. Како се Биномна расподјела може апроксимирати Пуасоновом расподјелом у случају када $np \longrightarrow \lambda$, може се задати да је $p = \dfrac{\lambda}{n}$ и исказати коначан резултат да степен чвора Ердеш-Ренијевог случајног графа има Пуасонову расподјелу са параметром λ, тј. за вјероватноћу да је степен чвора i једнак k важи:

$$P\{D_i = k\} = \frac{\lambda^k}{k!} e^{-k}.$$

Због овога се често Ердеш-Ренијев граф назива и Пуасонов граф.

Апроксимативно, Биномна расподјела брзо опада када је број експеримената већи од очекиване вриједности. Међутим, при истим условима расподјела степена чвора која се јавља у пракси није склона тако јаком паду. Због тога, иако је граф $\mathcal{G}_{n,p}$ важан математички модел, потребни су много комплекснији модели за представљање графова реалног свијета како би се обухватиле све њихове карактеристике.

На примјер, ако се посматра граф који описује авионску линију, степен његових чворова се креће у широком опсегу, од 1 или 2 за чворове који представљају мале градове, до 100 за главне чворове. Оваква расподјела степена чвора је Биномна.

Заједничка особина сложених мрежа које се јављају у пракси је присуство главних чворова, односно неколико чворова који су повезани са великим бројем других чворова у мрежи, те је њихов степен знатно већи у односу на степен осталих чворова. Углавном је то случај са мрежама које моделирају реалне проблеме, као на примјер мрежа телефонских позива. Такве мреже се називају *scale-free* мрежама и за њих важи нешто другачији закон расподјеле степена чвора (енг. *power law degree distribution*) по коме број чворова са датим степеном опада попут степене функције како степен расте, односно вјероватноћа да је степен чвора једнак k прати ткзв. *power law* расподјелу

$$P\{d(v) = k\} = c \cdot \frac{n}{k^{\gamma}},$$

гдје је γ позитивна, реална вриједност за коју се практично добија да је незнатно мања од 3 када је ријеч о сложеним мрежама.

Сљедећа теорема тврди да је степен чвора чврсто повезан са његовом очекиваном вриједношћу. Заправо, вјероватноћа да се степен чвора и његова очекивана вриједност np разликују за више да $\lambda\sqrt{np}$, опада експоненцијално брзо са порастом вриједности λ.

Теорема 4.2. *Нека је v чвор случајног графа $\mathcal{G}_{n,p}$. За $0 < \lambda < \sqrt{np}$ важи*

$$P\{|np - d(v)| \geq \lambda\sqrt{np}\} \leq 3e^{-\lambda^2/8}.$$

Претходна теорема важи само за један чвор, док се сљедећа посљедица односи на све чворове.

Посљедица 4.1. *Нека је ϵ позитивна константа. Ако је $p = \Omega(\frac{ln(n)}{n\epsilon^2})$ онда се скоро сигурно степен сваког чвора налази у интервалу $((1-\epsilon)np, (1+\epsilon)np)$.*

Претпоставка да је $p = \Omega(\frac{ln(n)}{n\epsilon^2})$ је неопходна. Ако је $p = \frac{d}{n}$, за d константно, онда степен неких чворова може бити ван наведеног интервала. За $p = \frac{1}{n}$ на основу посљедице, ниједан чвор нема степен већи од неке константе која је независна од n.

4.5 Анализа случајних графова

4.5.1 Коефицијент кластеровања

У теорији графова **коефицијент кластеровања** који се појављује и под називом индекс транзитивности представља мјеру која одређује колико су чворови у графу склони заједничком груписању у неке кластере. Након откривања овог својства, проучавање различитих графова сугерише да у већини случајева чворови теже креирању уско повезаних група карактеристичних по релативно високој густини повезаности p. Ове групе имају облик подграфа, па је вјероватноћа повезивања грана у оквиру њих углавном већа од просјечне вјероватноће успостављања случајне везе између два чвора почетног графа.

Посебно су социјалне мреже познате по високом коефицијенту кластеровања: ако су A и B пријатељи, као и B и C, онда је вјероватно да ће и A и C бити пријатељи. Зато се каже да коефицијент кластеровања мјери транзитивност графа.

Сложене мреже имају прилично висок број клика, које су заправо комплетни подграфови, односно у њима су свака два чвора спојена граном. Интуитивно, појединци у систему који имају нешто заједничко, образују групе. Ово својство графа се управо мјери коефицијентом кластеровања. То је мјера броја троуглова у графу, тј циклуса дужине 3.

Коефицијент кластеровања $\mathbb{C}$ се дефинише на 2 различита начина, као глобални и локални.

Глобални коефицијент кластеровања мјери удио сусједних парова грана таквих да све 3 гране које их повезују постоје:

$$\mathbb{C}_g = \frac{3 \cdot \textit{ број троуглова садржаних у графу}}{\textit{број парова сусједних грана}}.$$

У литератури се може наћи и сљедећа дефиниција која захтјева додатно објашњење

$$\mathbb{C}_{g,prob} = \frac{3 \cdot \textit{ број троуглова садржаних у графу}}{\textit{број повезаних тројки чворова}}.$$

Есенцијално, $\mathbb{C}$ треба да измјери средњу вриједност вјероватноће да су два сусједна чвора одређеног чвора (два која су повезана са истим чвором), такође, међусобно сусједни (повезани граном). Међутим, разматрајући граф K_3, по првој формули коефицијент кластеровања је очигледно 1, док се по другој формули у случају уређених тројки повезаних чворова добија $\frac{1}{2}$, а ако се посматрају неуређене тројке (тројке (1, 2, 3) и (3, 2, 1) су исте) добија се 3. Да би наведене формуле биле еквивалентне, повезана тројка се дефинише као уређен пар кога чине чвор и неуређен пар чворова који су повезани са чвором на првој координати, што би значило да су тројке (1, (2, 3)) и (1, (3, 2)) исте, те је у случају K_3 број повезаних тројки 3. Тада за различите парове тј. тројке исти чвор може бити бројан више пута.

Локални коефицијент кластеровања се дефинише као

$$\mathbb{C}_l(v) = \frac{\textit{број троуглова који садрже чвор } v}{\textit{број тројки са центром у } v}$$

или еквивалентно,

$$\mathbb{C}_l(v) = \frac{\textit{број грана које повезују сусједне чворове чвора } v}{\binom{d(v)}{2}},$$

гдје је $d(v)$ степен чвора v.

За граф $\mathcal{G}_{n,p}$ глобални коефицијент кластеровања може бити израчунат на основу локалног:

$$\mathbb{C}_g = \frac{\sum_{v=v_1}^{v_n} \binom{d(v)}{2} \cdot \mathbb{C}_l(v)}{\sum_{v=v_1}^{v_n} \binom{d(v)}{2}}$$

У наставку ознака $\mathbb{C}$ се односи на $\mathbb{C}_g$.

За израчунавање очекиване вриједности коефицијента кластеровања $\mathbb{C}$ претпоставља се да је вриједност p мала, тј. $p = p(n) \ll \frac{1}{n}$.
Нека је $\mathcal{T}$ скуп свих подграфова случајног графа $\mathcal{G}_{n,p}$ који су изоморфни комплентном графу K_3, чији чворови и гране образују троугао. Нека је $X_{\mathcal{T}}$ број таквих подграфова.

$$\mathcal{T} = \{H \subset \mathcal{G}_{n,p} | H \cong K_3\}$$

$$X_{\mathcal{T}} = |\mathcal{T}| = \sum_{1 \le i < j < k \le n} I\{(i, j, k) \in \mathcal{T}\}$$

$$E[X_{\mathcal{T}}] = E\left[\sum_{1 \le i < j < k \le n} I\{(i, j, k) \in \mathcal{T}\}\right] = \sum_{1 \le i < j < k \le n} P\{(i, j, k) \in \mathcal{T}\} = \binom{n}{3} p^3.$$

Број парова сусједних грана означен случајном величином $X_{\mathcal{W}}$ се израчунава на сличан начин. Нека је $\mathcal{W}$ скуп подграфова случајног графа $\mathcal{G}_{n,p}$ који су изоморфни случајном графу H са скупом чворова V и скупом грана E за које важи

$$V = \{v_1, v_2, v_3\}$$
$$E = \{(v_1, v_2), (v_1, v_3)\}$$

Онда је

$$\mathcal{W} = \{G \subset \mathcal{G}_{n,p} | G \cong H\}$$

$$X_{\mathcal{W}} = |\mathcal{W}|.$$

За скуп од 3 чвора $i_1 < i_2 < i_3$, дефинисана је случајна величина $\mathcal{W}_{i_1<i_2<i_3}$ као сума индикатора:

$$\mathcal{W}_{i_1<i_2<i_3} = I_{i_1} + I_{i_2} + I_{i_3}$$

гдје је, за $k = 1, 2, 3$

$$I_{i_k} = \begin{cases} 1, & \text{ако се у центру пара грана налази чвор } i_k \\ 0, & \text{иначе.} \end{cases}$$

Према томе,

$$X_{\mathcal{W}} = \sum_{1 \leq i_1 < i_2 < i_3 \leq n} \mathcal{W}_{i_1<i_2<i_3}$$

$$\mathbb{E}[X_{\mathcal{W}}] = \mathbb{E}\left[\sum_{1 \leq i_1 < i_2 < i_3 \leq n} \mathcal{W}_{i_1<i_2<i_3}\right] = \binom{n}{3} 3p^2.$$

Коначно, добија се очекивана вриједност коефицијента кластеровања:

$$\mathbb{E}[\mathbb{C}] = \frac{3 \cdot \mathbb{E}[X_{\mathcal{T}}]}{\mathbb{E}[X_{\mathcal{W}}]} = 3p.$$

Зависно од вриједности p за коју важи да $p(n) \longrightarrow 0$, када $n \longrightarrow \infty$, ова очекивана вриједност може бити јако мала. У реалним мрежама из праксе коефицијент кластеровања је углавном изнад 0,01. У моделу малог свијета ($small-world$ $model$) представљеног од стране $Watts$-а и $Strogatz$-а, јавља се особина високог коефицијента кластеровања која се комбинује са другим својствима. Иначе је низак коефицијент кластеровања један од разлога зашто су случајни графови лош модел за представљање комплексних мрежа, па се проналазе други математички модели који преузимају нека својства случајних графова.

4.5.2 Монотона својства

Дефиниција 4.1. *Подскуп Q скупа $\mathcal{G}_n$ се назива* ***својство*** *графа реда n ако $G \in Q$, $H \in \mathcal{G}_n$ и $G \simeq H$ повлачи $H \in Q$.*

Граф G има својство Q се записује и као $G \in Q$. Ако је Q неко својство, онда $P_M(Q)$ природно означава вјероватноћу припадања својства Q графу $\mathcal{G}_{n,M}$ и аналогно важи да $P_p(Q)$ означава вјероватноћу припадања својства Q графу $\mathcal{G}_{n,p}$.

Дефиниција 4.2. *Својство Q је* ***монотоно растуће*** *ако граф G има својство Q и $G \subset H$ тада важи да и H има својство Q тј. $H \in Q$.*

Другим ријечима, својство Q је монотоно растуће ако из чињенице да неки подграф графа G има својство Q слиједи да и цијели граф G има својство Q. Интуитивно, растуће својство остаје сачувано додавањем нових грана у граф, тј. када се својство једном појави у графу, више не може нестати. Оваква својства су од значаја јер се многи концепти случајних графова ослањају на њихово појављивање и присуство. Сличном логиком, супротно тврђење важи за монотоно опадајуће својство, које остаје сачувано ако се графу обришу неке гране.

Примјер растућег својства Q је својство да граф G *садржи k-циклус* или специјално за $k = 3$ својство да граф G *садржи троугао*, што ће бити касније показано. Ако подграф графа G *садржи k-циклус*, јасно је да онда граф G *садржи k-циклус.*
Још неки примјери растућег својства су: *фиксиран граф H је подграф у графу G, постојање огромних компонената у графу G, повезаност графа G, дијаметар графа G је највише d.*

Дефиниција 4.3. *Својство Q је* ***конвексно*** *ако $F \subset G \subset H$ и $F \in Q, H \in Q$ повлачи $G \in Q$.*

Растућа својства су дефинисана у терминима додавања грана у граф. Сљедећа теорема доказује да са порастом вјероватноће p у графу $\mathcal{G}_{n,p}$ расте и вјероватноћа да важи растуће својство, односно показује да случајан граф који има већу вјероватноћу p избора гране има и већу вјероватноћу да за њега важи растуће својство. Еквивалентно, додавање грана у граф ће вјероватније резултовати граф који има жељено растуће својство.

Теорема 4.3. *Нека је Q монотоно растуће својство, $0 \leq M_1 < M_2 \leq N$, $0 \leq p_1 < p_2 \leq 1$. Тада важи*

$$P_{M_1}(Q) \leq P_{M_2}(Q)$$

$$P_{p_1}(Q) \leq P_{p_2}(Q).$$

Доказ. (i) *Нека случајан граф формиран од M_1 грана има својство Q. Ако се приликом формирања другог графа изабере M_2 грана тако што се бира једна по једна грана и тај граф ће сигурно имати својство Q јер граф формиран од M_1 грана има то својство, а $M_1 < M_2$ и Q је монотоно растуће.*

(ii) Нека је $p = \frac{p_2 - p_1}{1 - p_1}$. Нека су $G_1 \in \mathcal{G}_{n,p_1}$ и $G \in \mathcal{G}_{n,p}$ два независно изабрана случајана графа и дефинисан је граф $G_2 = G_1 \cup G$. Онда су гране графа G_2 изабране независно и то са вјероватноћом $p_2 = p_1 + p - p_1 p$, па је граф G_2 елемент из $\mathcal{G}_{n,p_2}$. С обзиром на то да је Q монотоно растуће својство, ако G_1 има својство Q, онда то важи и за G_2. Према томе важи $P_{p_1}(Q) \leq P_{p_2}(Q)$.

□

4.5.3 Егзистенција троугла у графу

Из наведене теореме слиједи да је за монотоно растуће својство Q вјероватноћа $P_p(Q)$ монотона функција од p. Како се број чворова у графу повећава могло би се очекивати да расте и број могућих троуглова у графу, али то није случај. Наиме, својство Q које означава да граф $\mathcal{G}_{n,p}$ *не садржи троугао* важи са великом вјероватноћом, што значи да вјероватноћа супротног својства тј. да граф $\mathcal{G}_{n,p}$ *садржи троугао* тежи нули.

Нека је X_n број троуглова садржаних у графу $\mathcal{G}_{n,p}$. Максималан број троуглова у графу са n чворова је $\binom{n}{3}$. Нека су $Y_1, ..., Y_{\binom{n}{3}}$ случајне величине које детектују да ли граф $\mathcal{G}_{n,p}$ садржи i-ти могући троугао, тј. ако садржи онда индикатор Y_i узима вриједност 1, а иначе 0. Тада се X_n може записати као

$$X_n = \sum_{i=1}^{\binom{n}{3}} Y_i.$$

Вјероватноћа појављивање сваке гране је p, а усљед независности, вјероватноћа појављивања сваког троугла је p^3. Из дефиниције очекивања дискретне случајне величине и линеарности очекивања добија се да је

$$E[Y_i] = p^3, \text{ за } \forall i$$

$$E[X_n] = E\left[\sum_{i=1}^{\binom{n}{3}} Y_i\right] = \sum_{i=1}^{\binom{n}{3}} E[Y_i] = \binom{n}{3} p^3.$$

Ако се уведе параметризација $p = \dfrac{d}{n}$, за $d > 0$, онда вриједи

$$\lim_{n\to\infty} E[X_n] = \lim_{n\to\infty} \binom{n}{3} p^3 = \lim_{n\to\infty} \frac{n(n-1)(n-2)d^3}{6n^3} = \frac{d^3}{6}.$$

Може се доказати да низ случајних величина $\{X_n, n \in \mathbb{N}\}$ у расподјели конвергира Пуасоновој случајној величини X са параметром $\lambda = \dfrac{d^3}{6}$. Одатле слиједи

$$\lim_{n\to\infty} P\{\mathcal{G}_{n,p} \in Q\} = \lim_{n\to\infty} P\{X_n = 0\} = P\{X = 0\} = e^{\frac{-d^3}{6}}.$$

Ако је на примјер $p = \dfrac{10^6}{n}$ за свако довољно велико $n \in \mathbb{N}$ и како је

$$\lim_{d\to 0} e^{\frac{-d^3}{6}} = 1 \text{ и } \lim_{d\to\infty} e^{\frac{-d^3}{6}} = 0,$$

закључује се да је вјероватноћа да граф $\mathcal{G}_{n,p}$ *садржи троугао* јако мала, док је за $p = \dfrac{10^{-6}}{n}$ та вјероватноћа велика.

За $p = n^{-0.9}$, тј. $p \gg \frac{1}{n}$ закључује се да граф $\mathcal{G}_{n,p}$ са великом вјероватноћом *не садржи троугао*, тј. прецизније

$$\lim_{n\to\infty} P\{\mathcal{G}_{n,p} \in Q\} = 1.$$

Слично за $p \ll \frac{1}{n}$, на примјер $p = \frac{1}{n \cdot ln(n)}$, важи да граф са малом вјероватноћом садржи троугао. Ово је једна од првих представљених теорема из ове области.

4.5.4 Прелазне фунцкије за монотона својства

Да ли случајан граф $\mathcal{G}_{n,p}$ посједује неко својство може да зависи од вриједности вјероватноће настанка гране p, која у неким случајевима да би својство важило треба да задовољава одређен **праг** (енг. *threshold value*). Увођењем прага за p дефинише се граница у односу на коју својство важи у графу. Када се каже да граф скоро сигурно не садржи својство мисли се да је вјероватноћа да то својство припада графу блиска 0 када $n \longrightarrow \infty$. С обзиром на то да је интервал вриједности p углавном врло узак уводи се појам **прелазних функција**.

Дефиниција 4.4. *Реална функција $p^* = p^*(n)$ таква да $p^*(n) \longrightarrow \infty$ када $n \longrightarrow \infty$, назива се **прелазна функција** за својство Q графа $\mathcal{G}_{n,p}$ ако важи*

$$\lim_{n\to\infty} P\{\mathcal{G}_{n,p} \in Q\} = \begin{cases} 0, & \text{ако } \lim_{n\to\infty} \frac{p}{p^*(n)} = 0 \\ 1, & \text{ако } \lim_{n\to\infty} \frac{p}{p^*(n)} = \infty \end{cases}$$

Иста дефиниција важи и када је Q монотоно растуће својство, док у случају монотоно опадајућег својства услови о односу p и p^* у оквиру граничне вриједности вјероватноће су супротна.

Ако постоји прелазна функција $p^*(n)$ у односу на коју се скоро сигурно неко својства јавља у графу (или се губи), онда долази до **фазних прелаза**.

Сљедећу теорему која се односи на прелазне функције за монотона својства, изложили су Болобас [2] и Томасон.

Теорема 4.4. *Свако нетривијално монотоно својство има прелазну функцију.*

При одређивању прелазне функције за својство Q које важи за случајан граф $\mathcal{G}$ вјероватносног простора $\mathcal{G}_{n,p}$, корисно је посматрати то својство на сљедећи начин

$$Q = \{\mathcal{G} = (V, E) : X(\mathcal{G}) \geq 1\}$$

гдје је X одговарајућа ненегативна случајна величина дефинисана на вјероватносном простору $\mathcal{G}_{n,p}$. На примјер, ако се то својство односи на класу графова који садрже циклус, онда X може да означава број циклуса у графу $\mathcal{G}$. У већини случајева X се јасно може дефинисати из контекста и то је најчешће сума индикаторских случајних величина.

Прво, када је p мало у поређењу са p^* треба показати да асимптотски скоро сигурно граф не садржи својство Q, а затим када је p велико у поређењу са p^* треба показати да граф садржи то својство, тј.

$$\lim_{n\to\infty} P\{\mathcal{G}_{n,p} \in Q\} = \begin{cases} 0, & \text{ако } p = o(p^*) \\ 1, & \text{ако } p^* = o(p). \end{cases}$$

Еквивалентно, доказ да је p^* прелазна функција може се добити показујући да ако је p мало у прећењу са p^* онда је асимптотски скоро сигурно $X = 0$ и ако је p велико у прећењу са p^* онда је асимптотски скоро сигурно $X > 0$.

X је ненегативна случајна величина, тј. $X \geq 0$ па се овдје може користити Марковљева неједнакост за први дио доказа, односно метода првог момента описана у уводном дијелу. Ако је $E[X] = o(1)$ онда се добија да је $P\{X > 0\} = o(1)$ што значи да је $X = 0$ скоро сигурно.

За други дио доказа треба показати да је асимптотски скоро сигурно $X > 0$. Ако је $\frac{D[X]}{(E[X])^2} = o(1)$ може се примјенити посљедица 2.3 па је асимптотски скоро сигурно $X > 0$. Стога је довољно доказати $\frac{D[X]}{(E[X])^2} = o(1)$.

Једноставан примјер, али добар за илустрацију наведеног, би био да Q означава својство да граф $\mathcal{G} \in \mathcal{G}_{n,p}$ *има најмање 1 грану*, што се може записати као

$$Q = \{\mathcal{G} = (V, E) : X(\mathcal{G}) \geq 1\}$$

при чему је $X(\mathcal{G}) = |E(\mathcal{G})|$. Гране комплетног графа K_n се могу нумерисати $e_1, ...e_N$, па се случајна величина X може изразити као сума $N = \binom{n}{2}$ независних индикаторских случајних величина, тј.

$$X = \sum_{i=1}^{N} X_i, \text{ гдје је}$$

$$X_i = \begin{cases} 1, & \text{ако граф } \mathcal{G} \text{ садржи грану } e_i, \\ 0, & \text{иначе.} \end{cases}$$

Очекивање је дато са

$$E[X] = \sum_{i=1}^{N} E[X_i] = \sum_{i=1}^{N} p = Np = \binom{n}{2} p = \frac{n! \cdot p}{2!(n-2)!} = \frac{n(n-1)p}{2} = \Theta(n^2 p).$$

Ако је $p \ll \frac{1}{n^2}$ онда $E[X] \longrightarrow 0$, када $n \longrightarrow \infty$.
С друге стране, ако је $p \gg \frac{1}{n^2}$ тада $E[X] \longrightarrow \infty$, када $n \longrightarrow \infty$. Кандидат за прелазну функцију за својство Q је $p^*(n) = \frac{1}{n^2}$. Сада је потребно размотрити два случаја у складу са дефиницијом 4.4.

Прво, нека је $0 \leq p = p(n) \leq 1$ и $\frac{p}{p^*} \longrightarrow 0$ када $n \longrightarrow \infty$. Због $\frac{p}{p^*} = n^2 p$ и $E[X] = \Theta(n^2 p)$ на основу тога што $\frac{p}{p^*} \longrightarrow 0$ може се закључити да $E[X] \longrightarrow 0$ када $n \longrightarrow \infty$. Користећи Марковљеву неједнакост добија се

$$P\{X > 0\} \leq E[X].$$

Према томе, $P\{X > 0\} \longrightarrow 0$ када $n \longrightarrow \infty$, па је асимптотски скоро сигурно $X = 0$ и као посљедица тога $\mathcal{G}_{n,p} \notin Q$.

У другом случају важи да је $0 \leq p \leq 1$ и $\frac{p}{p^*} \longrightarrow \infty$ када $n \longrightarrow \infty$. Како је $\frac{p}{p^*} = n^2 p$ и $E[X] = \Theta(n^2 p)$ може се закључити да $E[X] \longrightarrow \infty$ када $n \longrightarrow \infty$. Како је случајна величина X сума N индикаторских случајних величина, њој одговара Биномна расподјела па је

$$D[X] = Np(1-p).$$

Тада је

$$\frac{D[X]}{(E[X])^2} = \frac{Np(1-p)}{(Np)^2} = \frac{1-p}{Np} \leq \frac{1}{Np} = \Theta\left(\frac{1}{n^2 p}\right),$$

а због $n^2 p \longrightarrow \infty$ када $n \longrightarrow \infty$, закључује се да је

$$\frac{D[X]}{(E[X])^2} = o(1).$$

Коначно, из посљедице 2.3 (метода другог момента) слиједи да је $P\{X = 0\} = o(1)$ и према томе $P\{X > 0\} \longrightarrow 1$ када $n \longrightarrow \infty$, па асимптотски скоро сигурно важи да је $X \geq 1$ и као посљедица тога $\mathcal{G}_{n,p} \in Q$.

У претходно разматраном примјеру својства да граф *не садржи троугао*, функција $p^*(n) = \dfrac{1}{n}$ је прелазна функција за то својство.

Може се примјетити да прелазна функција, ако постоји за неко својство, није јединствена. На примјер, за исто својство да граф *не садржи троугао* функција $p^* = \dfrac{10}{n}$ је, такође, прелазна.

4.5.5 Повезаност

Једно од растућих својстава која зависе од прелазне функције је повезаност графа. Може се дефинисати вриједност p тако да у зависности од ње са великом вјероватноћом граф $\mathcal{G}_{n,p}$ буде повезан или не:

Теорема 4.5. *За свако $\epsilon > 0$ важи*

$$P\{\mathcal{G}_{n,p} \text{ је повезан}\} \longrightarrow \begin{cases} 0, & \text{ако } p \leq \dfrac{(1-\epsilon)ln(n)}{n} \\ 1, & \text{ако } p \geq \dfrac{(1-\epsilon)ln(n)}{n}. \end{cases}$$

Сљедећа теорема обезбјеђује прелазну функцију за коју важи својство *да не постоје изоловани чворови* у графу $\mathcal{G}_{n,p}$, односно граф је скоро увијек повезан.

Теорема 4.6. *Нека је $p = \dfrac{c \cdot ln(n)}{n}$. Ако је $c \geq 3$ и $n \geq 100$ онда важи*

$$P\{\mathcal{G}_{n,p} \text{ је повезан}\} \longrightarrow 1$$

Егзактнији резултат у смислу теорије дат је наредном теоремом:

Теорема 4.7. *Нека је $p = \dfrac{c \cdot ln(n)}{n}$. Ако је $c > 1$ онда је граф $\mathcal{G}_{n,p}$ повезан са великом вјероватноћом.*

Теорема 4.8. *Нека је $p(n) = \dfrac{ln(n) + c + o(1)}{n}$. Тада важи*

$$P\{\mathcal{G}_{n,p} \text{ је повезан}\} \longrightarrow e^{-e^{-c}}$$

Специјално за $p(n) = \dfrac{ln(n)}{n}$ ова вјероватноћа тежи e^{-1}.

4.5.6 Одсуство изолованих чворова у графу $\mathcal{G}_{n,p}$

Прелазна функција за својство одсуства изолованих чворова у графу $\mathcal{G}_{n,p}$ је $\frac{ln(n)}{n}$. Одсуством изолованих чворова граф постаје повезан.

Нека је X број изолованих чворова у графу $\mathcal{G}_{n,p}$. Тада важи

$$E[X] = n(1-p)^{n-1}$$

Пошто је претпоставка да је функција прага $\dfrac{ln(n)}{n}$, нека је $p = c\dfrac{ln(n)}{n}$, па се добија

$$\lim_{n \longrightarrow \infty} E[X] = \lim_{n \longrightarrow \infty} n\left(1 - \frac{c \cdot ln(n)}{n}\right)^n = \lim_{n \longrightarrow \infty} ne^{-cln(n)} = \lim_{n \longrightarrow \infty} n^{1-c}$$

Ако је $c > 1$ очекивани број изолованих чворова тежи 0, а то значи да скоро сви могући графови немају изоловане чворове.
Ако је $c < 1$ очекивани број изолованих чворова тежи ∞ и тада се методом другог момента доказује да скоро сви графови имају изолован чвор, па се може записати да је

$$X = I_1 + I_2 + ... + I_n$$

гдје је I_i индикатор који указује да ли је чвор i изолован. Онда је

$$E[X^2] = \sum_{i=1}^{n} E[I_i^2] + 2\sum_{i<j} E[I_i I_j]$$

Пошто је $I_i^2 = I_i$, прва сума је једнака $E[X]$. У другој суми сви елементи су исти па се добија

$$E[X^2] = E[X] + 2\binom{n}{2}E[I_1 I_2] = E[X] + n(n-1)E[I_1 I_2] =$$

$$= E[X] + n(n-1)(1-p)^{n-1}(1-p)^{n-2} = E[X] + n(n-1)(1-p)^{2(n-1)-1}.$$

На крају се у експоненту одузима 1 јер је вјероватноћа да прва два чвора нису повезена већ укључена вјероватноћом да је први чвор изолован. Сада,

$$\frac{E[X^2]}{(E[X])^2} = \frac{n(1-p)^{n-1} + n(n-1)(1-p)^{2(n-1)-1}}{n^2(1-p)^{2(n-1)}} = \frac{1}{n(1-p)^{n-1}} + \left(1 - \frac{1}{n}\right)\frac{1}{1-p}$$

За $p = c\dfrac{ln(n)}{n}$, $c < 1$ и $\lim\limits_{n \longrightarrow \infty} E[X] = \infty$ важи

$$\lim_{n \longrightarrow \infty} \frac{E[X^2]}{(E[X])^2} = \lim_{n \longrightarrow \infty} \left[\frac{1}{n^{1-c}} + (1 - \frac{1}{n})\frac{1}{1 - c\frac{ln(n)}{n}}\right] = 1 + o(1)$$

На основу посљедице 2.4, вјероватноћа да је $X = 0$ тежи 0, што имплицира да скоро сви графови имају изолован чвор. Према томе, $\frac{ln(n)}{n}$ је прелазна функција за својство одсуства изолованог чвора.

4.5.7 Егзистенција циклуса у графу $\mathcal{G}_{n,p}$

Када је ријеч о проблему формирања циклуса у случајном графу посматра се сваки подскуп од k чворова и испитује се под којим условима они формирају или циклус или повезану компоненту.
Праг за својство присуства циклуса у случајном графу $\mathcal{G}_{n,p}$ је $\frac{1}{n}$, што се може показати примјеном метода првог и другог момента.

Нека је X број циклуса у графу $\mathcal{G}_{n,p}$. Формирање циклуса дужине k могуће је на $\binom{n}{k}$ начина. Датих k чворова може бити уређено произвољним избором првог чвора, затим избором другог од преосталих $k-1$, и тако редом. Како су почетак и крај циклуса спојени, циклус и његова обрнута варијанта се сматрају истим, па је број начина потребно подијелити са 2:

$$E[X] = \sum_{k=3}^{n} \binom{n}{k} \frac{(k-1)!}{2} p^k \leq \sum_{k=3}^{n} \frac{n^k p^k}{2k} \leq \sum_{k=3}^{n} n^k p^k = (np)^3 \frac{1-(np)^{n-2}}{1-np} \leq 2(np)^3$$

с тим да је $np < \frac{1}{2}$.
Када је p асимптотски мање од $\frac{1}{n}$, онда

$$\lim_{n\longrightarrow\infty} np = 0 \text{ и } \lim_{n\longrightarrow\infty} \sum_{k=3}^{n} (np)^k = 0.$$

Дакле, када $n \longrightarrow \infty$ онда $E[X] \longrightarrow 0$, па се на основу методе првог момента описане у уводу закључује да граф скоро сигурно нема циклусе. Помоћу методе другог момента може се показати да за $p = \frac{d}{n}$, $d > 1$, граф има циклус са вјероватноћом која тежи 1.
Претходно наведено не даје строг праг јер је тврђено да $E[X] \longrightarrow 0$ само под претпоставком да је p асимптотски мање од $\frac{1}{n}$. Строг праг захтјева да $E[X] \longrightarrow 0$ за $p = \frac{d}{n}$, $d < 1$.
У том случају за $p = \frac{d}{n}$, $d = const.$ важи

$$E[X] = \sum_{k=3}^{n} \binom{n}{k} \frac{(k-1)!}{2} p^k = \frac{1}{2} \sum_{k=3}^{n} \frac{n(n-1)(n-2)...(n-k+1)}{k!} (k-1)! \cdot p^k$$

$$= \frac{1}{2} \sum_{k=3}^{n} \frac{n(n-1)(n-2)...(n-k+1)}{n^k} \frac{d^k}{k}.$$

$E[X]$ конвергира ако је $d < 1$, а дивергира ако је $d \geq 1$.
Ако је $d < 1$, онда је $E[X] \leq \frac{1}{2} \sum_{k=3}^{n} \frac{d^k}{k}$ и гранична вриједност $E[X]$ тежи позитивној константи.

Ако је $d = 1$, онда је

$$E[X] = \frac{1}{2}\sum_{k=3}^{n} \frac{n(n-1)...(n-k+1)}{n^k}\frac{1}{k}.$$

Ако се у обзир узме само првих $ln(n)$ чланова суме, због $\frac{n}{n-i} = 1 + \frac{i}{n-i} \leq e^{\frac{i}{n-i}}$ слиједи да је $\frac{n(n-1)...(n-k+1)}{n^k} \geq \frac{1}{2}$. Стога,

$$E[X] \geq \frac{1}{2}\sum_{k=3}^{ln(n)} \frac{n(n-1)...(n-k+1)}{k \cdot n^k} \geq \frac{1}{4}\sum_{k=3}^{ln(n)} \frac{1}{k}.$$

$$\lim_{n\longrightarrow\infty} E[X] \geq \lim_{n\longrightarrow\infty} \frac{1}{4}\sum_{k=3}^{ln(n)} \frac{1}{k} \geq \lim_{n\longrightarrow\infty} ln(ln(n)) = \infty.$$

За $p = \frac{d}{n}$, $d < 1$, $E[X]$ конвергира константи различитој од 0 и са позитивном вјероватноћом графови ће имати константан број циклуса независно од величине графа.
За $d > 1$, $E[X]$ конвергира бесконачности, те по методи другог момента графови ће имати неограничен број циклуса који расте са n.

4.5.8 Асимптотска еквивалентност модела случајних графова

Прије навођења тврђења о асимптотској еквивалентности наведених модела, наводе се сљедећа два тврђења која пореде вјероватноће да неко својсво важи за дате моделе.

Тврђење 4.3. *Нека је Q неко својство графа $\mathcal{G}$, $p = \frac{M}{\binom{n}{2}}$, гдје је $M = M(n)$, и $\binom{n}{2} - M \longrightarrow \infty$. Тада асимптотски важи*

$$P\{\mathcal{G}_{n,M} \in Q\} \leq \sqrt{2\pi M} P\{\mathcal{G}_{n,p} \in Q\}$$

Тврђење 4.4. *Нека је Q монотоно растуће (опадајуће) својство графа $\mathcal{G}$, $p = \frac{M}{\binom{n}{2}}$. Тада асимптотски важи*

$$P\{\mathcal{G}_{n,M} \in Q\} \leq 3P\{\mathcal{G}_{n,p} \in Q\}.$$

Како је раније поменуто, корисно је примјетити да у већини случајева модели $\mathcal{G}_{n,M}$ и $\mathcal{G}_{n,p}$ могу практично замјенити један другог, под условом да је M близу $p\binom{n}{2}$. При анализи графа $\mathcal{G}_{n,p}$, ако се постави да је $p = \frac{M}{\binom{n}{2}}$, очекивани број грана у биномном случајном графу $\mathcal{G}_{n,p}$, као и у графу $\mathcal{G}_{n,M}$ једнак је M. За било које монотоно својство конвергенција графа $\mathcal{G}_{n,M}$ имплицира конвергенцију $\mathcal{G}_{n,p}$ ка истој граничној вриједности.

Тврђење 4.5. *Нека је Q произвољно својство, $p = p(n) \in [0,1]$ и $0 \leq a \leq 1$. Ако за сваки низ цијелих броја $M(n)$, $0 \leq M(n) \leq \binom{n}{2}$, такав да*

$$M = p\binom{n}{2} + O\left(\sqrt{\binom{n}{2}p(1-p)}\right) \text{ важи}$$

$$P\{\mathcal{G}_{n,M} \in Q\} \longrightarrow a, \text{ када } n \longrightarrow \infty, \text{ онда важи и да}$$

$$P\{\mathcal{G}_{n,p} \in Q\} \longrightarrow a, \text{ када } n \longrightarrow \infty.$$

4.5.9 Егзистенција k-клике у случајном графу

Као што је већ показано постоји праг за p тако да када $n \longrightarrow \infty$ случајан граф $\mathcal{G}_{n,p}$ скоро сигурно садржи троугао. Овај концепт може бити проширен проналаском вриједности p тако да са великом вјероватноћом граф $\mathcal{G}_{n,p}$ скоро сигурно или садржи или не садржи скуп од k чворова, таквих да су сви међусобно повезани, што се у литератури назива **k-клика**. Ова идеја мотивисана је претпоставком да додавањем нових грана у случајан граф, очекивано настају и нове клике, док у случају графа чије се гране појављују са малом вјероватноћом не постоји велики број клика. Од интереса је граница вриједности p у односу на коју се посматра егзистенција $k-$клике у графу. Јасно је да се за $p = 1$ може наћи било која клика величине мање или једнако n, али од интереса је најмања вриједност p за коју постоји $k-$клика у графу.

Теорема 4.9. *Нека је K_r комплетан граф са r чворова, $r \geq 3$. За вјероватноћу припадања $r-$клике графу $\mathcal{G}_{n,p}$ важи*

$$P(\mathcal{G}_{n,p} \supseteq K_r) \longrightarrow \begin{cases} 0, & \text{ако } p \ll n^{-2/(r-1)} \\ 1, & \text{ако } p \gg n^{-2/(r-1)} \end{cases}$$

Доказ ове теореме који је дао британски математичар Болобас [2] изложен је у раду [7].

У поређењу са својством повезаности интересанто је уочити да је граница вриједности p за егзистенцију k-клике асимтотски мања или једнака од функције облика n^{-x} за разлику од границе у Теореми 4.5, гдје је она дискретна функција која зависи од ϵ и $\frac{ln(n)}{n}$. Ово сугерише да је природа случајних графова *да буду повезани* осјетљивија када је у питању вриједност p од *својства да посједују подграф*.

Познавање својстава одређених подграфова је значајно у теорији графова и комбинаторици, као и у њиховој примјени. Модели случајних графова који генеришу обимне или честе подграфове (клике) су погодни за моделовање социјалних мрежа, као што су људске организације које се природно темеље на међусобно повезаним групама пријатеља или познаника.

4.5.10 Настајање K_4 клике у случајном графу $\mathcal{G}_{n,p}$

Нека је K_4 комплетан граф који има 4 чвора. У графу K_4 постоји $\binom{4}{2} = 6$ грана. Случајна величина X представља број графова облика K_4 коју су садржани у случајном графу $\mathcal{G}_{n,p}$. Биће показано да је одговарајућа прелазна функција $p^*(n) = n^{-2/3}$. Очекивана вриједност случајне величине X је

$$E[X] = \binom{n}{4} p^6 \sim n^4 p^6.$$

Ако је $p \ll p^*$ или еквивалентно $p = \dfrac{p^*}{\omega(n)}$, гдје $\omega(n)$ функција за коју важи

$$\lim_{n\to\infty} \omega(n) = \infty,$$

онда за очекивање случајне величине X важи

$$E[X] = \binom{n}{4} p^6 = \binom{n}{4} \frac{p^{*6}}{\omega(n)^6} = o(1).$$

Дакле, на основу метода првог момента може се закључити да када је $p \ll n^{-2/3}$ онда

$$P\{X = 0\} \leq E[X] = o(1),$$

или еквивалентно $X = 0$ са великом вјероватноћом.
Када је $p^* \ll p$ или еквивалентно $p = p^* \omega(n)$, онда је $X > 0$ са великом вјероватноћом. Очекивана вриједност броја K_4 подграфова тежи ∞, наиме

$$E[X] = \binom{n}{4} p^6 = \binom{n}{4} p^{*6} \omega(n)^6 = \Theta(\omega(n)^6) \longrightarrow \infty.$$

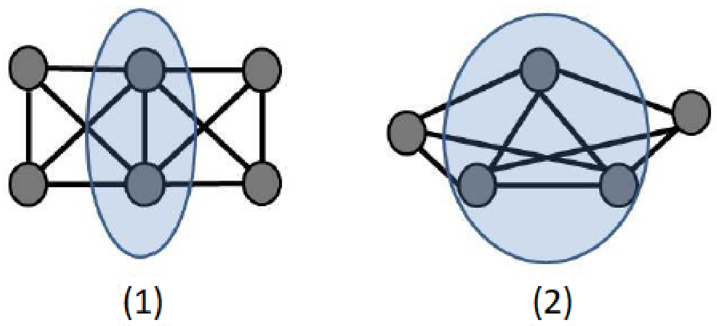
Међутим, из тога се не може закључити да је $X > 0$ са великом вјероватноћом. Овдје се примјењује метод другог момента. Прво, нека је X_i индикаторска промјенљива за присуство i-тог графа K_4 у графу, $i = 1, ..., \binom{n}{4}$. Случајна величина X која је дефинисана као број K_4 графова у графу $\mathcal{G}_{n,p}$ може се записати као

$$X = X_1 + X_2 + ... + X_{\binom{n}{4}}.$$

Када два графа K_4 не дијеле ни једну грану, онда су одговарајући индикатори независни, тј.

$$Cov[X_i, X_j] = P\{X_i = X_j = 1\} - P\{X_i\}P\{X_j\} = p^{12} - p^6p^6 = 0$$

Еквивалентно, то је случај када се два графа K_4 укрштају тј. сијеку у једном или ниједном чвору. Преостају 2 случаја, јер кардиналност скупа пресјека осим 0 и 1, може још бити 2 и 3. Та два случаја која су од интереса су приказана на слици, при чему су наглашени пресјеци два K_4 подграфа.



Слика 4.3: Два случаја пресјека K_4 подграфова

У првом случају два K_4 графа имају 2 заједничка чвора, односно 1 заједничку грану. Тада је укупан број грана 11, а оба индикатора узимају вриједност 1. Због тога се добија да је коваријација

$$Cov[X_i, X_j] = p^{11} - p^{12}.$$

Слично, у другом случају два K_4 графа имају 3 заједничка чвора и 3 заједничке гране. Тада је укупан број грана 9, па се добија да је коваријација

$$Cov[X_i, X_j] = p^9 - p^{12}.$$

Треба пребројати број комбинација које се појављују у односу на избор чворова у ова два случаја. У првом случају бира се 6 од n чворова, па има $\binom{n}{6}$ комбинација, и за сваку комбинацију бирају се 2 заједничка чвора, а затим и распоред преосталих. У другом случају бира се 5 од n чворова, а затим 3 заједничка чвора и слично.
Користећи методу другог момента објашњену у уводном дијелу, добија се

$$D[X] \leq E[X] + \sum_{i \neq j} Cov[X_i, X_j] \leq \binom{n}{4}p^6 + \binom{n}{6}\binom{6}{2}\binom{4}{2}p^{11} + \binom{n}{5}\binom{5}{3}\binom{2}{1}p^9 =$$

$$= o(n^8p^{12}) = o((E[X])^2).$$

Ово показује да је $X > 0$ са великом вјероватноћом када је $p^* \ll p$.

4.5.11 Највећа компонента случајног графа

Једно од најупечатљивих открића које су дали Ердеш и Рени односи се на присуство највеће компоненте у моделу случајног графа. Ово се доводи у везу са феноменом фазног прелаза, који се у овом случају односи на ред највеће компоненте. Укратко, ред компоненте у Ердеш-Ренијевом графу се мијења од логаритамског до линеарног када очекивани степен чвора са вриједности нешто мање од 1 прелази изнад 1 (на примјер од 0.99 до 1.01) приликом додавања грана. Према томе, сви случајни графови се понашају слично у смислу да постоји неколико компонената графа које су мале величине, и највећа компонента чија је величина по дефиницији већа од $n^{2/3}$.

Током процеса настанка графа, након одређеног времена једна компонента почиње да доминира у смислу своје величине, док друге компоненте расту знатно слабије. Испоставља се да када је ријеч о случајном графу који је тек у фази настанка, тачније када је просјечан степен одређеног чвора графа мањи од 1, раст свих компонената графа је реда $\mathcal{O}(ln(n))$. Међутим, како просјечан степен постаје већи од 1, компоненте почињу да расту линеарно.

Може се увести параметризација

$$M = \frac{dn}{2},$$

$$p = \frac{d}{n-1}$$

тако да d означава очекиван степен чвора случајног графа и може се узети у обзир када се разматрају својства која важе за граф са великом вјероватноћом, односно вјероватноћа тежи 1 када за број чворова важи $n \longrightarrow \infty$.

Прецизније, ако је у Ердеш-Ренијевом графу са $d > 0$ означен очекивани степен чвора и d је мање од критичне вриједности 1, случајан граф са великом вјероватноћом садржи стабла и нецикличне компоненте и тада највећу компоненту представља стабло реда $\mathcal{O}(ln(n))$. Такве компоненте се називају мале компоненте. С друге стране, ако је d веће од 1, са великом вјероватноћом постоји јединствена највећа компонента која је линеарног реда, док су све друге највеће компоненте нецикличне (стабло) реда $\mathcal{O}(ln(n))$. Другим ријечима, све компоненте осим највеће су мале. Ако је очекивани степен чвора d баш једнак критичној вриједности 1, ред највеће компоненте је реда већег од $n^{2/3}$.

Један од приступа разматрања највеће компоненте случајног графа је хеуристички. Претходно је показано да у случају када је $p = o(\frac{1}{n})$ не постоји ниједан троугао у случајном графу. То значи да су чворови углавном изоловани ($d = 0$) или образују стабла. Такође, наведено је да је за $p > \frac{c \cdot ln(n)}{n}$, $c \geq 1$ случајан граф повезан. Поставља се питање шта се дешава између ових фаза. Испоставља се да се јединствена највећа компонента појављује када је $p = \frac{c}{n}$ за $c > 1$. Највећа компонента графа се дефинише као компонента реда $\mathcal{O}(n)$.

Нека је u вјероватноћа да случајно изабран чвор не припада највећој компоненти. Било који други чвор може да буде садржан у највећој компоненти или да не буде, односно ова два чвора могу да буду неповезани или повезани. Ако други чвор припада највећој компоненти онда они нису повезани, а вјероватноћа овог догађаја је $1 - p$, а ако се други чвор не налази у највећој компоненти повезани су са вјероватноћом up. Након одабира првог, преостаје $n-1$ чворова, стога је

$$u = (1 - p + up)^{n-1}.$$

За уведену параметризацију $p = \dfrac{d}{n}$, за $d > 0$, гдје је d просјечан степен чвора, добија се

$$u = \left(1 - \frac{d}{n} + u\frac{d}{n}\right)^{n-1} = \left(1 - \frac{d}{n}(1-u)\right)^{n-1}$$

$$ln(u) = (n-1) \cdot ln\left(1 - \frac{d}{n}(1-u)\right)$$

Користећи $ln(1 + x) \approx x$ за мале вриједности x, добија се

$$ln(u) \approx -\frac{(n-1) \cdot d \cdot (1-u)}{n}$$

$$\Longrightarrow u \approx e^{-d(1-u)}$$

Коначно, вјероватноћа појављивања чвора у највећој компоненти $v = 1-u$ дата је једначином

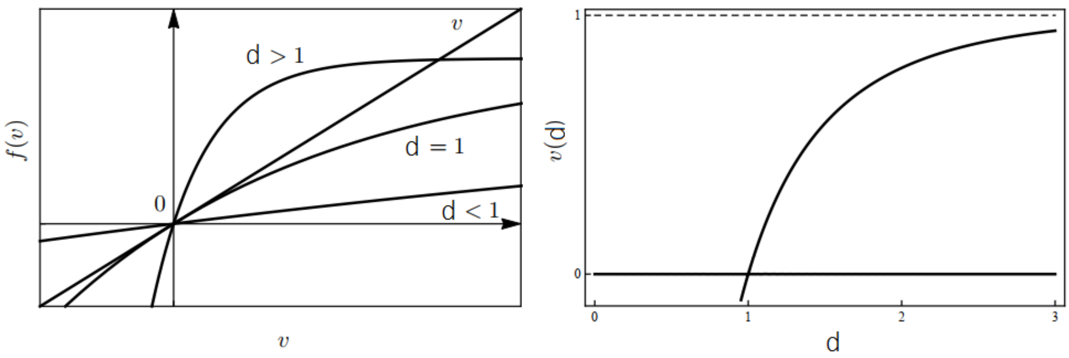
$$1 - v = e^{-d \cdot v}$$

Ова једначина увијек даје рјешење $v = 0$. Међутим, то није једино рјешење за све могуће вриједности d. Ово се може видјети посматрањем графикона двије функције $f(v) = 1 - e^{-dv}$ и $g(v) = v$. Њихови пресјеци дају рјешења почетне једначине. Сијеку се у 0 јер $f(0) = g(0) = 0$ као што је очекивано.

$$f'(v) = de^{-d \cdot v} > 0,$$

$$f''(v) = -d^2 e^{-d \cdot v} < 0$$

Слика 4.4 даје графиконе тих функција и омогућава анализу броја рјешења једначине $1-v = e^{-d \cdot v}$. На лијевом графикону се види да за $v \geq 0$ постоји једно



Слика 4.4: Анализа броја рјешења једначине $1 - v = e^{-d \cdot v}$

тривијално рјешење $v = 0$ када је $d \leq 1$, а када је $d > 1$ постоје два рјешења. Десно су рјешења дата као функција од d.

Посматрајући изводе, види се да за $v \leq 0$ извод не може бити већи него у $v = 0$ гдје је баш $f'(0) = d > 0$. Дакле, за $d \leq 0$ постоји јединствено тривијално рјешење $v = 0$, али за $d > 1$ појављују се и друга позитивна рјешења v тако да је $0 < v < 1$.

Технички, показано је да за $d \leq 1$ не постоји највећа компонента. Ако је $d > 1$, то значи да случајно изабран чвор са почетка има d сусједа у просјеку. Сваки његов сусјед има d^2 сусједа итд. Након s корака биће d^s чворова у оквиру удаљености s од почетног чвора. Број $d > 1$ расте експоненцијално и стога ће већина чворова бити повезана унутар највеће компоненте. Њихова учесталост се може наћи као ненула рјешење једначине $1 - v = e^{-dv}$.
Прецизније чињенице се могу постићи ако би се овај проблем сагледао из угла процеса гранања.

У циљу процјене дијаметра графа $\mathcal{G}_{n,p}$ може бити постављен још један хеуристички услов. Процес додавања нових потомака, тј. грана не траје бесконачно, већ само док се не достигне n чворова:

$$d^s = n$$

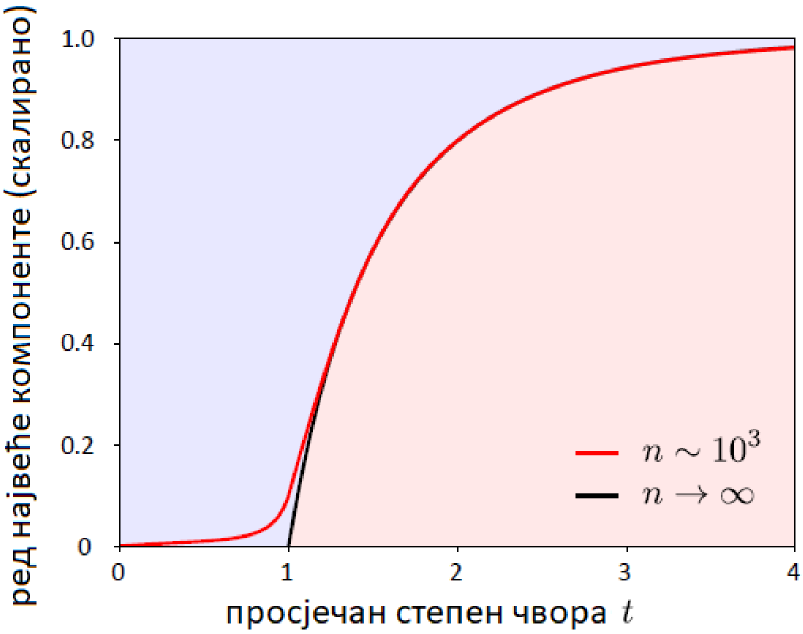
Након логаритмовања овог израза добија се

$$s \cdot ln(d) = ln(n).$$

$$s = \frac{ln(n)}{ln(d)}$$

апроксимира дијаметар Ердеш-Ренијевог графа.

С обзиром на то да је присуство највеће компоненте одређено вриједношћу прага, може се говорити о фазним прелазима. У овом случају се за праг узима установљена вриједност степена чвора, тј. за $d < 1$ највећа компонента не постоји, док за $d > 1$ она постоји.



Слика 4.5: Прелазне фазе у Ердеш-Ренијевом случајном граф

Приликом изучавања настанка највеће компоненте у графу $\mathcal{G}_{n,p}$ очекивани степен чвора је круцијалан. Настанак грана у графу може се повезати са процесом гранања. За дати чвор $\mathcal{V}$, прво се изложе његови потомци, а затим потомци сваког од њих и тако све док постоје чворови у компоненти којој припада чвор $\mathcal{V}$. Грубо речено, ако је до сада изложено $k = o(n)$ чворова, број потомака сваког од чворова је случајна величина са Биномном расподјелом и параметрима $n-k$ и p, те је очекивање $(n-k)p \sim d$. Како Биномна расподјела $\mathcal{B}(n-k,p)$ конвергира Пуасоновој расподјели $\mathcal{P}(d)$ са средњом вриједношћу $d = p(n-1)$ када $n \longrightarrow \infty$ и d је фиксирана константа, процес излагања компоненти се може апроксимирати $Galton-Watson$-овим процесом гранања који, такође, има Пуасонову расподјелу $\mathcal{P}(d)$. Процес гранања почиње једном јединком, односно у овом случају једним чвором, а даље се генеришу случајни чворови у складу са Пуасоновом расподјелом $\mathcal{P}(p)$. Класичан процес гранања са вјероватноћом 1 изумире ако је $d < 1$; ово одговара малим компонентама у графу $\mathcal{G}_{n,p}$. С друге стране, ако је $d > 1$ процес опстаје бесконачно дуго. Вјероватноћа опстанка дата је јединственим позитивним рјешењем $v = v(d) \in (0,1)$ једначине

$$1 - v = e^{-d \cdot v}.$$

Ова вјероватноћа одговара вјероватноћи да одабрани чвор графа $\mathcal{G}_{n,p}$ припада највећој компоненти. На овај начин добија се да је за $d > 1$ највећа компонента у случајном графу $\mathcal{G}_{n,p}$ реда $v \cdot n + o(n)$.

4.5.12 Највећа компонента случајног графа са датом расподјелом степена чвора

Амерички математичари *Molloy* и *Reed* су ријешили питање постојања највеће компоненте у случајном графу који има неуниформну расподјелу степена чвора. Нека је λ_i удио чворова чији је степен i, тада највећа компонента постоји ако и само ако важи

$$\sum_{i=0}^{\infty} i(i-2)\lambda_i > 0.$$

Интуитивно, чворови степена 0 се не појављују, осим ако је ријеч о изолованим чворовима. Ако постоји чвор степена 1, онда то значи смањење ширења грана које су на том чвору јер једна грана улази у чвор, али ни једна не излази или обрнуто. Степен 2 је неутралан јер се тај случај може посматрати као једна грана која улази и излази из чвора. У том случају нема повећања величине компоненте јер нема додатног гранања. Чворови са степеном већим од 2 повећавају границу компоненте за $i - 2$ чвора, јер у чвор улази једна грана, а излази $i - 1$ грана ка новим чворовима, односно излази та једна која је ушла и још $i - 2$ гране. У формули је $i\lambda_i$ пропорционално вјероватноћи достизања степена i, а $i - 2$ представља повећање или смањење величине границе када се достигне степен i.
Претходни услови могу бити примјењени на случајан граф $\mathcal{G}_{n,p}$. Фазни прелаз се дешаваје у случаја када $p = \frac{1}{n}$ и тада сљедећа сума достиже 0

$$\sum_{i=0}^{n} i(i-2)p_i = 0.$$

За $p = \frac{1}{n}$ просјечан степен сваког чвора је 1 и постоји $\frac{n}{2}$ грана. Заправо, расподјела степена чвора је Биномна гдје је вјероватноћа да степен чвора узима вриједност i дата са

$$p_i = \binom{n}{i} p^i (1-p)^{n-i}.$$

Треба показати да за $p_i = \binom{n}{i} p^i (1-p)^{n-i}$ и $p = \frac{1}{n}$ важи да је

$$\lim_{n \longrightarrow \infty} \sum_{i=0}^{n} i(i-2)p_i = 0.$$

$$\lim_{n\longrightarrow\infty}\sum_{i=0}^{n}\binom{n}{i}i(i-2)\left(\frac{1}{n}\right)^{i}\left(1-\frac{1}{n}\right)^{n-i}=$$

$$=\lim_{n\longrightarrow\infty}\sum_{i=0}^{n}i(i-2)\frac{n(n-1)...(n-i+1)}{i!\cdot n^{i}}\left(1-\frac{1}{n}\right)^{n}\left(1-\frac{1}{n}\right)^{-i}=$$

$$=\frac{1}{e}\lim_{n\longrightarrow\infty}\sum_{i=0}^{n}i(i-2)\frac{n(n-1)...(n-i+1)}{i!\cdot n^{i}}\left(\frac{n}{n-1}\right)^{i}\leq$$

$$\leq\sum_{i=0}^{\infty}\frac{i(i-2)}{i!}$$

Треба примјетити да је

$$\sum_{i=0}^{\infty}\frac{i}{i!}=\sum_{i=1}^{\infty}\frac{i}{i!}=\sum_{i=1}^{\infty}\frac{1}{(i-1)!}=\sum_{i=0}^{\infty}\frac{1}{i!}$$

И сљедеће

$$\sum_{i=0}^{\infty}\frac{i^2}{i!}=\sum_{i=1}^{\infty}\frac{i}{(i-1)!}=\sum_{i=0}^{\infty}\frac{i+1}{i!}=\sum_{i=0}^{\infty}\frac{i}{i!}+\sum_{i=0}^{\infty}\frac{1}{i!}=2\sum_{i=0}^{\infty}\frac{1}{i!}.$$

Стога важи

$$\sum_{i=0}^{\infty}\frac{i(i-2)}{i!}=\sum_{i=0}^{\infty}\frac{i^2}{i!}-2\sum_{i=0}^{\infty}\frac{i}{i!}=0.$$

4.6 Процес гранања за графове

Процес гранања се може примјенити приликом изучавања начина на који се формира случајан граф, односно стабло, при чему је као и до сада унапријед одређен број чворова, тј. број јединки у популацији. Почетни чвор који се још назива коријен и чворови који настају у наредним тренуцима као потомци, имају исту расподјелу која одређује број потомака сваког чвора.
Осим што се очигледно примјењује у студијама о популацији, процес гранања се користи и у изучавању повезаних компоненти у случајном графу. Основно питање које се овдје јавља односи се на вјероватноћу да стабло буде коначно, тј. вјероватноћу да након коначног броја генерација процес стане. Ова вјероватноћа се назива вјероватноћа изумирања. Анализа процеса гранања код графова подразумијева одређивање те вјероватноће условљене изумирањем популације, као и очекиване вриједности величине компонената. Важан алат који се овдје користи је генераторна функција.

Нека је Z_j број потомака генерације j и нека је $f_j(x)$ генераторна функција случајне величине Z_j. Онда је $f_1(x) = f(x)$ генераторна функција прве генерације, гдје је $f(x)$ генераторна функција за број потомака коријена стабла. Генераторна функција друге генерације је $f_2(x) = f(f(x))$. Генерално важи да је генераторна функција генерације $j+1$ дата са $f_{j+1}(x) = f_j(f(x))$. Ово се може видјети разматрајући генераторну функцију збира двије једнако расподјељене цјелобројне ненегативне случајне величине X_1 и X_2 која је једнака квадрату генераторне функције

$$f^2(x) = p_0^2 + (p_0p_1 + p_1p_0)x + (p_0p_2 + p_1p_1 + p_2p_0)x^2 + ...$$

Ако збир $X_1 + X_2$ узима вриједност 0, онда обје случајне величине узимају вриједност 0, а ако је вриједност збира 1, онда тачно једна случајна величина узима вриједност 0, а друга 1. Уопштено за независне случајне величине са генераторним функцијама $f(x)$ важи да је генераторна функција суме i таквих случајних величина заправо $f^i(x)$.
Коефицијент уз x_i у $f_j(x)$ означава вјероватноћу да је i-та јединка потомак j-е генерације. Ако постоји i потомака у j-ој генерацији, број потомака генерације $j+1$ је збир i независних случајних величина са генераторном функцијом $f(x)$. Према томе, генераторна функција $j+1$-ве генерације, при чему у j-ој генерацији има i потомака дата је са

$$f_{j+1}(x) = \sum_{i=0}^{\infty} P\{Z_j = i\} f^i(x).$$

Како су $f(x)$, као и итерације $f_2(x), f_3(x), ...$ полиноми по x са ненегативним коефицијентима, онда су те фунцкије монотоно растуће и конвексне на јединичном интервалу. Пошто се вјероватноћа броја потомака сумира на 1, ако је $p_0 < 1$ неки коефицијенти у $f(x)$ су различити од нуле и $f(x)$ је строго растућа.

Нека је $q = 1 - p$ вјероватноћа изумирања у процесу гранања. Ако у првој генерацији постоји i потомака, онда свако подстабло од њих i мора нестати тј. у неком тренутку престати да се грана, што се дешава са вјероватноћом q^i. Према томе, q је једнако суми (по i) производа вјероватноће постојања i потомака и вјероватноће да i-то подстабло изумре:

$$q = \sum_{i=0}^{\infty} p_i q^i.$$

Дакле, q је рјешење једначине

$$x = \sum_{i=0}^{\infty} p_i x^i,$$

гдје је заправо $x = f(x)$, што указује да рјешење задовољава једначину $f(x) = x$ на интервалу [0,1].
$x = 1$ је увијек рјешење јер је

$$f(1) = 1 = \sum_{i=0}^{\infty} p_i.$$

Нека m означава извод функције $f(x)$ у тачки $x = 1$ тј.

$$m = f'(1) = p_1 + 2p_2 + 3p_3 + ...$$

Тада је m очекивани број потомака.
Ако је $m > 1$ природно је очекивати да ће стабло бесконачно расти и да сваки чвор у тренутку j има више од једног потомка, али ово не значи да је вјероватноћа изумирања 0.
Како p_0 означава вјероватноћу да постоји 0 потомака, када је $p_0 > 0$, онда са позитивном вјероватноћом чвор нема потомке и процес гранање се зауставља.
Раније је d означавало степен чвора, а сада овдје m има ту улогу.
Ако је $m < 1$, и m је дефинисано као први извод функције $f(x)$ у тачки $x = 1$ онда је нагиб функције $f(x)$ у тачки $x = 1$ мањи од 1. Како је функција конвексна, то значи да је $f(x) > x$ за $x \in [0, 1)$ па онда једначина $f(x) = x$ нема рјешење на [0,1).
Ако је $m = 1$ и $p_1 < 1$, онда опет из конвексности слиједи да је $f(x) > x$ за $x \in [0, 1)$ и једначина $f(x) = x$ нема рјешење на [0,1).
Ако је $m = 1$ и $p_1 = 1$ онда је $f(x)$ права.
Ако је $m > 1$ онда је нагиб $f(x)$ већи од нагиба x у тачки $x = 1$. Због конвексности, једначина $f(x) = x$ има јединствено рјешење на [0,1). Када је $p_0 = 0$ рјешење је у тачки $x = 0$.
Нека је q најмање, ненегативно рјешење једначине $f(x) = x$. За $m < 1$ и за $m = 1$ и $p_0 < 1$, важи да је $q = 1$, а за $m > 1$ q је строго мање од 1. Управо рјешење q означава вјероватноћу изумирања, а $1-q$ вјероватноћу опстанка (бесмртности), тј. q је вјероватноћа да за неко j нема потомака j-е генерације. То се може видјети из напомене да за $m > 1$ и $0 \leq x < 1$ важи

$$\lim_{j \longrightarrow \infty} f_j(x) = q.$$

Како је раније дефинисано, $f_j(x)$ представља генераторну функцију, тј.

$$f_{j+1}(x) = \sum_{i=0}^{\infty} P\{Z_j = i\} f^i(x).$$

Чињеница да гранична вриједност генераторне функције не зависи од x већ је константа q, говори да је $P\{Z_j = 0\} = q$ и $P\{Z_j = i\} = 0$ за све коначне ненула вриједности i. Преостала вјероватноћа је вјероватноћа неограничене компоненте. Према томе, када је $m > 1$, q је вјероватноћа изумирања, а $1 - q$ вјероватноћа да Z_j расте без границе, тј бесмртно.

Теорема 4.10. *Нека је $f(x)$ генераторна функција броја потомака сваког чвора у стаблу које је генерисано процесом гранања.*
(а) Ако је очекивани број потомка сваког чвора мањи или једнак 1, онда је вјероватноћа изумирања 1, осим у случају да је вјероватноћа тачно једног потомка једнака 1.
(б) Ако је очекивани број потомка сваког чвора већи од 1, онда је вјероватноћа изумирања јединствено рјешење једначине $f(x) = x$ на интервалу [0, 1).

Процес гранања такав да је $m < 1$ или $m = 1$ и $p_1 < 1$ изумире са вјероватноћом 1.
Ако је $m = 1$ и $p_1 = 1$ онда процес гранања садржи бесконачан ланац. Ако је $m > 1$ онда до изумирања долази са неком вјероватноћом мањом од 1, осим у случају $p_0 = 0$ када не долази до изумирања јер чвор увијек има бар једног потомка.

За писање овог поглавља кориштена је литература [2], [3], [4], [5] и [9].

Поглавље 5

Сродне области и примјена

Од кад су представљени, случајни графови се интензивно проучавају. Постали су једна од централних тема савремене математике дијелом због тога што су уско повезани са бројним случајним структурама као што су површи настале на случајан начин, случајне мапе, матрице, системи, а дијелом и због тога што су корисне за моделовање, анализу и рјешавање структурних и алгоритамских проблема којих је све више у математици, информатици, природним и друштвеним наукама, као и у свакодневном животу. Интензивно проучавање случајних структура, посебно феномена фазних прелаза, спојило је више области истраживања, као што је теорија вјероватноћа, дискретна математика, програмирање, статистика.

5.1 Фазни прелази

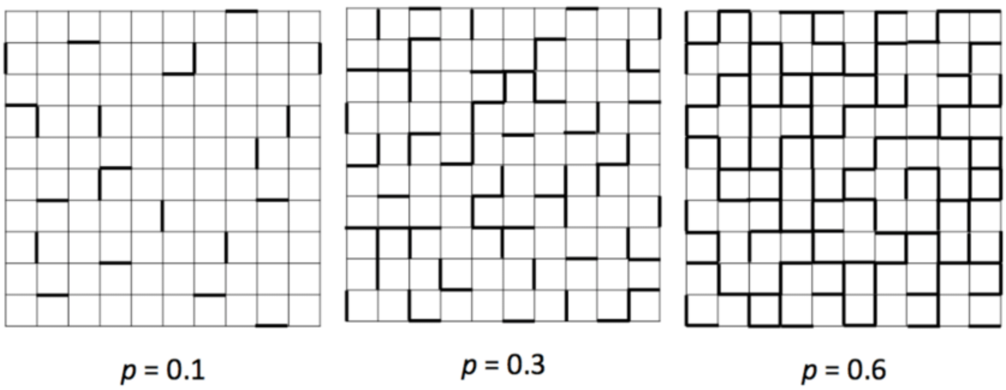
Фазни прелази се баве изненадним промјенама особина огромних структура, изазваних прекорачењем критичног параметра. То се може посматрати у математици и природним наукама из различитих углова. Фазни прелази у моделима случајних графова представљени су у подпоглављу 4.5.4 Прелазне фунцкије за монотона својства.

Примјер фазних прелаза из свакодневног живота је у вези са стањем воде: од леда (чврсто стање), преко воде (течно) до паре (гасовито стање). Овдје постоје двије критичне температуре: тачка смрзавања воде и тачка кључања воде. На ниским температурама у чврстом стању атоми и молекули узајамно јако дјелују са својим сусједима и густо су паковани. На средњим температурама, интеракције су слабије, што се резултује константним промјена на кратким удаљеностима. На високим температурама молекули образују ријетку, али јако динамичну структуру. Интригантна чињеница је да се ове промјене индуковане температуром не дешавају непрекидно, али доводе до два нагла шока тј. **фазна**

прелаза на температурама топљења и кључања. Ова својства се могу изучавати као својства случајних графова са прелазним функцијама.

Претпоставка је да се нека течност налази на врху неког предмета сачињеног од порозног материјала. Проблем који долази из области физике и гласи да ли ће течност наћи свој пут и стићи до дна предмета, може бити математички моделован тзв. границом перколације. Свака два сусједа у квадратној решетки су повезана граном. Грана може бити отворена са вјероватноћом p, што значи да пропушта течност. У супротном је затворена са вјероватноћом $1-p$ независно од осталих. Граница перколације комплетног графа K_n, који садржи n чворова и $\binom{n}{2}$ грана је биномни случајан граф $\mathcal{G}_{n,p}$.

Математичка формулација претходног питања односи се на вјероватноћу егзистенције отворене стазе која повезује врх и дно, за дато p. Испоставља се да за бесконачну решетку постоји критична вриједност p_c, тако да за $p < p_c$ вјероватноћа да таква стаза постоји је 0, а за $p > p_c$ вјероватноћа је 1.



Слика 5.1: Изглед решетке за различите вриједности p

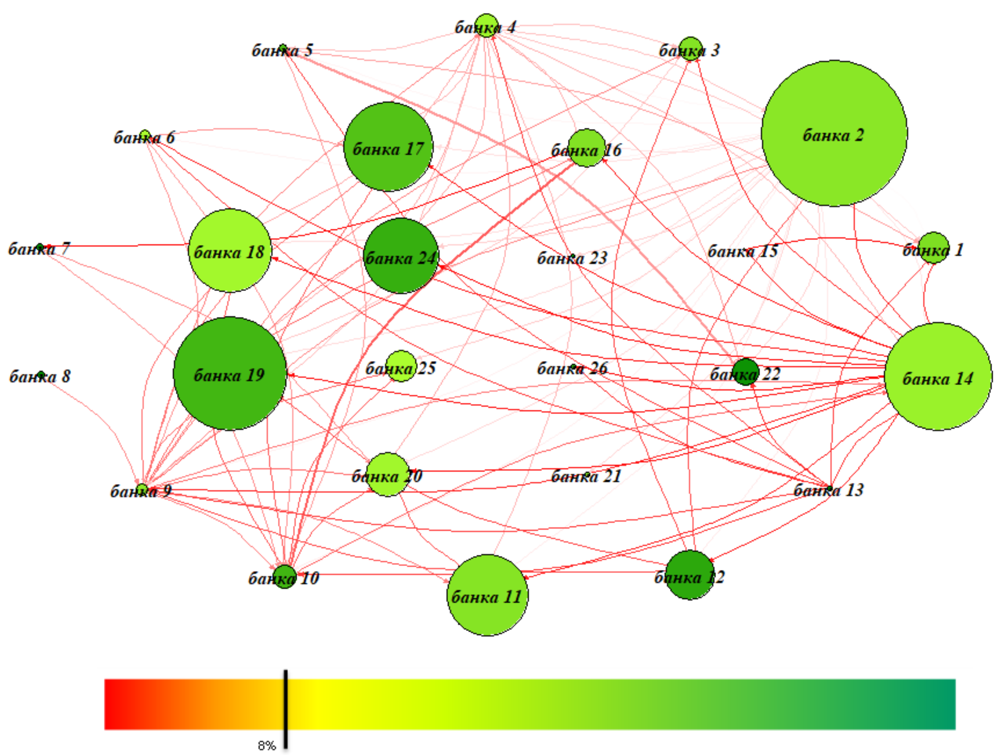
5.2 Друштвене науке

Анализа повезаности и комуникације међу људима из угла теорије детерминистичких и случајних графова је значајна у друштвеним наукама. Један од најстаријих примјера је феномен познат под називом шест степени раздвајања (енг. *six degrees of separation*), на основу кога било која два човјека могу бити повезана ланцем познанства преко 6 особа у просјеку. Оно по чему се препознаје сличност овог модела Ердеш-Ренијевом моделу јесте мала вриједност дијаметра, а разлику представља коефицијент кластеровања који је висок код оваквих модела тзв. мрежа малог свијета. Шанса да два човјека имају заједничког пријатеља је много већа него у случајном графу.

Свакодневно напредовање интернета веома користи друштвеним наукама: људска активност на *web* страницама и комуникација унутар друштвених заједница представљају богат извор података. У позадини ове структуре крију се мреже које могу помоћи у предвиђању будућег развоја и утицаја на друштво далеко изван граница интернета.

5.3 Мрежно моделовање банкарског сектора

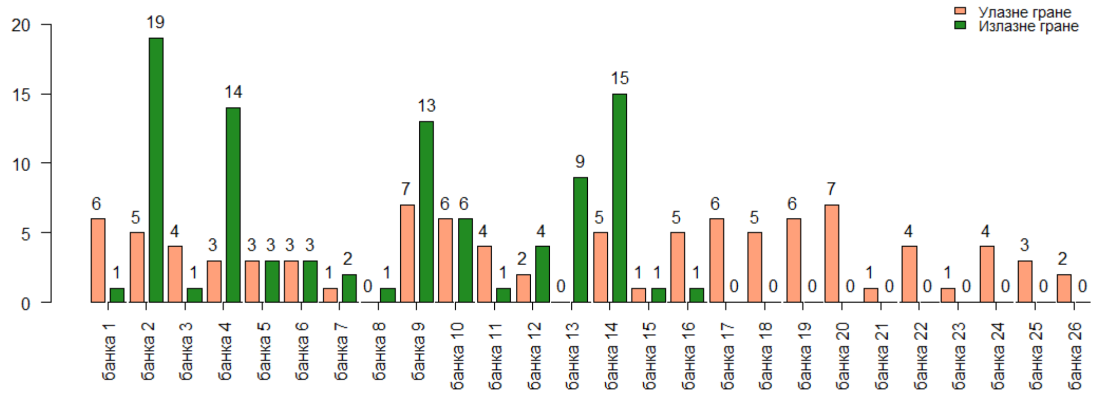
При описивању системског ризика значајно је посматрати динамику међусобних односа финансијских институција, односно бавити се структуром међусобних веза и отпорношћу система као цјелине на потенцијалне финансијске шокове. За моделовање таквог система прикладан је **оријентисан граф** чије гране имају тачно одређен смијер у смислу својеврсног дјеловања једног чвора на други. За вјеродостојно приказивање система, оријентисан граф треба да буде и **тежински** како би се представиле различите особине грана мреже у складу са интензитетом релације.



Слика 5.2: Мрежа банкарског сектора

Мрежу банкарског сектора чине пословне банке представљене чворовима финансијске мреже који су повезани гранама тако да свака грана описује неки однос прве банке према другој банци (на примјер губитак који би имала прва банка у случају пропасти друге). Управо та веза одређује смијер, дебљину и интензитет боје гране. Величина чвора може да представља на примјер износ којим банка располаже, а у складу са одређеним праговима дефинисан је спектар боја за чворове ради лакше анализе што се може видјети на слици 5.2 која приказује једну такву мрежу добијену над реалним подацима помоћу програмског језика R у коме постоје разни пакети за рад са графовима.

Како би се особине мреже одразиле на довољно добар начин, потребно је одредити одговарајуће статистике. Углавном се користе статистике као што су просјечан степен чвора, улазни и излазни степен чвора, степен повезаности, густина мреже, просјечна најкраћа стаза, глобална ефикасност и слично. На слици 5.3 дат је приказ броја улазних и излазних грана за сваки чвор мреже тј. број трансакција које свака банка прими, односно пружи.



Слика 5.3: Број улазних и излазних грана за сваки чвор графа

При посматрању системског ризика потребно је обратити пажњу на чворове кроз које се финансијски шок најлакше преноси. С тим циљем се из првобитне мреже издваја подграф кога чине банке које су посебно значајне за стабилност система као цјелине. Интуитивно, чворови већег степена имају значајнији положај у мрежи, тј. они остварују већи број веза са другим значајним чворовима. На тај начин добија се језгро мреже.

Овај проблем се може гледати и из угла теорије случајних графова, при чему се на основу претходних сценарија мреже одређује вјероватноћа настанка гране између два чвора, а затим се посматрају различити случајеви који би могли настати у будућем периоду.

Оваквом анализом се на основу статистика може закључити да ли је системска компонента ризика у банкарском сектору значајна или не.

За писање овог поглавља корисна је била литература [6].

Поглавље 6

Закључак

Циљ рада био је приказ Ердеш-Ренијевог модела случајних графова и њихових најзначајних особина. Под појомом Ердеш-Ренијев модел може се наћи униформни случајан граф, али је много чешћи биномни случајан граф са којим је рад из угла теорије вјероватноћа ближи и лакши. Показано је да су ова два модела асимптотски еквивалентна.

Својства случајних графова сагледана су из угла прелазних функција. Представљен је појам прелазне функције за монотона својства, као и централни резултат на овом пољу који су дали Болобас и Томасон. Они су тврдили да за свако монотоно својство постоји прелазна функција. За нека конкретна својства у раду показано је постојање прелазне функције и да заиста у односу на праг који она дефинише својство важи, односно не важи за случајан граф.

Теорија случајних графова је јако обимна и у овом раду поменут је само мали дио појмова и примјера. Осим униформног и биномног случајног графа, многи други модели се појављују. Неки од њих су случајан планарни граф, случајан регуларни граф, случајан хиперграф, нехомогени случајан граф и други чија је структура нешто сложенија и тежа за изучавање. Осим тога, постоје бројни показатељи који описују случајне графове, као што је еластичност случајног графа која мјери колико јако случајан граф садржи неко својство и слично.

Доступност података који имају структуру сложених мрежа и могућност њиховог складиштења омогућава детаљан рад и анализу реалних проблема помоћу програмских језика. У раду су приказане и могућности примјене неких претходно разматраних теоријских резултата.

Литература

[1] R. Diestel, Graph Theory, Springer 2000.

[2] B. Bollobás, Random Graphs, Second Edition, University of Memphis and Trinity College, Cambridge

[3] P. Erdös, A. Rényi, On the Evolution of Random Graphs, Publications of the Mathematical Institute of the Hungarian Academy of Sciences

[4] N. Alon, J. H. Spencer, The Probabilistic Method, Third Edition, 2008.

[5] G. Grimmett, Probability on Graphs, University of Cambridge, 2010.

[6] M. Kang, Z. Petráshek, Random Graphs: Theory and Applications from Nature to Society to the Brain, University of Technology, Graz, 2014.

[7] M. Tesliuc, Random Graphs and their Applications, 2017.

[8] J. A. Bondy, U. Murty, Graphs Theory with Applications, Oxford

[9] A. Frieze, M. Karoński, Introduction to Random Graphs

Биографија

Рођена сам 26. октобра 1995. године у Београду.
Гимназију "Филип Вишњић" у Бијељини завршила сам 2014. године. Након тога сам се уписала на Математички факултет у Београду. Дипломирала сам 2018. године на студијском програму Математика, на смијеру Статистика, актуарска и финансијска математика. Након завршених основних академских студија 2018. године образовање сам наставила на мастер академским студијама, такође, на Математичком факултету.
Живим у Београду. Тренутно радим у Народној банци Србије.