به نام خداوند جان و خرد

دانشگاه تهران

دانشکدگان فنی

دانشکده مهندسی برق و کامپیوتر

شبکه‌های اجتماعی

تمرین شماره ۲

نام و نام خانوادگی: علی خرم فر

شماره دانشجویی: ۸۱۰۱۰۲۱۲۹

آبان‌ماه ۱۴۰۲

فهرست مطالب

فهرست اشکال

۱_ پاسخ مسئله شماره ۱

Calculate the expected maximum degree kmax for the undirected networks listed in Table 4.1

در این سوال، باید Kmax برای شبکه‌های موجود در شکل 1 حساب شوند.

NETWORK	N	L	$\langle k \rangle$	$\langle k_{in}^2 \rangle$	$\langle k_{out}^2 \rangle$	$\langle k^2 \rangle$	γ_{in}	γ_{out}	γ
Internet	192,244	609,066	6.34	-	-	240.1	-	-	3.42*
WWW	325,729	1,497,134	4.60	1546.0	482.4	-	2.00	2.31	-
Power Grid	4,941	6,594	2.67	-	-	10.3	-	-	Exp.
Mobile Phone Calls	36,595	91,826	2.51	12.0	11.7	-	4.69*	5.01*	-
Email	57,194	103,731	1.81	94.7	1163.9	-	3.43*	2.03*	-
Science Collaboration	23,133	93,439	8.08	-	-	178.2	-	-	3.35*
Actor Network	702,388	29,397,908	83.71	-	-	47,353.7	-	-	2.12*
Citation Network	449,673	4,689,479	10.43	971.5	198.8	-	3.03**	4.00*	-
E. Coli Metabolism	1,039	5,802	5.58	535.7	396.7	-	2.43*	2.90*	-
Protein Interactions	2,018	2,930	2.90	-	-	32.3	-	-	2.89*

شکل ۱ تغییرات درجه در شبکه‌های واقعی

در شـــبکه‌هایی که از توزیع Pareto پیروی می‌کنند، از رابطه زیر برای محاســـبه بزرگترین درجه استفاده می‌شود:

$$\int_{kmax}^{\infty} p(k)dk = \frac{\gamma-1}{-\gamma+1} kmin^{\gamma-1}[k^{\gamma-1}]_{kmax}^{\infty} = -1\, kmin^{\gamma-1}\,(-kmax)^{\gamma-1} = kmax^{\gamma-1}$$

$$= \frac{1}{N} kmin^{\gamma-1} \rightarrow kmax = kmin\, N^{\frac{1}{\gamma-1}}$$

حال در این سوال Kmax را برای شبکه‌های غیرجهت‌دار در لیست ارائه شده محاسبه می‌کنیم. باتوجه به اعلام kmin برابر ۱، این مورد را لحاظ می‌کنیم.

Kmin = 1

Network	N	γ	Kmax
Internet	192,244	3.42	$\boldsymbol{k_{min} N^{\frac{1}{\gamma-1}} = 192244^{\frac{1}{3.42-1}} = 152.54 \approx 153}$
Science Collaboration	23,133	3.35	$\boldsymbol{k_{min} N^{\frac{1}{\gamma-1}} = 23133^{\frac{1}{3.35-1}} = 71.96 \approx 72}$

Actor Network	702,388	2.12	$\boldsymbol{k_{min}N^{\frac{1}{\gamma-1}} = 702388^{\frac{1}{2.12-1}} = 166019.1 \approx 166019}$
Protein Interactions	2,018	2.89	$\boldsymbol{k_{min}N^{\frac{1}{\gamma-1}} = 2018^{\frac{1}{2.89-1}} = 56.05 \approx 56}$
Power Grid	4,941	Exp.	**20**

باتوجه به توضیحات ارائه‌شده در ذیل جدول، شبکه Power Grid از نوع Scale Free نبوده و توزیع آن از نوع نمایی $e^{-\lambda k}$ است. باتوجه به توضیحات ارائه‌شده در متن کتاب اگر توزیع نمایی بود از فرمول ۴.۱۷ استفاده خواهدشد.(BARABÁSI, 2014)

$$kmax = kmin + \frac{lnN}{\lambda}$$

و در توزیع نمایی مقدار میانگین یا امیدریاضی یک متغیر تصادفی برابر با $\frac{1}{\lambda}$ می‌باشد. پس در این شبکه:

$$< k > = \frac{1}{\lambda} \rightarrow \lambda = \frac{1}{< k >} = \frac{1}{2.67} = 0.374$$

باتوجه به جدول زیر از (*BARABÁSI, 2014*) مقادیر λ و *kmin* برای شبکه *Power Grid* ارائه شده‌است. احتمالا در این جدول تخمین دیگری از λ ارائه‌شده پس برای محاسبات از آن استفاده می‌کنیم.

	λ	k_{min}	P-VALUE	PERCENTAGE
Power Grid	0.517	4	0.91	12%

شکل ۲ پارامترهای شبکه Power Grid

$$kmax = kmin + \frac{lnN}{\lambda} = 4 + \frac{ln4941}{0.517} = 4 + \frac{8.505}{0517} = 20.450 \approx \mathbf{20}$$

۲_پاسخ مسئله شماره ۲

The degree distribution pk expresses the probability that a randomly selected node has k neighbors. However, if we randomly select a link, the probability that a node at one of its ends has degree k is qk = Akpk, where A is a normalization factor.

باتوجه به اینکه پارامتر Pk نشـــان‌دهنده احتمال این اســت که یک گره انتخابی در شــبکه دارای k همسایه باشد، حال در این سوال بیان شده که اگر یک لینک به صورت تصادفی انتخاب شود، احتمال اینکه یکی از گره‌های متصل به آن دارای درجه k باشد برابر است با:

$q_k = Akp_k$

و A فاکتور نرمال‌سازی خواهدبود. حال به بررسی پرسش‌های این مسئله می‌پردازیم.

(a)Find the normalization factor *A*, assuming that the network has a power law degree distribution with 2 < γ < 3, with minimum degree *k*min and maximum degree *k*max.

باتوجه به اینکه در این سـوال مشـخص شـده که توزیع از نوع PowerLaw اسـت می‌توانیم از روابطی که در درس برای این نوع از توزیع خوانده‌ایم اسـتفاده کنیم. در ابتدا باتوجه به اینکه در رابطه این توزیع p_k رابطه مستقیمی با k^{γ} دارد این مورد را در نظر می‌گیریم.

$p_k = Ck^{-\gamma}$

حال این مورد را در فرمول اولیه سوال جایگذاری می‌کنیم:

$q_k = AkCk^{-\gamma} \rightarrow q_k = ACk^{1-\gamma}$

همانند یافتن فاکتور C در اســلایدهای درســی، در اینجا برای یافتن فاکتور نرمال‌ســازی A بر روی رابطه انتگرال گرفته و باتوجه به اینکه جمع همه احتمالات برابر ۱ است انتگرال را برابر ۱ قرار می‌دهیم.

$$\int_{Kmin}^{Kmax} q_k \ dk = 1 \rightarrow A = \frac{1}{\int_{Kmin}^{Kmax} CK^{1-\gamma} \ dk} = \frac{1}{C\int_{Kmin}^{Kmax} K^{1-\gamma} \ dk}$$

محاسبه انتگرال $\int_{Kmin}^{Kmax} K^{1-\gamma} \ dk$:

$$\int_{Kmin}^{Kmax} K^{1-\gamma} \ dk = \left[\frac{K^{2-\gamma}}{2-\gamma}\right]_{Kmin}^{Kmax} = \frac{Kmax^{2-\gamma}}{2-\gamma} - \frac{Kmin^{2-\gamma}}{2-\gamma} = \frac{Kmax^{2-\gamma} - Kmin^{2-\gamma}}{2-\gamma}$$

حال حاصل انتگرال را در رابطه قبلی برای بدست آوردن ضریب نرمال‌سازی A جایگذاری می‌کنیم:

$$A = \frac{1}{C\int_{Kmin}^{Kmax} K^{1-\gamma} \ dk} = \frac{1}{C\,\frac{Kmax^{2-\gamma} - Kmin^{2-\gamma}}{2-\gamma}} = \frac{2-\gamma}{C.(Kmax^{2-\gamma} - Kmin^{2-\gamma})}$$

باتوجه به اسلایدهای درسی و مرجع درس ضریب C برابر است با

$$C = (\gamma - 1)kmin^{-\gamma}$$

که در رابطه بالا جایگذاری می‌شود:

$$A = \frac{2-\gamma}{C.(Kmax^{2-\gamma} - Kmin^{2-\gamma})} = \frac{2-\gamma}{(\gamma-1)kmin^{-\gamma}.(Kmax^{2-\gamma} - Kmin^{2-\gamma})}$$

پس q_k به صورت زیر خواهدبود:

$$q_k = ACk^{1-\gamma} \rightarrow = \frac{2-\gamma}{C.(Kmax^{2-\gamma} - Kmin^{2-\gamma})}Ck^{1-\gamma} = \frac{2-\gamma}{(Kmax^{2-\gamma} - Kmin^{2-\gamma})}k^{1-\gamma}$$

(b) In the configuration model *qk* is also the probability that a randomly chosen node has a neighbor with degree *k*. What is the average degree of the neighbors of a randomly chosen node?

qk که در قسمت قبل بررسی شد، احتمال داشتن k درجه برای یکی از گره‌های متصل به یک یال تصادفی را نشان می‌دهد، بدیهی است که نشان دهنده احتمال وجود همسایه با درجه k برای یک گره انتخابی نیز خواهدبود. باتوجه به اینکه در این سوال متوسط درجه همسایه‌های یک گره تصادفی را مورد نظر دارد باید تمامی همسایه‌های یک گره بررسی شوند. یک نود که به صورت تصادفی انتخاب شده برای محاسبه اینکه همسایه آن به طور متوسط دارای چه درجه‌ای است کافی است برای همه k های ممکن احتمال داشتن همسایه با آن k را حساب کنیم.

حال این مورد را هم برای فضای پیوسته و هم گسسته بررسی می‌کنیم.

۱-۲_ فضای گسسته:

می‌دانیم که متوسط درجه هر شبکه به شرح زیر است:

$$< k > = \frac{\sum_{i=1}^{n} k_i}{n}$$

تعداد کل درجه‌های همسایه‌های گره‌ها در یک شبکه برابر $\sum_{i=1}^{n} k_i^2$ خواهدشد. دلیل این امر این است که هر گره با k_i گره دیگر همسایه است و اگر بخواهیم درجه‌های تولیدشده همسایگی برای این گره را حساب کنیم k_i گره همسایه هرکدام یک همسایه با k_i دارند که منظور همین گره مورد نظر است.

پس متوسط تعداد درجه‌های همسایه‌ها برابر است با :

$$< k >_{\text{neighbors}} = \frac{\sum_{i=1}^{n} k_i^2}{< k >} = \frac{< k^2 >}{< k >}$$

این فرمول برای زمانی است که شبکه کاملا مشخص باشد. حال به بررسی با حالتی که احتمال درجه‌های همسایگان را داریم می‌پردازیم که در این حالت باتوجه به اینکه امیدریاضی را حساب می‌کنیم نیازی به تقسیم بر تعداد نیست و کافی است احتمال هر درجه را در مقدارش ضرب کنیم :

$$< k >_{\text{neighbors}} = \sum_{k=Kmin}^{Kmax} k.q_k$$

۲-۲_ فضای پیوسته:

$$< k >_{\text{neighbors}} = \int_{Kmin}^{Kmax} k.q_k \, dk = \int_{Kmin}^{Kmax} k.\frac{2-\gamma}{(Kmax^{2-\gamma} - Kmin^{2-\gamma})} k^{1-\gamma} \, dk$$

$$= \frac{2-\gamma}{(Kmax^{2-\gamma} - Kmin^{2-\gamma})} \int_{Kmin}^{Kmax} k^{2-\gamma} \, dk = \frac{2-\gamma}{(Kmax^{2-\gamma} - Kmin^{2-\gamma})} \left[\frac{K^{3-\gamma}}{3-\gamma}\right]_{Kmin}^{Kmax}$$

$$\frac{2-\gamma}{(Kmax^{2-\gamma} - Kmin^{2-\gamma})} \left(\frac{Kmax^{3-\gamma} - Kmin^{3-\gamma}}{3-\gamma}\right) = \frac{(2-\gamma)(Kmax^{3-\gamma} - Kmin^{3-\gamma})}{(3-\gamma)(Kmax^{2-\gamma} - Kmin^{2-\gamma})}$$

هرچند که از راه دیگری هم این مسیر طی شد(بدون استفاده از q_k) که در زیر بررسی می‌کنیم.

در قسمت گسسته بررسی شد که

$$< k >_{\text{neighbors}} = \frac{< k^2 >}{< k >}$$

و با توجه به فرمول واریانس داریم:

$$\sigma^2 = < k^2 > - < k >^2$$

با تقسیم آن به $< k >$ خواهیم داشت:

$$\frac{\sigma^2}{< k >} = \frac{< k^2 >}{< k >} - < k >$$

باتوجه به فرمول $< k >_{\text{neighbors}}$ خواهیم داشت:

$$< k >_{\text{neighbors}} = \frac{\sigma^2}{< k >} + < k >$$

ممان اول یا میانگین برابر است با :

$$< k > = \int_{Kmin}^{Kmax} kp(k)dk \ = \mathrm{C}\left(\frac{Kmax^{2-\gamma} - \ Kmin^{2-\gamma}}{2-\gamma}\right)$$

ممان دوم برابر است با:

$$< k^2 > = \int_{Kmin}^{Kmax} k^2 p(k)dk \ = \mathrm{C}\left(\frac{Kmax^{3-\gamma} - \ Kmin^{3-\gamma}}{3-\gamma}\right)$$

$$\sigma^2 = < k^2 > - < k >^2 = \mathrm{C}\left(\frac{Kmax^{3-\gamma} - \ Kmin^{3-\gamma}}{3-\gamma}\right) - \left[\mathrm{C}\left(\frac{Kmax^{2-\gamma} - \ Kmin^{2-\gamma}}{2-\gamma}\right)\right]^2$$

$$< k >_{\text{neighbors}} = \frac{\sigma^2}{< k >} + < k > =$$

$$\frac{\mathrm{C}\left(\frac{Kmax^{3-\gamma} - \ Kmin^{3-\gamma}}{3-\gamma}\right) - \left[\mathrm{C}\left(\frac{Kmax^{2-\gamma} - \ Kmin^{2-\gamma}}{2-\gamma}\right)\right]^2}{\mathrm{C}\left(\frac{Kmax^{2-\gamma} - \ Kmin^{2-\gamma}}{2-\gamma}\right)} + \mathrm{C}\left(\frac{Kmax^{2-\gamma} - \ Kmin^{2-\gamma}}{2-\gamma}\right)$$

باتوجه به پایان ددلاین تمرین امکان ادامه از این راه وجود نداشت.

(c) Calculate the average degree of the neighbors of a randomly chosen node in a network with ***N* = 104**, **γ= 2.3**, ***k*min= 1** and ***k*max= 1, 000**. Compare the result with the average degree of the network, <*k*>.

$$< k >_{\text{neighbors}} = \frac{(2-\gamma)(Kmax^{3-\gamma} - \ Kmin^{3-\gamma})}{(3-\gamma)(Kmax^{2-\gamma} - \ Kmin^{2-\gamma}\ )}$$

$$= \frac{(2-2.3)(1000^{3-2.3} - \ 1^{3-2.3})}{(3-2.3)(1000^{2-2.3} - \ 1^{2-2.3}\ )} = \frac{(-0.3)(1000^{0.7} - \ 1^{0.7})}{(0.7)(1000^{-0.3} - \ 1^{-0.3}\ )} = \frac{(-0.3)124.89}{(0.7)(0.125-1\ )} =$$

$$\frac{-37.467}{-0.6125} = \ 61.1706 \ \approx 61$$

سعی کنیم این مقدار متوسط را با روش دومی که که بررسی کردیم هم محاسبه کنیم:

$$< k >_{\text{neighbors}} = \frac{\sigma^2}{< k >} + < k > =$$

$$\frac{\mathrm{C}\left(\frac{Kmax^{3-\gamma}-\ Kmin^{3-\gamma}}{3-\gamma}\right)-\left[\mathrm{C}\left(\frac{Kmax^{2-\gamma}-\ Kmin^{2-\gamma}}{2-\gamma}\right)\right]^2}{\mathrm{C}\left(\frac{Kmax^{2-\gamma}-\ Kmin^{2-\gamma}}{2-\gamma}\right)}+\mathrm{C}\left(\frac{Kmax^{2-\gamma}-\ Kmin^{2-\gamma}}{2-\gamma}\right)$$

$$\mathrm{C}=(\gamma-1)kmin^{-\gamma}=1.3*1=1.3$$

$$<k>_{\text{neighbors}}=\frac{1.3\left(\frac{124.89}{0.7}\right)-\left[1.3\left(\frac{-0.875}{-0.3}\right)\right]^2}{1.3\left(\frac{-0.875}{-0.3}\right)}+1.3\left(\frac{-0.875}{-0.3}\right)=491$$

حال برای محاسبه متوسط درجه باتوجه به فرمول متوسط درجه یا ممان اول k داریم:

$$p_k=Ck^{-\gamma}$$

$$<k>=\int_{Kmin}^{Kmax}kp(k)dk=\int_{Kmin}^{Kmax}Ck^{1-\gamma}dk=\ C\int_{Kmin}^{Kmax}k^{1-\gamma}dk=C\left[\frac{K^{2-\gamma}}{2-\gamma}\right]_{Kmin}^{Kmax}=$$

$$\mathrm{C}\left(\frac{Kmax^{2-\gamma}-\ Kmin^{2-\gamma}}{2-\gamma}\right)=(\gamma-1)kmin^{-\gamma}\left(\frac{Kmax^{2-\gamma}-\ Kmin^{2-\gamma}}{2-\gamma}\right)=$$

$$(2.3-1)1^{-2.3}\left(\frac{1000^{-0.3}-\ 1^{-0.3}}{-0.3}\right)=1.3\left(\frac{1000^{-0.3}-\ 1^{-0.3}}{-0.3}\right)=1.3\left(\frac{0.125-1}{-0.3}\right)=$$

$$1.3\left(\frac{-0.875}{-0.3}\right)=1.3\times2.91=3.783\ \approx4$$

باتوجه به نتایج به‌دست آمده، متوسط درجه در کل شبکه بسیار کمتر از متوسط درجه همسایه‌های هر گره است. یعنی هر گره انتخابی به صورت تصادفی تعداد درجه کمتری نسبت به همسایه‌های خود دارد.

(d) How can you explain the "paradox" of (c), that is a node's friends have more friends than the node itself?

دلایل مختلفی برای این موضوع وجود دارد. در شبکه‌هایی که از توزیع Power Law پیروی می‌کند، احتمال اینکه همسایه‌های گره تصادفی که انتخاب کرده‌ایم، درجه بالایی داشته باشند وجود دارد. زیرا گره‌های با درجه بالا هستند که ارتباطات بیشتری دارند. پس احتمال اینکه همسایه گره انتخابی باشند بالاست. همین دلیل منجر به بالارفتن متوسط درجه خواهدشد.

یک دلیل دیگر هم میتواند این باشد که در توزیع PowerLaw ۸۰ درصد لینک‌ها به ۲۰ درصد گره‌ها تعلق دارند. پس احتمال اینکه گره تصادفی که انتخاب کرده‌ایم جز آن ۸۰ درصدی باشد که درجه

کمتری نسبت به ۲۰ درصد بقیه گره‌ها داشته باشند بسیار بیشتر است. پس این گره‌ها، همسایه کمتری نسبت به آن ۲۰ درصد شبکه که همان هاب‌های شبکه هم محسوب می‌شوند داشته و علاوه بر این احتمالا چندین گره از آن ۲۰ درصد شبکه همسایه گره انتخابی ما هستند زیرا که هاب‌های شبکه با تعداد بسیار زیادی از گره‌ها ارتباط دارند، که همین مورد مزید بر علت خواهدشد.

۳_ مراجع

BARABÁSI, A.-L. (2014). *NETWORK SCIENCEACKNOWLEDGEMENTS PERSONAL .INTRODUCTION*