به نام خداوند جان و خرد

دانشگاه تهران

دانشکدگان فنی

دانشکده مهندسی برق و کامپیوتر

شبکه‌های اجتماعی

تمرین شماره ۱

نام و نام خانوادگی: علی خرم فر

شماره دانشجویی: ۸۱۰۱۰۲۱۲۹

آبان‌ماه ۱۴۰۲

فهرست مطالب

فهرست اشکال

۱_ جواب سوال ۱)

فرضیات این سوال به این صورت است که مدل Erdős-Rényi را با تعداد گره N=3000 که هر نود با احتمال $P=10^{-3}$ تشکیل لینک داده است بررسی می‌کنیم.

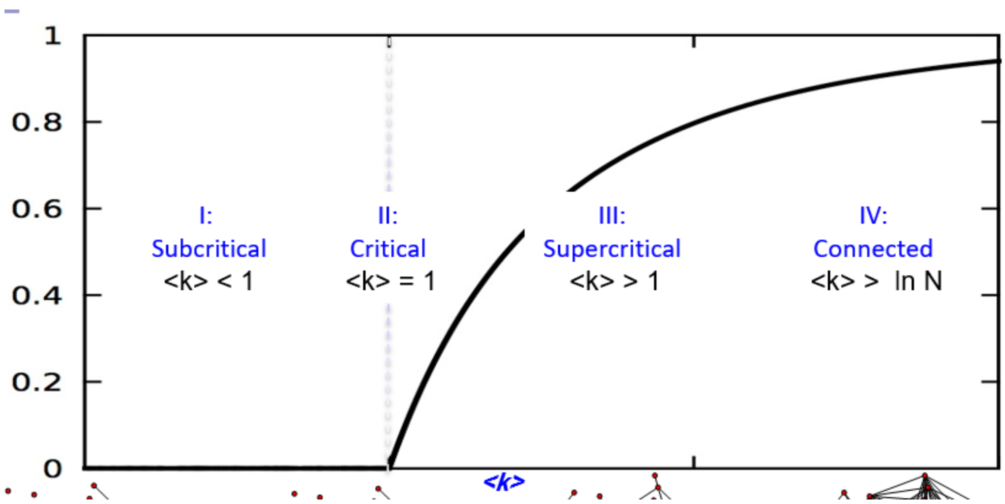
(a) What is the expected number of links, ⟨L⟩?

در حالتی که برای تشکیل یال‌ها یک احتمال به ما داده می‌شود از نوع G(N,P) خواهد بود لذا برای محاسبه ⟨L⟩ یک فرمول داریم متوسط کل یال‌ها را به ما می‌دهد:

$$\langle L \rangle = \frac{N(N-1)}{2} P = \frac{3000(3000-1)}{2} 10^{-3} = 4498500 * 0.001 = 4498.5$$

(b) In which regime is the network?

تصویر رژیم‌های مختلف را در زیر مشاهده می‌کنیم:



شکل ۱نمودار رژیم‌های مختلف گراف رندم

پس ابتدا درجه متوسط را حساب می‌کنیم:

$\langle K \rangle = (N-1)P = 2999 * 0.001 = 2.999$

از طرفی مقدار LN 3000 برابر حدودا ۸ است پس باتوجه به این موارد این گراف در رژیم Supercritical قرار دارد.

(c) Calculate the probability Pc so that the network is at the critical point.

باتوجه به فرمول‌های مذکور برای محاسبه احتمالی که در آن گراف در رژیم critical قرار گیرد باید متوسط درجه برابر ۱ باشد در نتیجه طبق فرمول احتمال برابر خواهد بود با:

$\langle K \rangle = (N-1)P = 1 \Rightarrow (2999)P = 1 \Rightarrow P = 1/2999 \Rightarrow P = 0.00033344448$

(d) Given the linking probability p = 10–3, calculate the number of nodes Ncr so that the network has only one component.

در رژیم Connected، داریم $\langle K \rangle > \ln N$ و گراف همبند خواهدبود و ما نود ایزوله نخواهیم داشت پس نتورک فقط ۱ Component خواهدداشت. این حالت زمانی اتفاق می‌افتد که علاوه بر شرط بالا و همینطور که در کلاس گفته شد N ما به سمت اعداد بزرگ میل کند، شرایط زیر برقرار باشد:

$\langle K \rangle > \ln N \rightarrow p > \ln N/N \rightarrow N * 10^{-3} > \ln N \rightarrow N > 10^{3} * \mathrm{Ln}N$

حال برای حل این نامساوی مقادیر مختلف را جایگذاری می‌کنیم تا نامعادله برقرار باشد و N مورد نظر را برمیگردانیم که با حل این سوال مشخص می‌شود. N های مختلفی با کمک پایتون تقریب زده شد و در نهایت به عدد ۹۱۱۹ رسیدیم.

(e) For the network in (d), calculate the average degree $\langle k_{cr} \rangle$ and the average distance between two randomly chosen nodes $\langle d \rangle$.

برای حل متوسط درجه داریم:

$\langle K \rangle = (N-1)P = 9118 * 10^{-3} = 9.118$

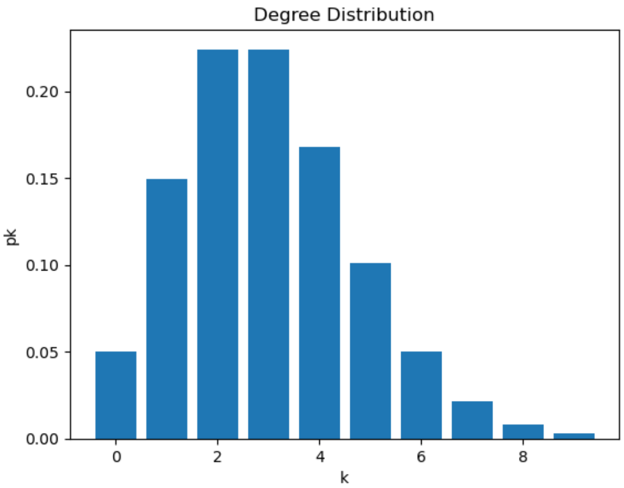
برای حل متوسط فاصله داریم:

$\langle d \rangle = \log N / \log\langle K \rangle = 3.95994721571 / 0.95989958786 = 4.12537651416$

(f) Calculate the degree distribution p_k of this network (approximate with a Poisson degree distribution).

$\langle K \rangle = 2.999$

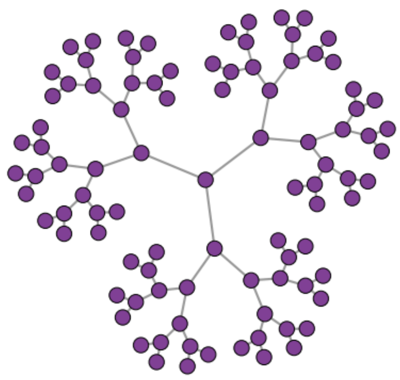
$$P(K) = \frac{e^{-<K>} <K>^{K}}{K!} = \frac{e^{-3}\, 3^{K}}{K!}$$



شکل ۲ توزیع پواسون برای درجه

سوال اضافه) احتمال اینکه یک نود دارای درجه ای بزرگتر یا مساوی ۱۰۰باشد چقدر است؟

$$P(k >= 100) = 1 - P(k < 100) = 1 - \frac{e^{-<2.999>} <2.999>^{100}}{100!}$$



شکل۳ نمونه درخت کیلی

۲_جواب سوال ۲)

(a) Calculate the number of nodes reachable in t steps from the central node.

همانطور که از ساختار درخت مشاهده می‌شود، هنگام حضور در ریشه به اندازه k گره قابل دسترس است. هرگاه به سمت برگ‌ها میل شود از آنجا که یکی از درجه رئوس مشترک خواهدبود، تعداد گره جدید قابل دسترس در این مراحل K-1 خواهدبود.

در این سوال t قدم طی می‌شود. اگر ریشه t=0 باشد حالتی است که با ۰ قدم قابل دسترس است. و همانطور که بررسی شد به ازای فرزندان ریشه، به اندازه قدم‌ها ضرب در k-1 می‌شود که تعداد گره‌هایی است که در هر سطح اضافه می‌شود. توجه شود که اولین قدم به دلیل حضور در ریشه برابر k خواهدبود. پس داریم:

$$f(x) = \begin{cases} 1, & root \\ k(k-1)^{t-1}, & not\ root \end{cases}$$

دلیل اینکه t-1 قرار گرفته‌شده، این است که ۱ واحد آن مرحله فرزندان ریشه است که k خواهدبود و این مقدار در $k(k-1)^{t-1}$ ضرب می‌شود.

(b) Calculate the degree distribution of the network.

ما به دنبال فرموله‌کردن حالتی هستیم که یک گره انتخابی احتمالا دارای درجه x باشد چقدر است. در درخت کیلی یا درجه یک گره ۱ است یا k. پس احتمالات فقط برای این دو مقدار دارد و برای بقیه موارد صفر خواهدبود. اگر گره برگ باشد، درجه آن ۱ است. اگر گره غیربرگ باشد k است. حال کافی است احتمال برگ بودن یک گره را حساب کنیم.

احتمال اینکه درجه یک گره انتخابی در درخت کیلی یک باشد برابر است با نسبت تعداد برگ‌ها به تعداد کل گره‌ها.

Leaves = $k\ (k-1)^{p-1}$

نحوه محاسبه تعداد برگ‌ها در قسمت قبل انجام شد.

باتوجه به بررسی‌های انجام شده در (Ostilli, 2012) تعداد کل گره‌های درخت کیلی برابر است با:

N = $\frac{k((k-1)^{p}-1)}{(k-2)}$

پس برای محاسبه احتمال برگ بودن یک گره یا همان ۱ بودن درجه آن leaves بر N تقسیم می‌شود:

$$P(X=1) = \frac{Leaves}{N} = \frac{k\,(k-1)^{p-1}}{\frac{k((k-1)^{p}-1)}{(k-2)}} = \frac{(k-2)k(k-1)^{p-1}}{(k-1)^{p}-1}$$

احتمال گره غیر برگ برابر است با تفاضل P(X=1) از ۱ پس احتمال اینکه نود درجه K داشته باشد برابر است با:

$$P(X=k) = 1-P(X=0) = 1-\frac{(k-2)k(k-1)^{p-1}}{(k-1)^{p}-1}$$

پس به طور کلی:

$$P(X) = \begin{cases} \dfrac{k((k-1)^{p}-1)}{(k-2)}, & X=1 \\ 1-\dfrac{(k-2)k(k-1)^{p-1}}{(k-1)^{p}-1}, & X=K \\ 0 & O.W \end{cases}$$

(c) Calculate the diameter d_{max}.

باتوجه به اینکه بیشترین فاصله بین دو گره را میتوان حالتی فرض کرد که یک گره تا ریشه و از ریشه تا گره زیردرخت دیگر فرض کنیم از مرتبه p و در مورد فرض این سوال برابر 2p خواهدبود. اگر درجه ۱ باشد برابر p خواهدبود.

$D_{max} = 2P.$

(d) Find an expression for the diameter d_{max} in terms of the total number of nodes N.

باتوجه به اینکه تعداد کل گره‌ها برابر است با:

$$N = \frac{k((k-1)^{p}-1)}{(k-2)}$$

ابتدا مقدار N را بر حسب P بدست می‌آوریم.

$$(k-2)N = K(K-1)^P - K \rightarrow (K-1)^P = \frac{(K-1)N+K}{K}$$

از طرفین لگاریتم گرفته می‌شود:

$$P = Log_{(k-1)}^{\frac{(K-1)N+K}{K}}$$

از طرفی در قسمت قبل محاسبه کردیم که Dmax = 2P.

$$D_{max} = 2Log_{(k-1)}^{\frac{(K-1)N+K}{K}}$$

(e) Does the network display the small-world property?

باتوجه به بررســـی‌های مختلفی که بر روی مدل‌های مختلف مثل گراف Random و یا -Watts Strogatz انجام دادیم شـــرایط small-world علاوه بر clustering coefficients بالا وجود فاصلـــه کم بین نودهای داخل شبکه است. اگر با شرایط این درخت مقایسه کنیم متوجه میشویم که برای مثال در شکل بالا باتوجه به اینکه مقدار درجه اختلاف زیادی با p ندارد و ما در خود هر ســـطح نیز تعداد زیادی گره داریم. باتوجه به وجود یک گره مرکزی که با همه گره ها دیگر با حداکثر مرتبه p فاصلـــه دارد و منجر به کاهش فاصله متوسط می‌شود ، که در هر سطح نیز این مورد تکرار میشود. باتوجه به محاسبات قسمت قبل حداکثر فاصله از مرتبه لگاریتمی است. پس یکی از شروط برقرار است.

و از طرفی یک ســـری Cluster های تشـــکیل می‌دهند زیرا که هر گره به اندازه درجه درخت گره همســـایه دارد و به هین ترتیب برای همســـایه‌های گره، میتوان حالت small world برای این درخت درنظر گرفت.

۳_ جواب سوال ۳)

(a) Calculate the average degree of the "blue" subnetwork made of only blue nodes, and the average degree in the full network.

برای محاسبه متوسط درجه زیرشبکه آبی را بدست آوریم حالتی است که همه N درجه آبی هرکدام با احتمال P با یکی دیگر از اعضـــای زیرشـــبکه آبی لینک دارند. پس هر نود با احتمال P با N-1 گره دیگر میتواند لینک داشته باشد. پس متوسط درجه برابر است با:

$\langle d_{blue} \rangle = (N-1)P$

برای محاسبه متوسط درجه در کل شبکه حالتی است که هر گره از آبی یا قرمز می‌توانند با N-1 گره از شبکه داخلی خود با احتمال P لینک داشته‌باشند. علاوه بر این هرگره می‌تواند با احتمال q با N گره از شبکه خارجی لینک داشته‌باشد.

$<d> = (N-1)P + Nq$

(b) Determine the minimal p and q required to have, with high probability, just one component.

زمانی این شبکه فقط یک component خواهد داشت که اولا احتمال تشکیل لینک‌های داخلی به حالتی برسد که مطمئن شویم هم زیرشبکه قرمز و هم زیرشبکه آبی فقط ۱ component دارند. مطابق مطالب درسی در رابطه با گراف‌های Random این حالت زمانی اتفاق می‌افتد که گراف در حالت connected باشد. شرط برقراری این حالت برای N های بزرگ برابر است با : P > LnN/N

پس باید این شرط برای زیرشبکه‌های داخلی برقرار باشد.

برای شرط q کافی است فقط یک گره از زیرشبکه آبی با یک گره از زیرشبکه قرمز لینک داشته‌باشد تا شرط دوم برای این مسئله برقرار شود.(**قدرت لینک‌های ضعیف**)

هرکدام از گره‌های یک زیرشبکه می‌توانند با N گره از زیرشبکه مقابل لینک داشته باشند، پس به طور کلی تمام لینک‌های ممکن برای برقراری این شرط N^2 خواهدبود. اگر یکی از این گره‌ها برقرار شود تمام شبکه ۱ component خواهدداشت. پس مقدار $q > \frac{1}{N^2}$

اگر بخواهیم حداقل ممکن برای اینکه با احتمال زیاد این شروط برقرار باشند را در نظر بگیریم:

$$p = \frac{(1+\epsilon)(lnN)}{N}$$

$$q = \frac{(1+\epsilon)}{N^2}$$

(c) Show that for large N even very snobbish networks (p≫q) display the small-world property.

ابتدا دوباره به بررسی شروط برقراری small-world می‌پردازیم. هم اینکه شبکه clustering coefficients بالا داشته باشد و علاوه بر این مقدار فاصله متوسط بین دو گره انتخابی کم باشد.

حالتی که این شبکه p های بسیار بزرگتر از q داشته باشد شرایطی است که زیرشبکه‌های قرمز و آبی به احتمال زیاد لینک‌های بسیار بیشتری نسبت به لینک‌های خارجی نسبت به زیرشبکه متفاوت از خود دارند. پس ویژگی small world همانند گراف random در زیرشبکه‌ها برقرار است. تنها موردی که نیاز است یک longrange link است که از طرف زیرشبکه قرمز به سمت آبی برقرار باشد. باتوجه برقراری ویژگی coefficients بالا و فاصله کم در هر یک از زیرشبکه‌ها، با برقراری این لینک ضعیف ویژگی small-world برای کل شبکه برقرار خواهد شد که احتمال بسیار کمی از $\frac{(1+\epsilon)}{N^2}$ می‌تواند این شرط را برقرار کند.

۴_مراجع

Ostilli, M. (2012). Cayley Trees and Bethe Lattices: A concise analysis for mathematicians and physicists. *Physica A: Statistical Mechanics and its Applications*, *391*(12), 3417-3423