به نام خدا

تمرین دوم شبکه‌های اجتماعی

محمد ناصری

۸۱۰۱۰۰۴۸۶

آذر ۱۴۰۰

سوال: فرض کنید در گراف تصادفی مدل Watts-Strogatz تعداد N نود داشته باشیم و <k>=2c باشد

الف: اثبات کنید برای حالت β=0 ضریب خوشه بندی عبارت خواهد بود از:

$$\mathrm{C}([\beta=0]) \;=\frac{3(c-1)}{2(2c-1)}$$

پاسخ:

برای بدست آوردن ضریب خوشه بندی نیازمند ۲ مقدار مشخص هستیم:

1. تعداد لینک‌های واقعی بین همسایگان گره
2. تعداد حداکثر لینک ممکن بین همسایگان گره

میدانیم زمانیکه از $\beta=0$ گراف حاصل یک Lattice منظم خواهد بود فلذا همسایگان یک گره به صورت مساوی بین همسایگان چپ و راست هستند. یعنی تعداد c عدد گره همسایه چپ و c عدد گره همسایه راست خواهیم داشت (<k> = 2c)

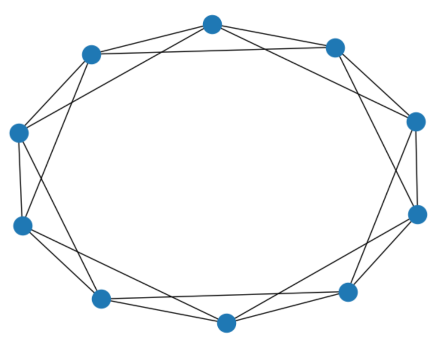
از طرفی میدانیم تعداد حداکثر لینک ممکن بین <k> = 2c گره، عبارت خواهد بود از:

$$\#E=\frac{k*(k-1)}{2}=\frac{2c(2c-1)}{2}=2c^2-c=c(2c-1)$$

برای بدست آوردن تعداد واقعی لینک ها از آزمون کردن چند مورد و اثبات استقرایی برای بدست آوردن رابطه استفاده میکنیم.

تست‌های زیر با در نظر گرفتن N=10 و $\beta=0$ انجام و آورده شده است:

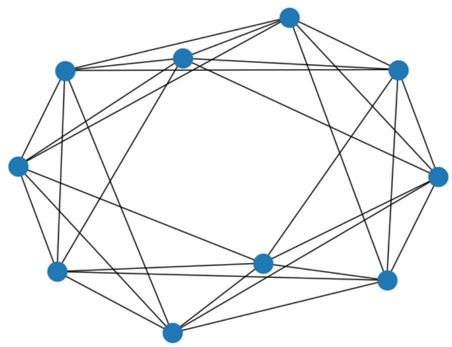
- با در نظر گرفتن 4 = <k> (c=2):



بطور واضح مشخص است که تعداد لینک میان همسایگان یک راس در این گراف برابر ۳ عدد است. همچنین ضریب خوشه بندی در این گراف به شرح زیر است:

$$C=\frac{3}{\frac{4(4-1)}{2}}=0.5$$

- با در نظر گرفتن 6 = <k> (c=3):



بطور واضح مشخص است که تعداد لینک میان همسایگان یک راس در این گراف برابر ۹ عدد است. همچنین ضریب خوشه بندی در این گراف به شرح زیر است:

$$C = \frac{9}{\frac{6(6-1)}{2}} = 15$$

با بررسی نمونه‌هایی از این قبیل به یک فرمول حدسی برای تعداد لینک‌های ممکن میرسیم که عبارت است از:

$$L_i = \frac{3c^2 - 3c}{2} = \frac{3}{2}\, c(c-1)$$

حال سعی میکنیم با استفاده از استقرای ریاضی فرمول بدست آمده را اثبات کنیم:

پایه استقرا: c = 2

$$(c=2) \rightarrow L_2 = \frac{3}{2}\, 2(2-1) = 3$$

فرض استقرا: c = k

$$(c=k) \rightarrow\ L_k = \frac{3}{2}\, k(k-1)$$

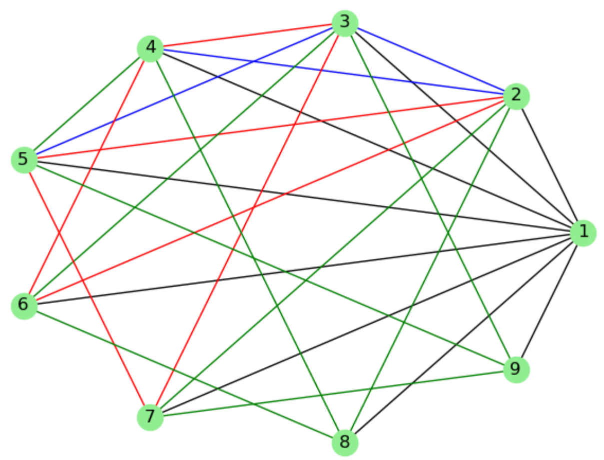
حکم استقرا: c = k+1

$$(c=k+1) \rightarrow\ L_{k+1} = \frac{3}{2}\, (k+1)k$$

با توجه به موارد بالا داریم:

$$L_{k+1} -\ L_k = \frac{3}{2} k\big((k+1)-(k-1)\big) = 3k$$

با بررسی Latticeهای ساخته شده توسط گراف Watts-Strogatz متوجه میشویم با اضافه شدن ۱ واحد به مقدار c دو گره به چپ و راست بعنوان همسایه ا ضافه میشود که این دو گره هرکدام به (c - 1) گره قبلی متصل میشوند. همچنین به علت بالا رفتن درجه گره‌های گراف، یک یال به یالهای (c - 1) گره قبلی اضافه میشود. بصورت کلی تعداد یالهای اضافه شده به گراف بصورت زیر است:



شکل 1 یالهای اضافه شده به ازای اضافه شدن یک واحد c در بازه c = [2,4]
(تنها یالهای اضافه شده به همسایگان یک عدد گره و تغییرات آن کشیده شده است و گراف بصورت کامل ترسیم نشده)

$$\#e = 2(c-1) + \frac{2(c-1)}{2} = 3(c-1)$$

این مقدار برای c = k+1 برابر 3k میشود. از این نتیجه‌گیری میتوان به صحت مقدار بدست آمده در اثبات استقرایی اطمینان پیدا کرده و درستی فرمول بدست آمده اثبات میشود.

حال با توجه به فرمول‌های بدست آمده داریم:

$$C_i = \frac{L_i}{\#E_i} = \frac{\frac{3}{2}\, c(c-1)}{c(2c-1)} = \frac{3(c-1)}{2(2c-1)}$$

که همان فرمول سوال است. بدین روش فرمول مربوطه اثبات میشود.

ب: برای حالت β=0 متوسط فاصله نودها عبارت خواهد بود از:

$$< \mathrm{d}[\beta = 0] > = \frac{N}{4c}$$

پاسخ:

برای بدست آوردن میانگین فاصله گره‌ها چند نکته را در نر میگیریم:

1 . گراف حاصل منتظم است. یعنی با محاسبه فاصله یک گره از سایر گره‌ها میتوان همه را برابر با آن مقدار در نظر گرفت.

2 . گراف را به ازای یک راس به ۲ نیم تقسیم میکنیم و اگر فاصله گره از نیمی را بدست بیاوریم با x2 کردن آن میتوان کل را حساب کرد

با احتساب فرض‌های فوق داریم که به ازای یک راس و گراف ۲نیم شده، فاصله گره از سایر گره‌ها بدن صورت است که از c گره اول فاصله ۱ دارد، از c گره دوم فاصله ۲ دارد و.... از c گره $\frac{\frac{n}{2}}{c}$ فاصله $\frac{\frac{n}{2}}{c}$ دارد. پس برای جمع نیمی از فواصل داریم:

$$\frac{d_i}{2} = (1 * c) + (2 * c) + \cdots + \left(\frac{n}{2c} * c\right) = c\left(1 + 2 + \cdots + \frac{n}{2c}\right) = \frac{c\left(\left(\frac{n}{2c} + 1\right) * \frac{n}{2c}\right)}{2}$$
$$= \frac{c\left(\frac{n^2}{4c^2} + \frac{n}{2c}\right)}{2} = \frac{(\frac{n^2}{4c} + \frac{n}{2})}{2}$$

در نتیجه داریم:

$$d_i = (\frac{n^2}{4c} + \frac{n}{2})$$

پس میانگین فاصله به عبارت زیر است:

$$< d > = \frac{n * (\frac{n^2}{4c} + \frac{n}{2})}{n(n-1)} = \frac{(\frac{n^2}{4c} + \frac{n}{2})}{n-1}$$

و با حد گرفتن از عبارت بالا داریم:

$$< d > = \frac{(\frac{n^2}{4c} + \frac{n}{2})}{n-1} \sim \frac{n}{4c} + \frac{1}{2} \sim \frac{n}{4c}$$

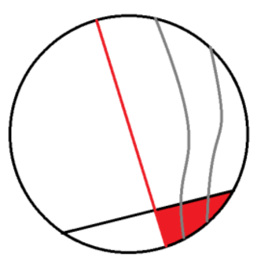
بدین ترتیب فرمول خواسته شده اثبات میشود.

ج: فرض کنید مرحله اول پروسه ایجاد گراف انجام گرفته یعنی گراف منظمی ایجاد شده که هر نود به c همسایه چپ و c همسایه راست لینک دارد. حال برای اولین بار یکی از لینکها rewire شده و انتهای آن یال به یک نود که بصورت تصادفی انتخاب شده وصل میشود. محاسبه کنید متوسط فاصله نودها بعد از rewire شدن چقدر خواهد شد؟

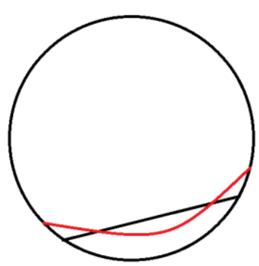
پاسخ:

برای حل این مساله حالت‌های مختلفی میتوان در نظر گرفت برای مثال بهترین حالت Rewire شدن حالتی است که لینک اضافه شده، قطر گراف و دورترین مسیر ممکن بین رئوس باشد. بدترین حالت، حالتی است که لینک جدید بین یک راس و c+1 امین راس بعد از این راس ایجاد شود.

متاسفانه برای بدست آوردن مقدار دقیق تفاوت حاصل شده در میانگین فاصله راه حل جامع و کاملی که همه احتمالات را شامل شود،به ذهن بنده نرسید و تمامی راه‌های بررسی شده از قبیل تقسیم کردن فضای گراف به چندین قسمت به صورت تقارنی یا از روی یال جدید ایجاد شده از نظر بنده احتمال‌هایی را شامل نمیشوند. موردی که من در بررسی‌ها به آن رسیدم این بود که با اضافه شدن لینک جدید و تقسیم گراف به دو بخش، اگر m گره که لینک جدید ایجاد کرده را بررسی کنیم به نیمی از آنها از سمت نیمه کلی گراف هر کدام یک واحد به مسیر اضافه میشود(بدلیل حذف یک لینک)



برای مثال پاسخ بررسی شده در گروه کلاس که گراف را توسط لینک جدید ۲ بخش در نظر میگیرد میتواند جواب دقیقی نداشته باشد چرا که مسیر برای مثال شکل زیر در نظر گرفته نشده است.



این مسیر بین دو گره در قسمت N-m هست ولی از فرمول کلی پیروی نمیکند و کوتاه تر شده است.