



درس شبکه‌های پیچیده

تمرین سوم

شبکه‌های جهان کوچک و ابزار SNAP

مهلت ارسال: ۱۴۰۳/۰۲/۰۵

مدرس: دکتر مصطفی صالحی

دانشکده‌گان علوم و فناوری‌های بین رشته‌ای

دانشکده سامانه‌های هوشمند

نیم‌سال دوم سالتحصیلی ۱۴۰۳-۱۴۰۲

۱. ضریب خوشه‌بندی و میانگین طول مسیر در مدل جهان کوچک نیومن-واتس را محاسبه کنید.
۲. فرض کنید در گراف تصادفی مدل Watts-Strogatz تعداد N نود داشته باشیم و $\langle k \rangle = 2c$
 - a. اثبات کنید برای حالت $\beta=0$ ضریب خوشه بندی عبارت خواهد بود از:

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

- b. اثبات کنید برای حالت $\beta=0$ متوسط فاصله نودها عبارت خواهد بود از:

$$\langle d[\beta = 0] \rangle = \frac{N}{4c}$$

- c. 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 rewired شده و انتهای آن یال به یک نود که بصورت تصادفی انتخاب شده است، وصل می‌شود. محاسبه کنید متوسط فاصله نودها بعد از rewired شدن چقدر است؟

۳. یکی از مجموعه داده‌های جهت‌دار از آدرس زیر <http://snap.stanford.edu/data/index.html#socnets> را انتخاب کنید. با استفاده از پکیج SNAP برای پایتون^۱ برای این مجموعه داده موارد زیر را محاسبه و گزارش کنید.^۲

- ا. تعداد نودها
- ب. تعداد یال‌های جهت‌دار و غیر جهت‌دار
- ج. میانگین ضریب خوشه‌بندی
- د. تعداد مولفه‌های متصل ضعیف
- ه. تعداد نودها در بزرگترین مولفه همبندی قوی
- و. توزیع درجات ورودی و توزیع درجات خروجی را به فرم log-log ترسیم نمایید.

۴. در این تمرین^۳ قصد داریم سه شبکه‌ی تصادفی اقدوش-رینی، مدل جهان کوچک و یک شبکه از دنیای واقعی را مقایسه کنیم. تمامی این شبکه‌ها بدون جهت هستند.

شبکه تصادفی: با استفاده از پکیج SNAP، یک شبکه‌ی تصادفی اقدوش-رینی $G(n,m)$ ایجاد کنید که $n=5242$ تعداد نودها و $m=14484$ تعداد یال‌ها است که یال‌ها به صورت تصادفی قرار می‌گیرند. برای تشکیل این شبکه از متدهای آماده در SNAP استفاده نکنید و شبکه را خودتان تشکیل دهید.

مدل جهان کوچک: یک شبکه از مدل جهان کوچک به این شکل ایجاد کنید. $n=5242$ تعداد نودهای شبکه است که این نودها مانند یک حلقه به هم متصل هستند (یعنی هر نود به نود سمت چپ و راستش متصل می‌شود؛ مثلاً نود ۳۹۹ به نودهای ۳۹۸ و ۴۰۰ متصل شود). در نتیجه در این مرحله ۵۲۴۲ یال خواهیم داشت. سپس هر نود را به همسایه‌ی همسایه‌هایش متصل

^۱ دانلود و مطالعه از لینک <http://snap.stanford.edu/snappy>.

^۲ این تمرین برگرفته از سوال‌های اول تا سوم تمرین زیر می‌باشد:

<http://snap.stanford.edu/class/cs224w-2017/homeworks/hw0/hw0.pdf>

^۳ این تمرین برگرفته از سوال اول تمرین زیر می‌باشد:

<http://snap.stanford.edu/class/cs224w-2017/homeworks/hw1/hw1.pdf>

کنید (مثلا نود ۳۹۹ به نودهای ۳۹۷ و ۴۰۱ متصل شود). اینجا نیز ۵۲۴۲ یال خواهیم داشت. در نهایت، ۴۰۰۰ جفت نود به صورت تصادفی انتخاب کرده و بین آن‌ها یال ایجاد می‌کنیم. این شبکه در آخر دارای ۱۴۴۸۴ یال می‌شود.

شبکه‌ی دنیای واقعی: فایل دیتاست با نام CA-GrQc.txt (که همراه فایل متن تمرین پیوست شده است) را باز کنید و یال‌های طوقه و تکراری را حذف کنید. این شبکه نیز باید شامل ۵۲۴۲ نود و ۱۴۴۸۴ یال باشد.

یکی از مفاهیم مهم در شبکه excess degree distribution است که احتمال اینکه یک یال به تصادف انتخاب شده در شبکه به یک نود با درجه $k+1$ متصل باشد را بیان می‌کند و با q_k به صورت زیر مشخص می‌شود:

$$q_k = \frac{q'_k}{\sum_i q'_i}, \quad q'_k = \sum_{i \in V} \sum_{(i,j) \in E} I_{[k_j=k+1]},$$

نشان دهید فقط با داشتن توزیع درجات p_k چگونه می‌توان q_k را محاسبه کرد. برای هر سه شبکه‌ی فوق نمودار توزیع excess degree distribution را به فرم log-log و در قالب یک نمودار ترسیم کنید. در مورد تفاوت بین این سه منحنی به دست آمده توضیح دهید.

گزارش کار:

گزارش تمرین را در قالب یک فایل PDF بنویسید و تنها به ارسال کد اکتفا نکنید، بطور مثال توابع، کلاسها و کتابخانه‌های مورد استفاده در کد یا روش اجرای کد و مشکلاتی که برخورد کردید را توضیح دهید. در صورتیکه از مرجعی برای نوشتن تمرین‌ها استفاده کرده‌اید، حتما در فایل گزارش به آن ارجاع دهید زیرا استفاده از مراجع اینترنتی و مقالات و ... بدون ارجاع به روش علمی نمره منفی خواهد داشت. فایل گزارش باید شامل نام، شماره دانشجویی و متن گزارش به تفکیک هر سؤال باشد. گزارش کار بخشی از نمره نهایی این تمرین خواهد بود. لذا در نگارش آن نهایت دقت را داشته باشید.

نحوه ارائه:

حتما به تمام بخش‌های تمرین خود مسلط باشید. در زمان ارائه ممکن است از شما توضیحاتی تکمیلی در مورد تمرین و یا تغییراتی در کد خواسته شود. نمره‌ی اختصاص داده شده برای افرادی که مسلط به کار خود باشند قابل تغییر است. در زمان مشخص شده جهت ارائه حتما حضور داشته باشید. غیبت در روز ارائه به منزله عدم ارسال تمرین است.

نکات ضروری:

- تمرین انفرادی بوده و تحویل آنلاین دارد که زمان آن اعلام خواهد شد.
- به ازای هر روز تاخیر در ارسال و حداکثر تا سقف ۵ روز، ۱۰ درصد از نمره کل تمرین کسر می‌شود.
- فایل ارسالی یک فایل فشرده شده شامل کدها، خروجی‌های هر تمرین و گزارش در قالب یک فایل PDF باشد.
- نام فایل مطابق الگوی ComplexNets_HW3_NAME_STDID.zip باشد. (بجای NAME نام خانوادگی خود و به جای STDID شماره دانشجویی خود را قرار دهید).

موفق باشید.