

تمرین‌های سری اول برنامه نویسی درس نظریه گراف

دانشکده مهندسی و علوم کامپیوتر

مدرس: فرشاد صفایی، اسفندماه ۱۴۰۱

در تمرین‌های زیر بنا داریم تا با برنامه نویسی پایتون تحت بسته نرم افزاری NetworkX هم با این زبان برنامه نویسی و بسته نرم افزاری NetworkX بیشتر آشنا شویم و هم اینکه با مفاهیمی که در کلاس به شکل نظری درباره شبکه ها آموختیم به صورت شهودی و عملی نیز آشنا شویم.

۱- به همراه این تمرین، یک متن انگلیسی (text-1) به شما تحویل میشود. این متن برگرفته از کتاب "منشاء بشر" نوشته چارلز داروین و تقریباً حاوی ۱ میلیون کارکتر اسکی (در حدود ۱۶۰۰۰۰ کلمه) است. فایل دوم (text-2) را قدری پیش پردازش کرده‌ایم تا تمامی کاراکترهای غیر الفبایی البته به جز کاما، نقطه، فاصله و بازگشت به سر سطر (CR) حذف شوند. مضافاً، همگی حروف بزرگ با نوع کوچک خود به نمایش در آمده‌اند. بدین سیاق، اندازه الفبا از ۴۰ سمبل (۲۶ حرف + ۱۰ عدد + ۴ کارکتر غیر الفبا) تشکیل شده است. در کلاس درس برای شما عزیزان درباره توزیع درجه در شبکه های مختلف و اینکه با هم تفاوت دارند صحبت کردیم. به ویژه درباره توزیعهای درجه دم کلفت در برخی شبکه ها سخن گفتیم. در این تمرین می‌خواهیم توزیع درجه شبکه را در مورد کتاب داروین مورد آزمون قرار دهیم.

خواسته ۱-۱: جدولی مطابق زیر فراهم کنید و برای کلمات مندرج در آن، رتبه و فراوانی هر کدام را از متن اصلی استخراج کرده و در ستون مربوطه درج کنید.

کلمه	رتبه	فراوانی
the		
eye		
of		
existence		
hair		
and		
nature		
month		
head		
darwin		
size		
touch		
second		
conclusion		
revolution		
muscles		
necessity		
public		
body		
Within		
to		
in		
distribution		
words		

خواسته ۱-۲: نموداری ترسیم کنید که محور عمودی آن، فراوانی کلمات مندرج در جدول بالا و محور افقی آن بیانگر رتبه آنان باشد.^۱ برای وضوح بهتر میتوانید نمودار به دست آمده را به شکل لگاریتم مضاعف (double-log) ترسیم کنید. تفسیر این نمودار چیست؟ با کمک نرم افزار NetWorkX شبکه‌های دنیای کوچک، تصادفی و مقیاس-آزاد را با فرض تعداد گره و یال تقریباً مساوی با کلمات کتاب داروین و با یک متوسط درجه گراف تولید کرده و توزیع درجه آنها را ترسیم کنید. چه فرقی یا شباهتی بین توزیع واقعی کلمات کتاب داروین و شبکه‌های مصنوعی ساخته شده توسط شما وجود دارد؟

۲- در سر کلاس بارها درباره احتمالات ناهمبندی (*disconnection*) و ایزوله بودن (*isolation*) در گرافها صحبت کردیم. با نرم افزار مربوطه تعدادی گراف از هر سه مدل را بسازید. توجه داشته باشید که در هر بار آزمایش، شبکه‌ها بایستی به شکل تصادفی و مستقل تولید گردند.

خواسته ۲-۱: نمودار احتمال ناهمبندی شبکه را برحسب درصد خرابی تصادفی در یالها و گره‌ها برای هر یک از شبکه‌های فوق ترسیم کنید. منظور از خرابی تصادفی آن است که هر گره یا یال با احتمال برابر ممکن است به خرابی دچار گردد. نمودارهایی فراهم کنید که محور افقی آن درصد خرابی در مولفه‌های (یال/گره) شبکه‌های مورد نظر و محور عمودی آن برابر با احتمال ناهمبندی شبکه باشد. نمودارهای به دست آمده را برای هر شبکه ترسیم و با هم مقایسه و تجزیه و تحلیل کنید. درباره نتایج هر نمودار تحلیل خود را بنویسید و میزان آسیب پذیری یا استحکام شبکه‌ها را با یکدیگر مقایسه کنید. محور افقی و عمودی را به شکل مناسب مقیاس‌بندی کنید تا نمودارهای واضح و خوانایی تولید گردد.

خواسته ۲-۲: احتمال ایزوله‌شدن شبکه‌هایی را که در بخش پیشین ساختید برآثر خرابیهای تصادفی محاسبه کنید. مقصود از ایزوله شدن یعنی اینکه گره‌ای در شبکه یافت شود که درجه آن برابر صفر باشد. برای هر یک از شبکه‌های فوق نموداری ترسیم کنید که محور افقی آن درصد خرابی در مولفه-های (یال/گره) شبکه و محور عمودی آن برابر با احتمال ایزوله‌شدن باشد.

در هر کدام از خواسته‌های قبلی لازم است که برای افزایش دقت شبیه‌سازی، تعداد دفعات اجرای شبیه‌ساز به عنوان یک پارامتر ورودی قابل تعیین باشد و هم اینکه نشان داده شود که چگونه با افزایش تعداد دفعات شبیه‌سازی به دقت نتایج آن افزوده میگردد. ضمناً لازم است در پایان شبیه-سازی به کمک توزیع‌های نرمال یا t دانشجویی، دقت و نیز فاصله اطمینان شبیه‌سازی محاسبه گردد.

نکات مهم: شما دانشجویان عزیز میتوانید به بخشهای مختلف به سلیقه خود افزونه‌هایی را بیفزایند که نمره مثبت به آنها تعلق خواهد گرفت. برای مثال، ایجاد فرم GUI برای وارد کردن مناسب داده‌ها و پارامترهای مساله، امکان اعمال خرابیهای عمدی به جای تصادفی در مولفه‌های گراف، نمایش بصری هر شبکه، امکان تعریف یک گراف دلخواه توسط کاربر و امثال آن.

موفق باشید

^۱- می‌توانید نمودار را در مقیاس log-log (یعنی هر دو محور لگاریتمی باشد) ترسیم کنید.