



دانشگاه صنعتی امیرکبیر

(پلی تکنیک تهران)

دانشکده مهندسی کامپیوتر و فناوری اطلاعات

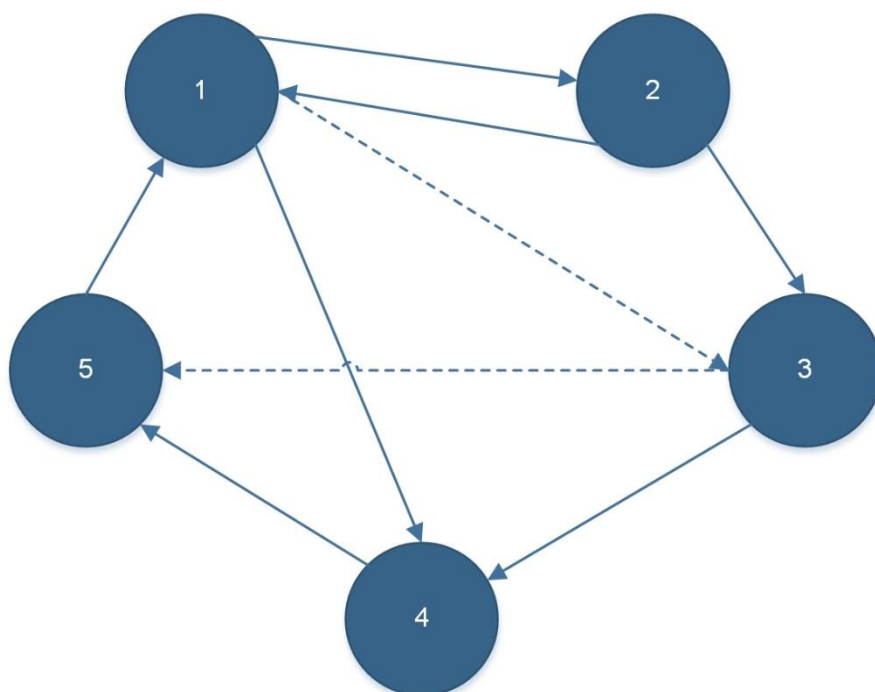
تمرین سری دوم الگوریتم‌های شبکه‌های پیچیده

بهار ۱۳۹۷

سؤال ۱

گراف زیر را در نظر بگیرید.

احتمال یا امتیازی که روش‌های زیر برای وجود یال برای گراف زیر می‌دهند را محاسبه کنید. این محاسبه را فقط برای یال‌هایی که با خط‌چین مشخص شده‌اند انجام دهید. برای هر مورد محاسبات را انجام داده و در گزارش ثبت کنید.



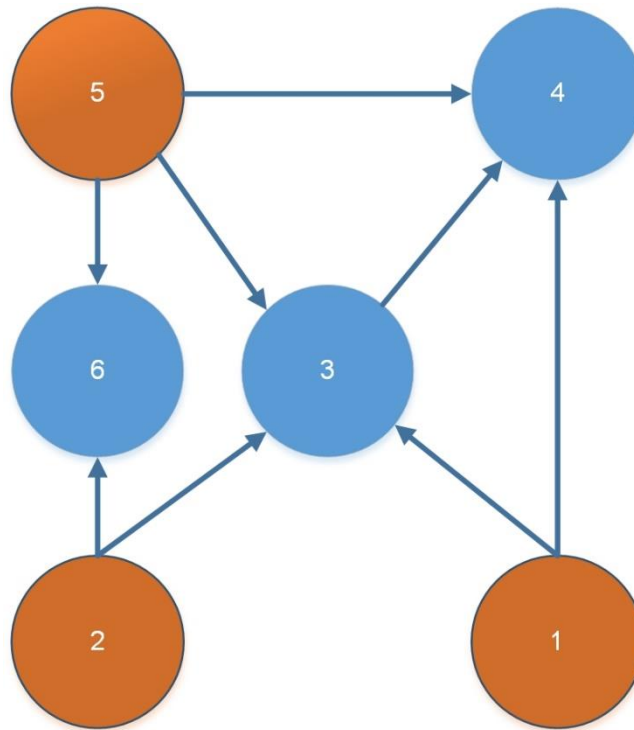
- بیشینه درست‌نمایی (MLE)
- تعداد همسایه مشترک نرمال شده
- کتر^۱ (تا ۲ سطح)

بعد از محاسبه امتیازات یادشده، فرض کنید یال‌ها در گراف وجود دارند. یک حد آستانه تعیین کنید و برای هر ۳ امتیاز از همان حد آستانه استفاده کرده و دقت الگوریتم‌ها را گزارش کنید.

سؤال ۲

¹ Katz

مدل آبخاری مستقل^۲ را برای گراف زیر در نظر بگیرید.



تابع انتشار σ را برای گراف فوق محاسبه کنید (گره‌های ۱ و ۲ و ۵ گره‌های آلوده اولیه هستند). برای راهنمایی می‌توانید به لینک <https://www.utdallas.edu/~dxd056000/cs6301/unit2-2.ppt> مراجعه کنید.

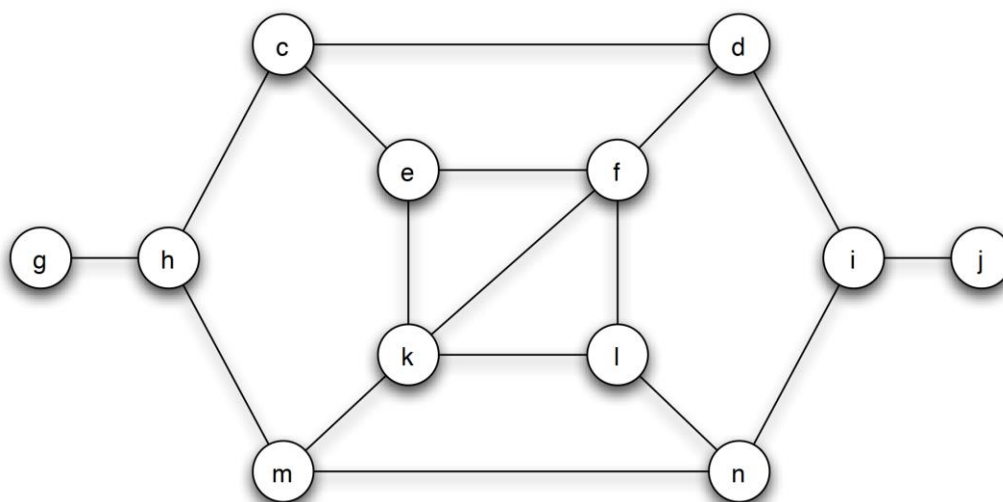
قسمت اختیاری: برنامه‌ای بنویسید که ماتریس مجاورت را ورودی گرفته و تابع انتشار σ را به عنوان خروجی محاسبه کند.

سؤال ۳

گراف زیر را در نظر بگیرید. برای انتشار در این گراف از مدل آستانه خطی^۳ با حد آستانه $\frac{2}{5}$ استفاده می‌کنیم. فرض کنید دو گره e و f گره‌های آلوده اولیه هستند.

^۲ ICM

^۳ LT



الف) یک خوشه با ظرفیت $1 - q = \frac{3}{5}$ پیدا کنید که از آلودگی کل گراف جلوگیری می‌کند.

ب) فرض کنید اجازه دارید یک گره به مجموعه گره‌های آلوده اضافه کنید. آیا گره‌ای وجود دارد که با اضافه کردن آن به مجموعه گره‌های اولیه، آلودگی در کل گراف منتشر شود؟

ج) حداقل میزانی که باید حد آستانه تغییر کند تا کل گراف آلوده شود، را محاسبه کنید.

سؤال ۴

دیتاست گراف out.actor-movi را در نظر بگیرید. هر سطر شامل دو عدد و بیانگر یک یال است که عدد اول گره مبدأ و عدد دوم گره مقصد را نمایش می‌دهد. این گراف جهت‌دار است. برای این گراف نمونه‌برداری را با روش‌های زیر انجام دهید. برای هر روش به ازای [5%, 10%, 15%, 20%] اندازه گراف نمونه‌برداری را انجام داده و نمودار فراسنج‌های میانگین ضریب خوشه‌بندی، میانگین درجه ورودی، میانگین درجه خروجی، قطر گراف را رسم کنید (در تمامی نمودارهای این سوال، مقدار فراسنج به ازای گراف اصلی را نیز به عنوان مجانب افقی رسم کنید تا ببینید با افزایش اندازه نمونه به مقدار اصلی همگرا می‌شوید؟). محور افقی در این نمودارها اندازه نمونه و محور عمودی مقدار فراسنج مدنظر است. نمودارهای رسم شده را تحلیل کنید.

روش‌های مدنظر:

- گره تصادفی
- یال تصادفی
- پرش تصادفی
- پیمایش اول عمق

• آتش‌سوزی جنگل

خروجی مدنظر: نمودار فراسنج‌های میانگین ضریب خوشه‌بندی، میانگین درجه ورودی، میانگین درجه خروجی، قطر گراف به ازای روش‌های فوق و به ازای [5%, 10%, 15%, 20%] اندازه گراف، تحلیل نمودارهای یادشده (هر روش در کدام ویژگی بهتر عمل کرده است و چرا؟)

سؤال ۵

برای انتشار در گراف‌هایی که بزرگ نیستند می‌توان از روش حریصانه استفاده کرد (هر مرحله گره‌ای را انتخاب کرد که با مجموعه گره آلوده فعلی، بیشترین آلودگی را در پی دارد)؛ اما در گراف‌های بزرگ این روش عملی نیست. یک روش برای محدود کردن دامنه بررسی این الگوریتم ارائه دهید. روش را در گراف out.actor-movie اجرا کنید. برای این سؤال نمودار تعداد نود اولیه آلوده - تعداد نود نهایی آلوده را رسم کنید. برای تعداد نود اولیه آلوده [1%, 2%, ..., 10%] رسم کنید. همچنین حداقل درصد نود آلوده اولیه که به ازای آن بیشتر از ۹۵٪ گراف آلوده می‌شود را گزارش کنید.

خروجی مدنظر: شرح کامل روش پیشنهادی، نمودار تعداد نود اولیه آلوده - تعداد نود نهایی برای تعداد نود اولیه آلوده [1%, 2%, ..., 10%]، همچنین حداقل درصد نود آلوده اولیه که به ازای آن بیشتر از ۹۵٪ گراف آلوده می‌شود.

قسمت اختیاری: آیا می‌توانید با استفاده از ایده برنامه‌نویسی پویا روشی برای بهتر شدن الگوریتم ارائه دهید؟

سؤال ۶

سؤال ۵ را با روش پیشنهادی خودتان که مبتنی بر مفهوم خوشه‌بندی است تکرار کنید.

خروجی مدنظر: شرح کامل روش، نمودار تعداد نود اولیه آلوده - تعداد نود نهایی برای تعداد نود اولیه آلوده [1%, 2%, ..., 10%]، همچنین حداقل درصد نود آلوده اولیه که به ازای آن بیشتر از ۹۵٪ گراف آلوده می‌شود.

سؤال ۷

در این سؤال، هدف پیش‌بینی لینک با استفاده از معیارهای شباهت در گراف و خوشه‌بندی سلسله مراتبی است. گراف out.actor-movie را در نظر بگیرید. برای معیارهای شباهت ذکرشده، ماتریس شباهت را محاسبه کنید. هر سطر (گره) این ماتریس بیانگر یک داده است. خوشه‌بندی سلسله مراتبی (با معیار average link) را بروی ماتریس انجام داده و درخت سلسله مراتبی را به دست بیاورید. با استفاده از درخت سلسله مراتبی احتمال وجود یال را بین دو گره‌ای که قصد پیش‌بینی یال برای آن دارید، را محاسبه کنید. بعد از انجام فرآیند یادشده، به ازای هر معیار شباهت، برای هر زوج یال یک احتمال به دست می‌آید. با تجمیع نتایج (روش پیشنهادی خودتان) احتمال وجود یال را گزارش کنید. معیار precision را گزارش کنید.

معیارهای شباهت:

- طول کوتاه‌ترین مسیر
- معیار Katz
- تعداد همسایه مشترک نرمال شده

خروجی مدنظر: شرح روش پیشنهادی، گزارش معیار precision.

نکته: در این سوال ۷۰٪ ماتریس مجاورت را برای آموزش و ۳۰٪ مابقی را برای آزمون استفاده کنید. در نمونه‌برداری برای آموزش با مشکل عدم تعادل در داده‌گان دو کلاس مواجه می‌شوید. راهی برای رفع آن ارائه دهید و در گزارش روش خود را شرح دهید.

سؤال ۸

دیتاست **digits** را در نظر بگیرید. با استفاده از روش پیشنهادی خودتان، دیتاست را به گراف تبدیل کنید.

خروجی مدنظر: شرح کامل روش.

حال گراف ایجادشده در مرحله قبل را با روش پیشنهادی خودتان دسته‌بندی^۴ کنید. سپس با استفاده از فایل **realidx** مقدار دقت^۵ را گزارش کنید.

خروجی مدنظر: شرح روش پیشنهادی، گزارش معیار دقت.

- در صورت هرگونه سؤال یا ابهام به idin47@gmail.com ایمیل بزنید.

⁴ classify

⁵ accuracy