به نام پاک آفریدگار

پروژه درس شبیه‌سازی کامپیوتری

دانشکدهٔ مهندسی و علوم کامپیوتر

مدرس: فرشاد صفایی، اردیبهشت ماه ۱۴۰۲

شبیه‌سازی و تحلیل تاب آوری و قابلیت اطمینان سیستم‌های مبتنی بر شبکه

۱- مقدمه

در شبکه‌های اجتماعی، افراد و سازمان‌ها به‌واسطه علایق، سلایق، عقاید و منافع مشترک با یکدیگر در ارتباط‌اند و بسیاری از سیستم‌هایی که به عنوان یک سیستم پیچیده ادراک می‌کنیم در حقیقت درجهٔ زیادی از تحمل‌پذیری در برابر اشکالات را از خویش به نمایش می‌گذارند. در اثر خطا و خرابی، گره‌ها ممکن است از شبکه حذف شده و در نتیجه، ویژگی‌های محلی یا سراسری شبکه و نیز عملکرد صحیح گره‌های سالمِ آن‌را دستخوش تغییر سازند. در شبکه ممکن است دو مدل اِشکال اتفاق بیفتد که شامل خرابی در یال‌ها و گره‌های شبکه است. هنگام وقوع خرابی در گره، تمامی یال‌های متصل بدان گره نیز خراب شده و به صورت اشکال‌دار (معیوب) برچسب زده می‌شوند. اشکالات نه تنها ممکن است توان عملیاتی و محاسباتی شبکه را کاهش دهند، بلکه همچنین ممکن است ساختار و توپولوژی شبکه را نیز دستخوش تغییر ساخته و به ناهمبندی[1] شبکه منجر شوند. برطبق تعریف، یک شبکه را در صورتی همبند گویند که دست‌کم مسیری بین هر دو مؤلفهٔ سالم آن وجود داشته باشد؛ در غیر این‌صورت گراف ناهمبند است.

اشکالات و حملات در شبکه می‌توانند به کاهش کارآیی آن منجر شوند که این امر با توجه به حذف انتخابی گره‌ها یا یال‌ها در شبکه صورت میپذیرد. در حالت کلی، این امر بیانگر معیاری از کاهش عملکرد شبکه در برابر خرابی‌های تصادفی یا حملات بدخواهانه و هدف‌مند است. اهمیت ارزیابی آسیب‌پذیری شبکه در برابر اشکالات و حملات به این دلیل است که بتوانیم راه‌کارها و راهبردهایی را اتخاذ کنیم تا شبکه را در برابر این قبیل حملات و اشکالات مورد محافظت قرار دهیم. بدین سیاق، بایستی با انواع روش‌های راهبردی حمله آشنا باشیم و دانش خویش را در این زمینه تقویت کنیم.

توپولوژی و ساختار شبکه‌ها معیار مهمی برای درک و شناخت آنها است و برای تبیین و استخراج تعاملات و خواص موجود در شبکه‌ها، شماری از معیارهای مبتنی بر آماره‌های شبکه توسعه یافته‌اند که هریک از آن‌ها در زمینه‌ای مختلف و مشخصی، موضوعیت، کاربرد و سودمندی دارند. در این پروژه تلاش بر این است که با شبیه‌سازی و تحلیل داده‌های حاصل از آزمونهای شبیه‌سازی، ببینیم که معنای دقیق هر یک از معیارها و فرمول‌بندی آنها را درک کرده و بفهمیم که چگونه میتوان نتایج هر معیار را به نحو مناسبی تفسیر کرد. بدین ترتیب، این زمینهٔ پژوهشی اصولاً به دنبال آن است تا راه‌کارها و مکانیزمهایی را جهت بهبود اتصال‌پذیری و استحکام گرافها و شبکه‌های پیچیده در برابر خرابی‌های تصادفی و حملات سیستماتیک جستجو کند. این معیارهای مختلف و متنوعی عمدتاً به نظریهٔ گراف و مطالعه آن اختصاص دارند.

در گزینش و دسته‌بندی معیارهای استحکام، تمرکز با اتکا بر جستجو در ادبیات تحقیق و اساساً از طریق گوگل اسکولار به کمک طیف وسیعی از کلمات کلیدیِ مرتبط میتواند صورت بپذیرد. میتوان تعدادی معیاری استحکام را استخراج کرد که برخی از آن‌ها ممکن است مشخصات مشابهی را به اشتراک بگذارند. در این پروژه از میان چندین دسته معیاری که به منظور ارزیابی و تحلیل قابلیت اطمینان گرافها و شبکه‌ها وجود دارند، ما دسته معیارهای مبتنی بر طیف[2] گراف و شبکه‌ها را مورد توجه قرار میدهیم.

۲- کلاس معیارهای مبتنی بر طیف گراف‌ها و شبکه‌ها

امروزه بخش مهمی از نظریه جبری گراف به بررسی و مطالعه طیفِ ماتریس مجاورت گراف (طیف گراف) اختصاص یافته است. برطبق تعریف، منظور از طیف گراف، ماتریس‌هایی است که به شکل یکتایی ساختار آن گراف را نمایش میدهند. برای مثال، طیف ماتریس مجاورت، ماتریس لاپلاسین و ماتریس فاصله از این دست به شمار میروند. اگر برای بردار n بُعدی ناصفر Ci و عددی مانند λi، تساوی ACi=λiCi برقرار باشد، آنگاه Ci بردار ویژه ماتریس مجاورت A و متناظر با مقدار ویژه λi نامیده میشود. در نظریهٔ طیفی گراف، Ci بردار ویژه و λi مقدار ویژه گراف G نام دارد. اگر G گرافی با n راس باشد، ماتریس مجاورت آن n×n است و دارای n بردار ویژه مستقل خطی است. به دنباله $\lambda_1, \ldots, \lambda_n$ طیف گراف گفته میشود و دو گراف n راسی که طیف یکسانی دارند را هم‌طیف[3] میگویند. مقادیر ویژه همگی حقیقی هستند اما الزاماً ممکن است از یکدیگر متمایز نباشند؛ یعنی $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$.

گراف‌ها را میتوان به شکل مستقل یا به کمک ماتریس‌های مرتبط با آنها مطالعه کرد. یکی از این مهمترین این ماتریسها پس از ماتریس مجاورت، ماتریس لاپلاسین گراف نام دارد. ماتریس لاپلاسین، L، عبارت از Δ-A است که Δ ماتریس قطری درجه است که قطر اصلی شامل درجاتی از رئوس گراف و سایر درایه‌ها شامل مقادیر صفر هستند. ماتریس L، یک ماتریس متقارن و نیمه معین مثبت است که مجموع سطرهای آن برابر صفر است؛ در نتیجه، مقادیر ویژه L، حقیقی و مثبت هستند و کوچکترین آنها برابر با صفر است. در ادامه، برخی از مهمترین معیارهای ارزیابی استحکام مبتنی بر طیف گراف معرفی میشوند.

---

[1] Disconnection
[2] Spectrum
[3] Co-spectral

۲-۱. شکاف طیفی[۴]

در گراف‌های غیرجهت‌دار، بزرگترین زوجِ ویژهٔ (مقدار و بردار ویژه) ماتریس مجاورتِ متقارن از درجه اهمیت زیادی برخوردارند. بزرگترین مقدار ویژه ماتریس مجاورت، شعاع طیفی نام دارد که کاربرد آن در مدل انتشار است. هر قدر شعاع طیفی گراف بزرگتر باشد، استحکام آن گراف نیز بیشتر خواهد بود. برطبق تعریف، شکاف طیفی، $\Delta\lambda$، معیاری مستخرج از طیف گراف و برابر با اختلاف بین بزرگترین دو مقدار ویژه ماتریس مجاورت گراف است. به بیان ریاضی

$$\Delta\lambda = \lambda_n - \lambda_{n-1}$$

این معیار میتواند میزان استحکام شبکه را در برابر تغییرات توپولوژیکی اندازه‌گیری و تحلیل کند. بدین ترتیب، شکاف طیفی کم، بیانگر تعداد نقاط مفصلی[۵] کمتر است که در هنگام وقوع خرابی لینک/ گره موجب اِفراز شبکه میشوند. نقاط مزبور نقاطی هستند که حذف آنها همراه با حذف یال‌های متصل به آن‌ها سبب ناهمبندی در گراف خواهد شد.

۲-۲. اتصال‌پذیری جبری[۶]

برطبق تعریف، اتصالپذیری جبری عبارت از دومین کوچکترین مقدار ویژه ماتریس لاپلاسین گراف است. مقادیر ویژه L را میتوان از کوچک به بزرگ به شکل $0 = \mu_1 \le \mu_2 \le \ldots \le \mu_n$ مرتب ساخت. بدین سیاق، اتصالپذیری جبری برابر با μ_2 است. اگر و فقط اگر گراف ناهمبند باشد، آنگاه $\mu_2=0$ خواهد بود و وقتی همگی گره‌ها متصل باشند، با تعداد کل رئوس گراف برابر خواهد بود. اتصال پذیری جبری هر گراف ناکامل همیشه از اتصالپذیری راسی (κ_v) و یالی (κ_e) آن بزرگتر نخواهد بود. یعنی همواره $\mu_2 \le \kappa_v \le \kappa_e \le \delta$ که δ مینیمم درجه گراف است. مضافاً، عدد تکرار ریشه‌های صفر مقادیر ویژه با تعداد مولفه-های همبند گراف متناظر است. براین اساس، طبیعی به نظر میرسد که بزرگتر بودن معیار اتصال‌پذیری جبری میتواند متناظر با استحکام بیشتر گراف باشد.

در این پروژه تمرکز ما بیشتر بر گرافهای تصادفی اسپارس با سایز n گره و m یال است؛ یعنی $m=O(n)$ یا به عبارتی، یالها به شکل خطی با تعداد رئوس n مقیاسبندی میشوند. در این گرافها، متوسط درجه با رابطه زیر قابل بیان است

$$\prec d \succ = 2m / n = O(1) \quad as \quad n \to \infty$$

در شبکه های تصادفی مدل اردوش-رینی (ER)، متوسط فاصله با $\ln n$ متناسب است. به همین دلیل $m \sim O(n\log n)$. به بیان دیگر، اگر در این قبیل شبکه‌ها $O(n\log n)$ یال وجود داشته باشد، آنگاه شبکه تصادفی ER کاملاً متصل خواهد بود.

در گرافهای تصادفی δ-منتظم برای $\delta \ge 3$ داریم

$$\mu_2 = \delta - 2\sqrt{\delta - 1} \quad \text{as} \quad n \to \infty$$

در گرافهای شبه-منتظم با درجه بین ۲ و ۳، متوسط درجه برابر $\prec d \succ = 2 + p$ است که $0<p<1$ و در حالت مجانبی وقتی $p\to 0, n\to\infty$، اتصالپذیری جبری تقریباً برابر $\mu_2 \approx p^2 / 4$ خواهد شد.

در انتها، لازم است اشاره کنیم که معیار اتصال‌پذیری جبری تنها به توپولوژی وابسته است و در ضمن معیاری غیرنرمال است. از سویی، برای یک شبکه ناهمبند، مقدار آن صفر میشود درحالی‌که ممکن است هنوز مولفه‌های همبندی با سایز به قدر کافی بزرگ در شبکه وجود داشته باشند. به همین دلیل گاهی اوقات بهتر است استحکام شبکه را براساس تمامی مقادیر ویژهٔ لاپلاسین، و نه فقط دومین آنها (اتصال‌پذیری جبری)، ارزیابی کنیم.

۲-۳. اتصال‌پذیری (مقدار ویژهٔ) طبیعی[۷]

اتصالپذیری طبیعی از شاخصی موسوم به مرکزیت زیرگراف سراسری الهام گرفته شده است. معنای فیزیکی این معیار ساده و روشن است و میتواند به عنوان شاخصی برای بیان افزونگی مسیرهای آلترناتیو با تعیین تعداد گام‌های وزن‌دار برای تمامی طول‌ها در گشتهای گراف باشد.

معیار اتصال‌پذیری طبیعی (مقدار ویژه طبیعی) برحسب تعداد گامهای بسته در یک گراف است و با مجموع مقادیر ویژه در ارتباط است. یعنی میتوان گام‌زدن ذره یا تعداد مدارهای اولری در شبکه را بررسی کرد. مدارهای اولری به طول k هستند ($k\ge 3$) و با دور همیلتونی تفاوت دارند. مقدار $k=2$ به معنای رفتن به یک راس و بازگشت از همان راس و پیمایش دوباره یک یال است. مزیت این معیار آن است که به شکل اکیداً یکنوا[۸] با افزودن تعداد یال‌ها افزایش می‌یابد و در هر دو حالت شبکه همبند و ناهمبند کار میکند و برخلاف اتصالپذیری جبری حتی برای شبکه ناهمبند مقدار آن صفر نمیشود. اتصال پذیری طبیعی افزونگی، تعدد و تنوع مسیرهای جایگزین را با مسور ساختن[۹] عدد وزنی گشت‌هایی[۱۰] از تمامی طولهای ممکن تبیین میکند. این معیار مُلهم از طیف گراف است و متوسط مقدار ویژه را ارائه میدهد.

[4] Spectral gap
[5] Articulation points
[6] Algebraic connectivity
[7] Natural Connectivity (Natural Eigenvalue)
[8] Strictly monotonically
[9] Quantify
[10] Walk

یک نکته جالب توجه در گرافها این است که برطبق دانش جبرخطی، trace توان s ام ماتریس مجاورت گراف، یعنی $Tr(A^s)$، عبارتست از

$$n_s = \mathrm{Tr}(A^s) = \sum_{i=1}^{n} \lambda_i^s$$

که λi مقدار ویژه i ام ماتریس مجاورت گراف است. معادله بالا در واقع، تعداد گشت‌های بسته به طول s را با استفاده از ماتریس مجاورت گراف به ما عرضه میکند.

اگر $\phi_s = Tr(A^s) / n$ تعریف کنیم، به معنای متوسط تعداد گشتهای بسته به طول s در گراف G خواهد بود. همچنین اگر $\rho(x)$ بیانگر تابع چگالی طیف گراف G باشد، میتوان نوشت

$$\lim_{n\to\infty} \int x^s \rho(x) dx = \phi_s$$

به عبارت دیگر، مقدار این متوسط، همان گشتاور s ام تابع چگالی طیف گراف حول مبدأ است. لازم به ذکر است که از منظر شهودی، در شبکه‌های تصادفی هرقدر که سایز شبکه افزایش پیدا میکند ($n\rightarrow\infty$)، با فرض ثابت بودن s (طول گشت)، احتمال یافتن حلقه‌هایی به طول s کم خواهد شد و گراف از این منظر مشابه با یک درخت عمل میکند.

معمولاً فراوانی گشت‌هایی به طول کم یا متوسط محتملاً بیشتر است (توزیع طول گشت)، به همین دلیل برای پرهیز از واگرایی، معمولاً ns را به عوض n به !s تقسیم کرده و مجموع وزنی تعداد گشت‌های بسته n_sبه طول s را در گراف تعریف میکنند.

$$S = \sum_{s=1}^{\infty} \frac{n_s}{s!}$$

رابطهٔ فوق را میتوان به شکل زیر بازنویسی کرد

$$S = \sum_{s=1}^{\infty} \frac{n_s}{s!} = \sum_{s=0}^{\infty}\sum_{i=1}^{n} \frac{\lambda_i^s}{s!} = \sum_{i=1}^{n}\sum_{s=0}^{\infty} \frac{\lambda_i^s}{s!} = \sum_{i=1}^{n} e^{\lambda_i}$$

که فرمول فوق همان معیار مرکزیت زیرگراف سراسری است که در آن، به گشت‌های کوتاه‌تر، وزنهای بیشتر (اهمیت بیشتر) داده شده است. به دلیل افزایش مقدار S با افزایش n، پارامتر S را مقیاس‌بندی کرده و معیار مقدار ویژه متوسط یا اتصالپذیری طبیعی (مقدار ویژه طبیعی) را به صورت زیر تعریف میکنیم

$$\bar{\lambda} = \ln(S / n) = \ln[\sum_{i=1}^{n} e^{\lambda_i} / n]$$

هرقدر که اتصال پذیری طبیعی شبکه‌ای بیشتر باشد، استحکام آن شبکه نیز بیشتر خواهد بود.

در بین گرافها، گراف کامل K_n بیشترین مقدار اتصالپذیری طبیعی و گراف پوچ[11] (شامل گره های ایزوله) مینیمم اتصالپذیری طبیعی را دارد.

$$\begin{cases} K_n : \lambda_1 = n-1, \lambda_2 = \lambda_3 = \ldots = \lambda_n = -1 \\ \text{Null Graph} : \lambda_1 = \lambda_2 = \ldots = \lambda_n = 0 \end{cases}$$

همچنین، کران بالای اتصال‌پذیری طبیعی عبارتست از

$$0 \le \bar{\lambda} \le \ln((n-1)e^{-1} + e^{n-1}) - \ln n \approx n - \ln n$$

۲-۴. نسبت تقارن[12]

برای آنکه درک درستی از ویژگی تقارن در گراف داشته باشیم، ابتدا لازم است تعاریف زیر را ارائه دهیم.

تعریف ۱ (خودریختی[13]): *جایگشتی مانند π از مجموعه رئوس گراف، V_G، است که لینک‌ها را حفظ میکند. بدین معنی که، اگر u,v دو راس مجاور هم (یال) از گراف G باشند، آنگاه π_u و π_v نیز مجاور هم (یالی از G) خواهند بود. به بیان دیگر، G با خودش یکریخت است و داریم*

$$\begin{aligned} &\mathrm{Aut}_G \times V_G \to V_G \\ &\mathrm{Aut}_G \triangleq \{\exists \pi : V_G \xrightarrow[\text{onto}]{1-1} V_G \;\; \ni \;\; \forall u,v \in V_G, (u,v) \in E_G \Rightarrow (\pi_u, \pi_v) \in E_G\} \end{aligned}$$

بایستی اشاره کنیم که خودریختی شکلی از تقارن است که گراف به خودش نگاشت میشود و همزمان اتصالپذیری یالی-راسی آن حفظ میگردد. پس میتوان خودریختی را به یک معنا، تقارن در شی دانست و آنرا به روشی برای نگاشت یک شئ به خودش تعبیر کرد به نحوی که تمامی ساختار شئ مورد نظر دست نخورده باقی بماند. در واقع، خودریختی، یعنی جایگشتی از شماره رئوس یک گراف. یعنی اگر گرافی دارای راسهایی با برچسب ۱ تا n-1 باشد و هر جایگشتی از

[11] Null graph
[12] Symmetry ratio
[13] Automorphism

این اعداد یک گراف را شکل دهد، آنگاه این گرافها با یکدیگر یکریخت[۱۴] خواهند بود. لازم است اشاره کنیم که مجموعه خودریختی‌های یک گراف تحت عمل ترکیب توابع، تشکیل یک گروه میدهند. توجه داشته باشید که نگاشت همانی یک گراف به خودش نیز همیشه یک خودریختی است که گاهی اوقات به آن خودریختی بدیهی[۱۵] گفته میشود.

تعریف ۲: *گراف G را مشابهت گره‌ای (NS) گویند اگر و تنها اگر برای هر دو گره دلخواه u,v داشته باشیم*

$$\forall u,v \in V_G, \exists \pi \in \mathrm{Aut}_G \ni \pi_u = v$$

بدین ترتیب، ویژگی NS در گراف به معنای آن است که همگی گره‌ها در آن مشابه با یکدیگر به نظر برسند. معمولاً به ویژگی NS، راس–انتقالی[۱۶] نیز گفته میشود. در حقیقت، گراف راس–انتقالی، گرافی است که هر زوج راس آن تحت برخی عناصر گروه خودریختی‌اش با یکدیگر معادل باشند. برای مثال، گراف رادو (Rado)، گراف مسیر، درختهای منتظم، گراف کیلی (Cayley) و گراف تتراهدرون بریده شده[۱۷]، همگی نمونه‌هایی از گرافهای راس–انتقالی (NS) به شمار میروند. توجه داشته باشید که ویژگی NS الزاماً به معنای وجود تقارن در گراف نیست. ویژگی مشابهت گره‌ای در کاربردهای پردازش موازی ، مسیریابی بهینه و توزن بار کاربرد دارد. ویژگی NS در گرافها همچنین میتواند به معنای یکسان دیده شدن اهمیت گره‌ها از دید مهاجم و در نتیجه تحمل پذیری بیشتر این نوع گرافها در برابر حملات تصادفی و هدفمند باشد.

تعریف ۳: *گراف G متقارن است اگر و تنها اگر برای هر دو لینک (u,v) و (x,y) از مجموعه یالهای G داشته باشیم*

$$\forall (u,v),(x,y) \in E_G, \exists \pi \in \mathrm{Aut}_G \ni \ \pi_u = x \wedge \pi_v = y$$

بدین ترتیب میبینیم که ویژگی تقارن، حافظ لینک است و وجود آن میتواند به این معنا باشد که همگی لینکها مشابه به نظر میرسند. به عبارت دیگر، برای هر دو لینک e1=(u,v) و e2=(x,y) از مجموعه یالهای گراف G، جایگشتی مانند π وجود دارد که $\pi_{e1}=e_2$. به همین دلیل به ویژگی تقارن گاهی اوقات یال–انتقالی نیز گفته میشود. برای مثال، گراف کامل K_n و گراف‌های سیکل Cn متقارن هستند.

براین اساس و با توجه به اهمیت ویژگی تقارن در شبکه‌های مستحکم، میتوان معیار نسبت تقارن، r≥1، را به شکل زیر تعریف کرد

$$r = \varepsilon \,/\, (D+1)$$

طوریکه نماد ε به تعداد مقادیر ویژه متمایز ماتریس مجاورت گراف G اشاره دارد و D نیز بیانگر قطر گراف است. برای شبکه هایی با تقارن بالا، 1≤r≤3 است؛ بدین معنی که، همگی گره‌ها مستقل از درجه‌ای که دارند دارای اهمیت هستند و در نتیجه حملات هدفمند و خرابیهای تصادفی تاثیر یکسانی برجای خواهند گذاشت.

۲-۵. انرژی و انرژی لاپلاسین گراف

برطبق تعریف اگر فرض کنید $G(V,\mathrm{E})$ یک گراف سادهٔ همبند بدون‌جهت با n گره و m یال باشد و مقادیر ویژه ماتریس مجاورت گراف $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$ باشند، در اینصورت انرژی گراف G عبارتست از

$$\mathscr{E}(G) = 2\sum\nolimits_{j:\lambda_j>0} \lambda_j = \sum\nolimits_{i=1}^{n} \left|\lambda_i\right|$$

همچنین اگر $G(V,\mathrm{E})$ یک گراف با مشخصات بالا و مقادیر ویژه لاپلاسین $\mu_1,\ldots,\mu_n$ باشد (طیف لاپلاسین گراف)، در اینصورت انرژی لاپلاسین گراف G عبارتست از

$$\mathscr{E}_L(G) = \sum\nolimits_{i=1}^{n} \left|\gamma_i\right|$$

که در آن $\gamma_i = \mu_i - 2m \,/\, n$ به ازای $i = 1,2,\ldots,n$ است که مقادیر ویژه کمکی[۱۸] نام دارند. برطبق یک قضیه، اگر G یک گراف منتظم باشد در اینصورت انرژی و انرژی لاپلاسین آن با هم مساوی هستند.

خواسته ۱: اگر متوسط درجه گرافها را ثابت فرض کنیم، به کمک شبیه‌سازی و تحلیل‌های ریاضی در حالت حدی n→∞ (افزایش سایز گرافها)، بررسی کنید و ببینید که از بین گرافهای مدلهای تصادفی ER، مقیاس–آزاد (SF) و دنیای کوچک واتس–استروگاتز، کدامیک دارای بیشترین اتصالپذیری جبری خواهد بود؟

---

[14] Isomorphic
[15] Trivial automorphism
[16] Vertex-transitive
[17] Truncated tetrahedron
[18] Auxiliary eigenvalue

پارامترهای شکاف طبیعی، اتصالپذیری طبیعی، انرژی و انرژی لاپلاسین این گرافها را محاسبه و با یکدیگر مقایسه کنید. مدل این گرافها در بسته نرم افزاری NetworkX موجود است لذا به برنامه نویسی برای ایجاد آنها نیازی نیست. شما میتوانید شبکه‌هایی با سایزهای مختلف برای مثال، ۱۰ تا ۱۰۰۰ را تولید کنید. برای مدل ER احتمال همبندی را برای مثال $p=0.5$ و برای مدل دنیای کوچک واتس-استروگاتر، احتمال سیمبندی مجدد را برای مثال برابر 0.3 در نظر بگیرید. در مدل مقیاس-آزاد از مدل شبکه باراباشی-آلبرت استفاده کنید و برای مثال پارامتر $m_0=3$ (هسته اولیه شبکه) فرض کنید.

تعریف ۴ (گراف دوبخشی تصادفی شبه منتظم[۱۹] (RSRBG)): *فرض کنید G یک گراف دوبخشی باشد طوریکه در یک بخش، n_1 گره با درجه d_1 و در بخش دیگر n_2 گره با درجه d_2 داشته باشیم. تعداد کل گره‌ها برابر $n=n_1+n_2$ و رابطه $n_1d_1=n_2d_2$ برقرار است. بدین ترتیب میتوان نوشت*

$$\begin{cases} n_1 = nd_2 / (d_1 + d_2) \\ n_2 = nd_1 / (d_1 + d_2) \end{cases}$$

قضیه ۱: متوسط درجه گراف دوبخشی تصادفی شبه منتظم عبارتست از $\prec d \succ = 2d_1d_2 / (d_1 + d_2)$

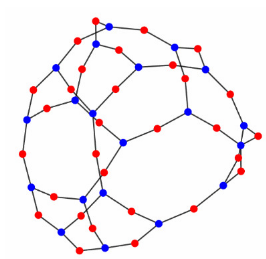
قضیه ۲: در گرافهای $K_{b,n-b}$ که $b<n/2$ و دوبخشی کامل هستند؛ اتصالپذیری جبری $\mu_2=b$ و در حالت حدی $n\to\infty$، متوسط درجه به سمت 2b میل میکند و اتصالپذیری جبری $\delta/2$ خواهد شد که δ مینیمم درجه گراف است.

قضیه ۳: اتصالپذیری جبری در گرافهای دوبخشی $K_{\delta/2,n-\delta/2}$، با شرط $\delta<15$، از گرافهای δ-منتظم بیشتر است.

۳- الگوریتم تولید گراف دوبخشی تصادفی شبه منتظم

دو ظرف (مجموعه/گروه) را در نظر بگیرید. در ظرف اول، d_1 کپی از n_1 راس را قرار داده و از شماره ۱ تا n_1 برچسب گذاری کنید. در ظرف دوم، d_2 کپی از n_2 راس را قرار داده و با برچسبهای n_1+1 تا n_1+n_2 مشخص کنید. سپس از یک مجموعه به سایز n_1+n_2 شروع کنید که در آغاز تهی باشد. یک راس را به تصادف و بدون جایگذاری از ظرف اول انتخاب کنید. راس دوم را نیز به تصادف و بدون جایگذاری از ظرف شماره ۲ انتخاب کنید. بین این دو راس انتخابی، یک یال ترسیم کرده و این کار را آنقدر ادامه دهید تا در ظرف‌های اول و دوم هیچ گره‌ای باقی نمانده باشد. در ضمن ساخت گراف، توجه داشته باشید که امکان تشکیل طوقه و یال چندتایی وجود ندارد. گراف حاصل دوبخشی تصادفی شبه منتظم نامیده شده که با نماد RSRBG(d1,d2) نمایش میدهیم.

خواسته ۲: برنامه‌ای بنویسید که مطابق الگوریتم شرح داده شده در بالا، گرافهای دوبخشی مدل $RSRBG(d_1, d_2)$ را تولید کند. برای مثال، یک نمونه از گراف RSRBG(2, 3) در شکل ۱ نشان داده شده است. پیچیدگی الگوریتمیک برنامه تولید گراف RSRBG را به دست آورید. با استفاده از برنامه نوشته شده، به تعداد کافی از این نوع گرافها را با پارامترهای متفاوت و مشخص تولید کرده و نمودار توزیع درجه، توزیع مقادیر ویژه و نیز پارامترهای شکاف طیفی، اتصالپذیری جبری، اتصالپذیری طبیعی، نسبت تقارن، انرژی و انرژی لاپلاسین را در آنها محاسبه کنید.



شکل ۱. یک نمونه از گراف مدل دوبخشی تصادفی شبه منتظم؛ RSRBG(2, 3)

۴- گراف‌های تصادفی شبه منتظم

این مدل گراف با نماد $RSRG(p;d_1,d_2)$ نمایش داده میشود. فرض براین است که $n_1 = \lfloor n(1-p) \rfloor$ و $n_2=n-n_1$ و $0<p<1$. مدل این گرافها طوری است که امید داریم به تولید گرافهایی با متوسط درجه کم بینجامد.

قضیه ۴: متوسط درجه در گراف مدل $RSRG(p;d_1,d_2)$ عبارتست از $\prec d \succ = d_1(1-p) + pd_2$

[۱۹] Random semi regular bipartite graph

۴-۱. الگوریتم تولید گراف تصادفی شبه منتظم

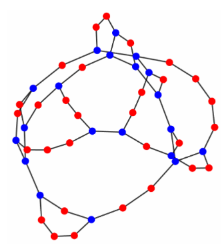
مانند دفعه قبل عمل میکنیم؛ منتها به جای دو ظرف این بار از یک ظرف استفاده میکنیم. در این ظرف، d_1 کپی از n_1 راس را با برچسبهای $1…n_1$ و d_2 کپی از n_2 راس را با برچسبهای $n_1+1…n$ مشخص میکنیم. سپس، دو راس را به تصادف و بدون جایگذاری از ظرف بیرون می‌آوریم و بین آنها با احتمال p یالی را برقرار میکنیم. این کار را آنقدر ادامه میدهیم تا درون ظرف، یک یا هیچ گره‌ای باقی نماند. شبکه حاصل یک $RSRG(p;d_1,d_2)$ است.

قضیه ۵: در مدل $RSRG(p;d_1,d_2)$، تابع ϕ_s عبارتست از

$$\phi_s = (d_1\Phi + d_2\Omega) / (d_1 + d_2)$$

که در آن Φ تعداد گشتهای بسته به طول s با شروع از یک راس به درجه d_2 و Ω تعداد رئوس با درجه d_1 است.

خواسته ۳: برنامه‌ای بنویسید که مطابق الگوریتم شرح داده شده در بالا، گرافهای تصادفی شبه منتظم مدل $RSRG(p;d_1,d_2)$ را تولید کند. برای مثال، یک نمونه از گراف $RSRG(0.4;2,3)$ در شکل ۲ نشان داده شده است. پیچیدگی الگوریتمیک برنامه تولید گراف RSRG را به دست آورید. با استفاده از برنامه نوشته شده، به تعداد کافی از این نوع گرافها را با گزینش مقادیر p، d_1 و d_2، تولید کنید. نمودار توزیع درجه، توزیع مقادیر ویژه و نیز پارامترهای شکاف طیفی، اتصالپذیری طبیعی، نسبت تقارن، انرژی و انرژی لاپلاسین را در آنها محاسبه کنید. چه تفاوتی بین این نمودارها با نمودارهای مدل تصادفی ER وجود دارد؟ همچنین، نتایج و نمودارها را با مدل دوبخشی تصادفی شبه منتظم نیز مقایسه کنید. همچنین، با کمک آزمونهای شبیه‌سازی، نمودارهای تابع توزیع اتصالپذیری جبری را در گراف‌های مدل $RSRG(p;d_1,d_2)$ ترسیم کنید. متوسط درجه را برابر 2+4p است. در آزمونهای شبیه‌سازی $p=1, 0.75, 0.5, 0.25, 0.125$ فرض کنید. سایز گراف $n=500$ گره و تعداد شبیه‌سازیها برای مثال تا ۱۰۰۰ بار انجام دهید. متوسط اتصالپذیری جبری را به کمک شبیه‌سازی مونت-کارلو محاسبه کنید و نمودار pdf توزیع اتصال پذیری جبری را به ازای متوسط درجه $5 \le \prec d \succ$ ترسیم کنید. آیا نمودارهایی که به دست آورده‌اید از توزیع نرمال تبعیت میکنند؟ چرا برخی از این نمودارها ممکن است توزیعهای چندمُدی باشند که حول میانگین متمرکز نشده‌اند.



شکل۲. یک نمونه از گراف مدل تصادفی شبه منتظم؛ $RSRG(0.4; 2, 3)$

تعریف ۵: گراف خط[20] یک گراف بدون جهت G، گرافی است که با L(G) نشان داده میشود بیانگر همسایگی بین یالهای G است. گراف خط L(G) به این صورت ساخته میشود که هر یال G به یک راس در L(G) نگاشت میشود و هر دو یالی که یک راس مشترک دارند، یک یال را بین رئوس متناظر در L(G) تشکیل میدهند.

خواسته ۴: برنامه‌ای بنویسید که گراف خط شبکه‌های تصادفی مدل ER، SF، WS، RSRBG و RSRG را تولید کند. این برنامه در نرم‌افزار NetworkX برای گراف عمومی G وجود دارد. سپس بررسی کنید و ببینید که آیا گراف خط این مدلها به خود آنها شباهت دارد؟ نمودار تابع چگالی احتمال توزیع درجه در مدل گراف خط این مدلها را با توزیع درجه در گراف اصلی مقایسه کرده و شباهتها یا تفاوتها را گزارش کنید. اکنون، نمودار تابع چگالی احتمال توزیع مقادیر ویژه این گرافها را فراهم سازید. چه تفاوتها یا شباهتهایی با یکدیگر دارند؟ در آزمونهای شبیه‌سازی خود میتوانید گرافها را با سایزهای مناسب و گزینش پارامترهای مناسب تولید کنید. برای مثال سایز ۲۰۰ و احتمال همبندی $p=2p_c$ که $p_c=\ln n/n$ نقطه احتمال بحرانی همبندی در گرافهای تصادفی مدل ER است. توجه دارید که اگر مدل گراف SF باشد در اینصورت که توزیع درجه از قانون توانی با نمای مقیاسبندی γ تبعیت میکند.

قضیه ۶: با فرض استقلال درجه گره‌ها در گراف G، تابع چگالی احتمال توزیع درجه در گراف خط L(G) در مدل شبکه های مقیاس-آزاد در حالت حدی $n\to\infty$ عبارتست از

$$\Pr\{D_L = k\} \approx (k+2)^{-\gamma}$$

[20] line graph

اگر G گراف مقیاس–آزاد مدل BA باشد دارای نمای $\gamma\approx 2$ است؛ در حالیکه گراف خط متناظر با آن یعنی L(G)، دارای $\gamma\approx 3$ است. آیا میتوانید تابع چگالی احتمال را برای شبکه های تصادفی و مدل دنیای کوچک مانند رابطه بالا بنویسید؟

خواسته ۵: نموداری ترسیم کنید و در آن تابع چگالی توزیع درجه گراف خط در مدل مقیاس–آزاد را برحسب پارامترهای مختلف γ ترسیم کنید (هر دو محور را لگاریتمی فرض کنید). به تعداد ۱۰۰۰ یا بیشتر گراف مدل SF با سایز مشخص (مثلاً N=500 گره) و متوسط درجه مشخص (برای مثال ۴) تولید کنید و نتایج تجربی حاصل از آزمونهای شبیه سازی و مدل تحلیلی $\Pr\{D_L = k\} \approx (k+2)^{-\gamma}$ از قضیه بالا را با یکدیگر مقایسه کنید. اختلافهای مشاهده شده ناشی از چیست؟ یک راه جالب و ساده برای برای نمایش اختلاف بین دو توزیع یا مدل شبیه‌سازی و مدل تحلیلی میتواند استفاده از معیار فاصله واریانس کل[21] (d_{TV}) باشد که به صورت $d_{TV} = \sum_{k=-\infty}^{\infty} |\Pr\{X=k\} - \Pr\{Y=k\}|$ تعریف میشود. در علم داده‌ها، فاصلهٔ واریانس کل، معیاری برای نمایش اختلاف و فاصله بین دو توزیع احتمال است. یعنی، یک معیار فاصله آماری محسوب میشود. از d_{TV} میتوان برای محاسبه اختلاف بین نتایج شبیه سازی و مدل تحلیلی استفاده کرد. در فرمول d_{TV}، احتمالات Pr{X=k} و Pr{Y=k} میتوانند توابع چگالی احتمال دو توزیع برای مقایسه باشند. شما در آزمایش خود، تعداد گره‌ها را از ۵۰۰ تا ۱۰۰۰ تغییر دهید و نموداری ترسیم کنید که محور عمودی آن d_{TV} و محور افقی آن برحسب سایز شبکه‌های مدل SF باشد. به ازای هر نقطه از نمودار (یعنی به ازای هر سایز شبکه) میتوانید ۱۰۰۰ بار گراف را تولید و سپس نتایج را متوسط‌گیری کنید. هر قدر که نمودار افقی‌تر باشد، اختلاف کمتر و در نتیجه دقت بیشتر خواهد بود. با این حال، توجه داشته باشید که ممکن است شبکه‌هایی باشند که علیرغم توزیع درجه یکسان، عملاً خواص توپولوژیکی متفاوتی داشته باشند.

خواسته ۶: به کمک آزمون‌های شبیه‌سازی یک گراف δ–منتظم (برای مثال ۶–منتظم با ۵۰۰ گره و ۱۵۰۰ یال) بسازید و سپس یکی یکی یالها را به طور تصادفی قطع کنید تا گراف ناهمبند شود. این آزمون را از ۱۰۰ تا ۱۰۰۰ بار انجام داده و احتمال ناهمبندی را محاسبه کنید. همین کار را برای مدل RSRBG(4,12) (برای مثال با سایز ۵۰۰ گره و تعداد یال ۱۵۰۰) انجام دهید. اتصالپذیری جبری، یالی و گره‌ای را در هر بار محاسبه کنید. چه تفاوتی وجود دارد؟ چه اتفاقی برای اتصالپذیری جبری خواهد افتاد؟ چه گرافهای از این جنبه دارای بیشترین قابلیت اطمینان هستند؟

تعریف ۶: *کمر (girth) گراف عبارتست از طول کوتاهترین دور در گراف. اگر گراف فاقد دور باشد، آنگاه کمرش بینهایت است.*

تعریف cage: *گراف δ–منتظم با کمر مشخص g که کمترین تعداد رئوس را داشته باشد یک g-cage نام دارد. برای مثال، گراف ۶–منتظم 5-cage دارای ۴۰ راس است.*

خواسته ۷: از بین مدلهای مختلف گراف که در خواسته‌های ۱ تا ۸ ساختید، کدامیک کمر بیشتری دارند؟ آیا میتوانید گراف یا گرافهایی ارائه دهید که با یک متوسط درجه و سایز مشخص، بیشترین کمر باشند. تاثیر cage و کمر در معیارهای استحکام و اتصالپذیری جبری گراف چیست؟ برای نمونه، با فرض تساوی رئوس، آیا گرافهای RSRBG(2,6) با متوسط درجه ۳ و گراف RSRG(2,6) در قیاس با گرافهای مکعبی ($d_1=d_2=3$) دارای cage و کمر بهتری هستند؟

۵– خرابی گره‌ها براساس توزیع طول عمر و قابلیت اطمینان استوکستیک شبکه‌ها

در بخش‌های پیشین، قابلیت اطمینان گرافها و شبکه‌ها را به شکل ایستا مورد ارزیابی قرار دادیم؛ یعنی، هر مولفه از شبکه (گره) یک متغیر تصادفی برنولی است که تنها میتواند در دو وضعیت on (عملیاتی) یا off (غیر عملیاتی) قرار داشته باشد. در این بخش، وضعیتهایی را بررسی خواهیم کرد که در آن هر یک از مولفه‌های شبکه (گره) و در نتیجه خود شبکه، دارای طول عمر مشخصی است. با این مطالعه میتوانیم قابلیت اطمینان شبکه‌های مختلف را در هر لحظه از زمان برحسب تابعی از زمان محاسبه کنیم و یا کرانهایی را برای آن به دست آوریم. میخواهیم قابلیت اطمینان و تاب آوری شبکه‌هایی را که در بخشهای قبل ساخته بودیم با مطالعه احتمال ناهمبندی، احتمال ایزوله شدن و زمان طول عمر روی تک گره به شکلی پویا آنالیز کنیم. توزیع طول عمر کاربران میتواند دارای خاصیتهای مختلفی مانند NWUE[22] یا NBUE[23] باشد. قرار است بررسی کنیم و ببینیم که با داشتن کدامیک از این خاصیتها، گراف و شبکه مستحکمتر خواهد بود. همچنین در این پروژه با روشهای گرافیکی و متدهای ناپارامتریک برای آزمون مرتبه NWUE آشنا خواهید شد.

مدل خرابی‌های گره‌ها مبتنی بر طول عمر گره‌ها است و بدین ترتیب میتوان تاب آوری استوکستیک شبکه‌ها را مورد بررسی قرار داد. مدلی که مورد استفاده قرار میدهیم passive نام دارد؛ یعنی فرض میشود که لینک معیوب هرگز تعمیر نمیشود و گره برای مدت زمانی در شبکه باقی میماند تا اینکه همه یا درصدی از همسایگان او به وضعیت offline بروند و به محض وقوع چنین پیشامدی گره (کاربر) شبکه را ترک میکند. به همین دلیل فرض میکنیم که کاربرانی که به شبکه وارد شده و گراف را تشکیل داده‌اند، در ابتدا دارای یک طول عمر تصادفی هستند که از یک توزیع احتمال مشخص پیروی میکند. توزیع طول عمر، رفتار کاربر و مدت زمان سرویس (durable time) کل شبکه را نشان میدهد. مدل passive برای سناریوهایی مناسب است که لینکی در شبکه

[21] total variance distance
[22] New worse than used in expectation
[23] New better than used in expectation

حذف یا خراب میشود و زمان تعمیر لینک معیوب یا جستجو برای برقراری لینک جدید، بسیار بیشتر از طول عمر کاربر است. یا اینکه اساساً گره‌های شبکه مجهز به راهبردی جهت ترمیم یا جستجو برای همسایه مناسب نیستند. البته، مدل پیشنهادی میتواند همچنین active باشد؛ بدین معنی که گره‌ها قادر به ترمیم هستند و میتوانند با صرف یک زمان جستجوی تصادفی که از یک توزیع احتمال به دست می آید، به گره‌های مناسب در شبکه متصل گردند.

در این پروژه ما از مدل passive استفاده میکنیم و هنگامیکه درصد مشخصی از همسایگان یک گره شبکه را ترک کنند یا خراب شوند، گره ایزوله خواهد شد.

۵-۱. توصیف مدل

فرض کنید T یک متغیر تصادفی نامنفی و بیانگر طول عمر گره‌ای باشد که میخواهد در شبکه online بماند. این طول عمر میتواند در برخی شبکه‌ها دارای خاصیت دُم کلفت (مانند توزیع پارتو و ویبل) با ویژگی NWUE باشد؛ یا از توزیعهای دُم نازک (مانند توزیع نمایی) پیروی کند. برای مثال، توزیع طول عمر فرآیندها در سیستم عاملهایی مانند یونیکس دارای ویژگی NWUE است. البته در این جا میتوانیم توزیع طول عمر کاربر را نوع عمومی فرض کنیم تا مدل پیشنهادی برای هر دو نوع شبکه مبتنی بر کاربرهای انسانی و غیر-انسانی قابل استفاده باشد.

اگر T متغیر تصادفی نامنفی طول عمر گره با تابع توزیع F و قابلیت اطمینان $\overline{F}$ باشد، تابع m(t) که به صورت $m(t)=\int_t^\infty \overline{F}(x)dx \,/\, \overline{F}(t)$ تعریف میشود، میانگین عمر مانده[۲۴] نام دارد. برطبق تعریف، گوییم T ویژگی NWUE دارد، هرگاه $\forall t\geq 0$ $m(t)\geq E[T]$ باشد. مشابهاً، گوییم T خاصیت NBUE دارد، هرگاه $\forall t\geq 0$ $m(t)\leq E[T]$ باشد. برای مقایسه دو توزیع طول عمر از منظر NBUE-ness، یک ترتیب جزئی[۲۵] وجود دارد و از آنجاکه NWUE و NBUE دوگان هم هستند، ما از مرتبه NWUE استفاده میکنیم و برای مثال اگر بخواهیم دو متغیر تصادفی طول عمر X و Y را از نظر مرتبه با یکدیگر مقایسه کنیم از نماد $X\geq_{NWUE} Y$ استفاده میکنیم.

خواسته ۸: برای هر یک از شبکه‌هایی که در بخش‌های پیشین ساختید، فرض کنید گره‌ها دارای طول عمر T با توزیعهای پارتوی شیفت یافته، ویبل، نمایی و یکنواخت باشند. در اینصورت برنامه‌ای بنویسید و به کمک آزمونهای شبیه‌سازی، احتمال ایزوله شدن شبکه‌ها را برای هریک از این توزیعهای طول عمر در یک نمودار ترسیم کنید. نکته مهم این است که توجه داشته باشید که اگر درجه گره k باشد، در مدل passive گره در صورتی ایزوله فرض میشود که آخرین همسایه او خراب شود. بنابراین، گره زمانی به وضعیت ایزوله وارد میشود اگر و تنها اگر $1\leq r\leq k$ همسایه‌اش در وضعیت خرابی قرار گرفته باشند. برنامه شما بایستی قابلیت تنظیم این پارامتر را داشته باشد. همچنین برای انجام فرآیند شبیه‌سازی، پارامترهای توزیع‌ها را طوری تنظیم کنید که برای مقایسه عادلانه، مقادیر میانگین و نما معمولاً با یکدیگر برابر باشند.

پارتوی شیفت یافته	نمایی	ویبل
$1-(1+t/\beta)^{-\alpha}, t>0, \alpha>1, \beta>0$	$1-e^{-\lambda t}, t\geq 0$	$1-e^{-(t/\beta)^{\alpha}}$

احتمال ایزوله شدن شبکه برای کدامیک از توزیعها بیشتر و برای کدامیک کمتر است؟ مرتبه این توزیعها را با یکدیگر مقایسه کنید. آیا میتوان نتیجه گرفت شبکه‌ای که خاصیت NWUE بیشتری دارد، مستحکمتر است؟ به عبارت دیگر، اگر احتمال ایزوله شدن شبکه را تحت توزیع‌های طول عمر X و Y به ترتیب با $\pi_r(X), \pi_r(Y)$ نمایش دهیم، آیا میتوان نتیجه گرفت که $X\geq_{NWUE} Y \Rightarrow \pi_r(X)\leq \pi_r(Y)$

۵-۲. تشخیص مرتبه NWUE یک توزیع

از آنجا که NWUE-ness طول عمر گره، تاثیر مستقیمی بر تاب آوری شبکه دارد، شاید برایمان جالب باشد تا ببینیم توزیعهای طول عمر مورد نظر از نظر NWUE بودن چه ترتیبی نسبت به هم دارند. در این بخش، روشهای آماری و شبیه‌سازی را برای تشخیص مرتبه NWUE مابین دو توزیع طول عمر براساس نمونه برداری‌های $X_n=(X_1,\ldots,X_n)$ و $Y_m=(Y_1,\ldots,Y_m)$ به ترتیب از جمعیت‌های مستقل و پیوسته X و Y، بیان میکنیم.

۵-۲-۱. روش گرافیکی مبتنی بر نمودار TTT

در آزمونهای طول عمر، کمیتی موسوم به زمان کل آزمون[۲۶] (TTT) وجود دارد که نقشی محوری در تحلیل قابلیت اطمینان ایفا میکند. TTT مقیاسبندی شده[۲۷] (STTT) متغیر تصادفی نامنفی طول عمر T با توزیع F عبارتست از

[24] Mean residue lifetime
[25] partial order
[26] Total time on test
[27] Scaled TTT

$$H(u)=\int_0^{F^{-1}(u)}\overline{F}(t)dt \qquad u\in(0,1)$$

که در آن u یک عدد تصادفی، H(u) تابع توزیع نرخ خطر h(t) و $F^{-1}(u)=\mathrm{Min}\{t,F(t)>u\}$ است. از آنجا که F و H(u) به شکل منحصر به‌فردی یکدیگر را مشخص میکنند، تبدیل TTT میتواند بخوبی خواص سالمندی در شبکه را قضاوت کند. متغیر تصادفی طول عمر NWUE است، اگر و تنها اگر $H(u)\le u$ for all $u\in(0,1)$.

بدین ترتیب تبدیل TTT دو ویژگی مهم دارد. اول اینکه، مشتق H(u) به ازای u=F(t) با وارون تابع نرخ خطر h(t) برابر است. دوم اینکه، H(1) همواره با مقدار میانگین برابر است. معمولاً H(u) به H(1) تقسیم میشود و حاصل این تقسیم با $M(u)=H(u)/H(1)=H(u)/\mu$ نمایش داده شده و به آن تبدیل TTT مقیاسبندی شده (STTT) گفته میشود. توجه داریم که $M(0)=0, M(1)=1$ است. اگر مشتق M(u) را نسبت به u محاسبه کنیم داریم

$$\frac{d}{du}M(u)=\frac{1}{\mu h(F^{-1}(u))}$$

بدین ترتیب، اگر M(u) یک خط مستقیم باشد، توزیع F نمایی خواهد بود. معمولاً در صورتیکه توزیع F موجود نباشد، برای تبدیل TTT بایستی برآورد نمونه ای انجام شود. یعنی از F به طور تصادفی نمونه برداری کنیم تا تابع توزیع تجربی $\hat{F}$ یافت شود. اگر $\hat{F}$ تابع توزیع تجربی متناظر با نمونه تصادفی $T_1, \ldots, T_c$ و $T_{1:c}\le T_{2:c}\le\ldots\le T_{c:c}$ آماره های ترتیبی[28] باشند، آنگاه $\hat{F}(t)=j/c$. چنانچه، $\hat{H}(u)$ را در نقطه u=j/c محاسبه کنیم، به معنای تابعی از زمان کل آزمایش تا لحظه خرابی مولفه j ام در شبکه‌ای با c مولفه خواهد بود. بنابراین داریم

$$\hat{H}(u=j/n)=\frac{1}{n}T(t_j)$$

بدین ترتیب میتوان $\hat{M}(u)$ ، یعنی برآوردی از TTT مقیاسبندی شده را به صورت زیر محاسبه کرد

$$\hat{M}(u=j/c)=\frac{\hat{H}(j/c)}{\hat{H}(1)}=\frac{T(t_j)}{T(t_c)} \qquad j=1,2,\ldots,c$$

وقتی سایز نمونه‌ها زیاد باشد، بر طبق قضیه حد مرکزی، به نسخه جامعه همگرا خواهد شد؛ یعنی $\hat{M}(u)\to M(u)$. در نتیجه، برای بررسی رفتار تابع نرخ خطر، کمیت $\hat{M}(j/c)$ را ملاک قرار داده و نمودار STTT را برحسب زوج نقاط $(j/c,\hat{M}(j/c))$ ترسیم میکنیم.

خواسته ۹: نمودارهای تبدیلات TTT مقیاسبندی شده (نمونه) را برای توزیع‌های طول عمر گره‌ها در شبکه‌های ساخته شده در کنار یکدیگر ترسیم کنید. مقادیر عددی پارامترهای توزیع را طوری تنظیم کنید که میانگین و میانه توزیعها به طور تقریبی با یکدیگر مساوی باشند. در توزیعهای دم کلفت مانند پارتو و ویبل اگر پارامتر شکل را ثابت و پارامتر مقیاس را تغییر دهید چه تاثیری بر روی خاصیت NWUE-ness خواهد داشت؟ برای مثال پارامتر α را یکبار برابر 2 و بار دیگر برابر 1.25 فرض کنید و سایز نمونه ها را برای X برابر ۳۰ و برای Y برابر ۳۲ فرض کنید. با فرض خطای ۵٪، مقدار p-value را برای آزمون فرض مساوی بودن مرتبه NWUE دو توزیع $(H_0{:}X\sim Y)$ در برابر فرض متقابل $(H_A{:}X\ge_{NWUE}Y)$ به دست آورید. به ازای کدام مقدار α توزیعها مرتبه NWUE بیشتری دارند؟ آیا با توجه به مقدار p-value، فرض صفر یعنی برابر بودن مرتبه NWUE قابل رد است؟

یاداشت: **این پروژه به تفصیل برای شما دانشجویان عزیز شرح داده شده تا کامل باشد و تقریباً نیازهای شما را از منظر مطالعه مطالب لازم برآورده سازد. به همین دلیل تعداد صفحات آن قدری بیشتر شده است. امیدمندم که کوشش انجام شده در تعریف این پروژه و انجام آن توسط شما عزیزان منجر به شکوفایی علایق و رشد توامندیهای شما در این عرصه گردد که جزو آرزوهای بزرگ بنده است. با این حال، نکات مهمی را که به ذهن من میرسد، به شکل بندهای مختلف در زیر نوشته‌ام که امیدمندم مورد مطالعه و توجه شما نورچشمان قرار بگیرد.**

۱- قضیه‌های نوشته شده در متن این گزارش تنها برای درک بیشتر است و به همین دلیل بدون اثبات آورده شده است. شما عزیزان لازم نیست آنها را اثبات کنید؛ هرچند که اثبات آنها کار خیلی دشواری نیست. با این حال، دانشجویان عزیز و علاقه مندی که دوست دارند از نمره اضافی و امتیازی برخوردار شوند، میتوانند در گزارش خود اثبات آنها را بیاورند.

۲- تعداد ۹ خواسته این پروژه اگرچه به یکدیگر وابسته‌اند، شما عزیزان میتوانید هریک از خواسته‌ها به طور جداگانه انجام داده و نمره آنرا دریافت کنید. بنابراین، آنها را طوری فهرست کرده‌ام که هر خواسته به‌طور جداگانه قابل نمره‌دهی باشد تا اگر کسی به هر دلیلی مایل به انجام برخی بندها نبود، بتواند نمره سایر بخشها را دریافت کند.

۳- تعداد نفرات اعضای این پروژه حداکثر ۲ نفر است و مهلت ارسال آن تا پایان ترم خواهد بود و تمدید نخواهد شد؛ ضمن اینکه در موعد مشخص (معمولا ۱۰ روز پس از ازمون پایانترم)، از تک تک اعضای پروژه پرسش خواهد شد. بدین ترتیب، لازم است که دانشجویان عزیز و محترم نسبت به چگونگی انجام پروژه خود دانش و آگاهی لازم را داشته باشند و استدعا دارم که از کُپی و رونویسی بدون یادگیری جداً پرهیز نمایند.

[28] Order statistics

۴- نُه خواسته پروژه، جزو خواسته‌های ضروری است. با این حال، دانشجویان عزیز میتوانند بسته به ذوق و سلیقه خود، به بخشهای مختلف پروژه، افزونه‌هایی بیفزایند که مجدداً نمره امتیازی و اضافی به ایشان تعلق خواهد گرفت. برای مثال، ایجاد فرم *GUI* برای ورود مناسب داده‌ها و پارامترهای مساله، امکان نمایش بصری گرافها و شبکه‌ها، امکان تعریف و افزودن سایر گراف‌ها توسط کاربر و ...

۵- لطفاً در همه حال در هر زمانی که پرسش داشتید حتما به بنده مراجعه کرده یا از طریق ایمیل *f_safaei@sbu.ac.ir* بنده را در جریان اشکالات و ابهامات خود قرار دهید. آرزو دارم که انشالله به درس علاقه مند شده باشید و با انجام این پروژه گامی در راستای افزایش و رشد توانمندیهای خود بردارید.

با آرزوی سرفرازی برای همه شما نازنینان

دوستدار همگی شما

صفایی