



تاثیر فرآیند ورود گره‌ها و رشد تصادفی یال‌ها در شبکه‌های مقیاس-آزاد مدل باراباشی-آلبرت

۱- مقدمه مدل گراف‌های باراباشی-آلبرت^۱ (BA) در شبکه‌های مقیاس-آزاد

یکی از مدل‌های شناخته شده در شبکه‌ها، مدل باراباشی-آلبرت (BA) است که رشد و اتصال ترجیحی خطی^۲ (LPA) را به عنوان دو عنصر سازنده اصلی یک شبکه خود-سازمانده^۳ در ساختار مقیاس-آزاد^۴ (SF) پیشنهاد کرده است. این عوامل ناظر به این واقعیت است که در اغلب شبکه‌ها، عامل رشد ناشی از افزودن گره‌های جدیدی است که به شکل LPA به گره‌های موجود در شبکه که تعداد زیادی همسایه دارند، اتصال می‌یابند. چگونگی فرآیند ورود گره‌ها به شبکه‌های واقعی و توزیع احتمال رشد یال‌ها و نیز توزیع احتمال طول عمر گره‌ها در یک شبکه از اهمیت به سزایی برخوردار است و همچنین، یکی از پرسش‌های کلیدی که در این پروژه شبیه‌سازی قرار است به آن پاسخ دهیم این است که آیا مدل BA و اساساً پدیده LPA صرف‌نظر از فرآیند ورود گره‌ها و توزیع احتمال رشد یال‌ها کماکان قادر به تولید گرافی مستقل از مقیاس هست یا خیر؟ صرف‌نظر از مسائلی مختلف در هر شبکه، چگالی احتمال توزیع درجه گره‌ها یک امر مهم ساختاری است که در شبکه‌ها بسیار مورد توجه قرار گرفته است. بسیاری از شبکه‌های دنیای واقعی (البته نه همه) به داشتن توزیع قانون توانی درجه (power-law) شناخته شده‌اند که این امر طبیعت شبکه‌های مقیاس-آزاد (SF) را بیان می‌کند.

یکی از شناخته‌شده‌ترین گرافی که ویژگی مقیاس-آزاد (SF) را از خویش به نمایش گذاشته، توسط باراباشی و آلبرت پیشنهاد شده است که به اختصار BA نام دارد. در این مدل، مفهومی به نام اتصال ترجیحی خطی (LPA) حکم‌فرما است؛ بدین معنی که بسیار محتمل است که گره‌های جدید با گره‌هایی که قبلاً در شبکه استقرار یافته‌اند، ارتباط برقرار سازند. به واژه دیگر، SF مدلی است که در آن، احتمال یافتن گره‌ای با درجه k برحسب قانون-قدرت کاهش پیدا می‌کند. مشخص گردیده که توزیع درجه در این مدل از توزیع قدرت معینی یعنی توزیع پارتو پیروی می‌کند.

مدل BA مدلی از شبکه‌های SF را به نمایش می‌گذارد که با سه گام شرایط اولیه^۵، رشد^۶ و اتصال ترجیحی^۷ ساخته می‌شوند. مدل BA بر دو مکانیزم اصلی اتکاء دارد که بر ویژگی SF شبکه‌های واقعی حکم‌فرما است: (۱) شبکه‌ها به طور پیوسته با افزودن گره‌های جدید توسعه می‌یابند و (۲) گره‌های جدیدالورود به شکل ترجیحی به سایر گره‌هایی که قبلاً در شبکه به خوبی متصل بوده‌اند، اتصال می‌یابند. به این ترتیب شبکه با شروع از یک مجموعه متشکل از m_0 گره ساخته می‌شود پس از آن در هر مرحله، با افزودن گره‌های جدید، شبکه رشد می‌کند. برای هر گره جدید، m لبه جدید ($m \leq m_0$)، بین گره تازه اضافه شده و برخی از گره‌های قبلی شبکه اضافه می‌شود. گره‌هایی که لبه‌های جدید به آن‌ها مرتبط است مطابق یک قانون اتصال ترجیحی خطی انتخاب می‌شوند، یعنی احتمال اینکه گره جدید i به یک گره موجود j متصل شود متناسب با درجه گره j است. مضافاً، گره‌های جدید به یکدیگر اتصال نمی‌یابند تا شاهد وقوع طوقه و لینک‌های چندگانه در گراف تکامل یافته و پدید آمده از مدل پیشنهادی در طی زمان نباشیم. بنابراین گره‌هایی که بیشترین ارتباطات را دارند با احتمال بیشتری به گره‌های جدید متصل می‌شوند. پس به طور خلاصه، این مدل از تعدادی گره اولیه در گراف (m_0) موسوم به هسته اولیه گراف آغاز می‌شود، سپس به شکل یک فرآیند تکراری و با افزودن یک تک گره جدید u با m لبه در هر گام و اتصال آن به گره‌های موجود $v \in V$ و براساس احتمال LPA در معادله زیر، فرآیند رشد در این مدل استمرار می‌یابد

$$\pi_u = k_v / \sum_j k_j \quad (1)$$

بنابراین، اگر m لبه در t تکرار اضافه شده باشد، اندازه گراف $n = m_0 + t$ و کل تعداد لبه‌ها $|E| = mt$ شده و در نتیجه درجه کل گراف $\sum k_i = 2mt$ خواهد شد.

¹ Barabási-Albert

² Linear preferential attachment

³ Self-organizing

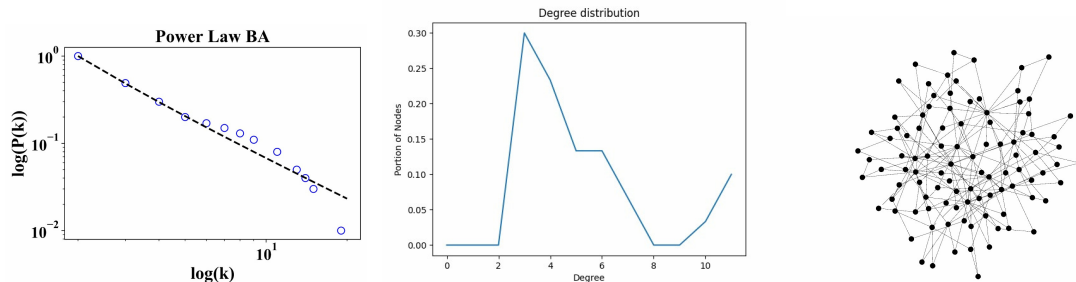
⁴ Scale-free

⁵ Initialization

⁶ Growth

⁷ Preferential Attachment

در شکل زیر به ازای پارامترهای $m_0=3, m=2, n=100$ خروجی شبیه‌سازی مدل BA همراه با توزیع درجه در دو حالت معمولی و در مقیاس $\log\text{-}\log$ نشان داده شده است.



۲- روش ساخت مدل پیشنهادی

در طی سال‌ها، مکانیزم مزیت تجمعی که توسط باراباشی و آلبرت پیشنهاد شده به شکل گسترده‌ای به عنوان یک توضیح قابل قبول برای توزیع درجه قانون-قدرت که در شبکه‌هایی همچون WWW و اینترنت مشاهده گردیده پذیرفته شده است. همان‌طور که گفته شد، هدف ما در این پروژه شبیه‌سازی، توسعه مدل BA است که جزئیات ساخت آن به تفصیل مطرح خواهد شد. برای توصیف رقابت بین گره‌ها ما در مدل BA تغییراتی داده‌ایم یعنی برخلاف BA اولیه، ورود گره‌ها در مدل ما از یک فرآیند تصادفی برای مثال پواسن پیروی میکنند.

فرآیند ورود گره‌ها می‌تواند از توزیع‌های مختلف احتمال پیروی کند که جزئیات برخی از مهمترین آن‌ها همراه با مقادیر پارامترهایشان در زیر آورده شده است. فرآیند ورود گره‌ها می‌تواند براساس توزیع‌های احتمال زیرین باشد:

- **توزیع پواسن با نرخ 2.5**
- توزیع تیهگن^۹ (برای ورود گره برابر با ۱ و برای رشد یال برابر با ۲)؛ بد نیست اشاره کنیم که اگر فرآیند ورود گره از توزیع تیهگن و برابر با ۱ پیروی کند، به همان مدل BA کلاسیک دست خواهیم یافت.
- توزیع نمایی با نرخ ورود 0.7
- توزیع لوی^{۱۰} با شاخص پایداری^{۱۲} $\alpha=0.5$ ، پارامتر چولگی $\beta=1$ ، پارامتر مقیاس $\gamma=1$ و پارامتر مکان $\delta=0$
- توزیع کوشی^{۱۱} با شاخص پایداری $\alpha=1$ ، پارامتر چولگی $\beta=0$ ، پارامتر مقیاس $\gamma=1$ و پارامتر مکان $\delta=2.5$
- توزیع نرمال با شاخص پایداری $\alpha=2$ ، پارامتر چولگی $\beta=0$ ، پارامتر مقیاس $\gamma=4$ و پارامتر مکان $\delta=2.5$
- توزیع پارتو با پارامتر مقیاس ۱ و پارامتر چولگی ۱
- توزیع لجستیک با میانه 2.5 و پارامتر مقیاس 2؛ لازم به ذکر است که با توجه به ذات رشد جمعیت گره‌ها در شبکه، از توزیع لجستیک تنها برای رشد گره و نه رشد یال باید استفاده کرد و برای توزیع احتمال رشد یال نیز از توزیع دوگمله‌ای با پارامترهای $p=0.1$ و n (اندازه گراف) استفاده نمود.

تذکر ۱: در این پروژه تنها فرآیند پواسن برای ورود گره‌ها الزامی است و مابقی توزیع‌ها در صورت تمایل به پیاده‌سازی نمره امتیازی خواهند داشت.

یادداشت ۱: توزیع‌های کوشی، لوی و نرمال به خانواده‌ای از توزیع‌ها موسوم به توزیع‌های پایدار^{۱۲} تعلق دارند که این توزیع‌ها به دلیل چولگی و داشتن جرم احتمال زیاد در دُم‌های تابع چگالی، مدل‌های مناسبی برای بسیاری از پدیده‌ها در مهندسی، اقتصاد، فیزیک و هیدرولوژی محسوب می‌شوند. از آن‌جا که به غیر از توزیع کوشی، فرم تحلیلی برای معکوس CDF متغیر تصادفی توزیع پایدار وجود ندارد، نمیتوان به مانند سایر توزیع‌ها از روش تبدیل معکوس برای تولید اعداد تصادفی این دسته از توزیع‌ها استفاده کرد. براین اساس، به منظور تولید اعداد تصادفی از تابع چگالی متغیرهای تصادفی پایدار باید استفاده کرد.

یادداشت ۲: در تنظیم پارامترهای توزیع‌های مختلف احتمال در فرآیند ورود گره‌ها باید بازه عددی خروجی این توابع مشابه با یکدیگر باشد و نیز میانه توزیع‌های احتمال با یکدیگر برابر باشد (برای مثال در شبیه‌سازی مقدار 2.5 برای میانه فرض شود) تا بتوان به مقایسه‌ای عادلانه در میان این توابع توزیع احتمال دست یافت. مضافاً، محدودیت دیگری نیز اعمال شده و آن این است که حد رشد شبکه یا تعداد تجمعی گره‌ها را متناهی فرض کرده‌ایم طوری‌که در هر گام از شبیه‌سازی حداکثر ۱۰ گره بتوانند به طور هم‌زمان وارد شبکه گردند و بدین سیاق در روند شبیه‌سازی، گراف را تا حداکثر ۲۰۰ گره رشد داد. البته

⁸ Degenerate

⁹ Levy

¹⁰ Index of stability

¹¹ Cauchy

¹² Stable distribution

واضح و منطقی هم به نظر می‌رسد که در هر گام از مراحل ساخت، به بیش از ۱۰ گره اجازه ورود به سیستم را ندهیم چراکه اگر ۲۰۰ گره به‌طور همزمان و در یک گام وارد سیستم شوند دیگر قادر به بررسی و تحلیل فرآیند رشد شبکه نخواهیم بود.

۳- نحوه پیادسازی مدل اصلاح شده باراباشی-آلبرت (Modified BA)

الف. رشد تصادفی

رشد تصادفی بدین معنی است که با تعداد اندکی گره ($m_0=3$) کار را آغاز کنیم و در لحظه زمانی t ، تعدادی گره جدیدی را که از توزیع احتمال مشخصی اخذ میشوند به سیستم وارد کنیم. بایستی توجه داشت که هر گره به شکل مستقلی از سایر گره‌ها به m گره متمایز (که پارامتر m خود نیز از یک توزیع احتمال پیروی می‌کند) موجود در سیستم متصل میشود

ب. اتصال ترجیحی

در گزینش گره‌هایی که قرار است گره یا گره‌های جدید بدان‌ها متصل گردد، شیوه عملکرد ما متکی بر همان اسلوب LPA است که در مدل سنتی BA وجود دارد. هدف از این کار، بررسی فرآیند ورود گره‌ها به شیوه‌ای مستقل از نحوه انتخاب آن‌ها است. بدین ترتیب فرض میکنیم که با احتمال π_u گره جدیدالورود u به گره‌ای مانند v بسته به درجه آن، $k_v(t)$ ، اتصال یافته است. یعنی، وقتی گره‌هایی را برای اتصال گره یا گره‌های جدید به آنها در شبکه گزینش میکنیم فرض بر این است که احتمال اتصال گره یا گره‌های جدید به گره انتخابی، به درجه این گره انتخاب شده وابسته بوده و از سایر بررسی موارد پرهیز شده است. همچنین، فرض کرده‌ایم که تعداد یال‌های گره یا گره‌های جدیدالورود می‌تواند از یک توزیع احتمال پیروی کند. لازم به ذکر است که اگر تعداد یال‌ها براساس توزیع احتمال تبهگن اختیار گردد، مجدداً به همان مدل پایه BA دست خواهیم یافت.

<p>الگوریتم: مدل اصلاح شده باراباشی-آلبرت (BA) مبتنی بر فرآیند تصادفی ورود گره‌ها و توزیع احتمال رشد یال‌ها</p> <p>ورودی: هسته اولیه گراف ($m_0=3$)، تعداد گام اجرایی الگوریتم (t_{max})، توزیع احتمال ورود گره، توزیع احتمال رشد یال، حداکثر تعداد گره ورودی در هر گام (n_{max})</p> <p>خروجی: گراف تکامل یافته پس از t_{max} گام زمانی</p> <p>از $t=1$ تا t_{max} انجام بده</p> <p>(الف) رشد تصادفی</p> <p>الف-۱. با استفاده از توزیع احتمال مفروض مربوط به فرآیند ورود گره، عددی را به شکل تصادفی یکنواخت اختیار می‌کنیم</p> <p>الف-۱-۱. اگر این عدد تصادفی بزرگتر از n_{max} بود از آن صرف‌نظر کرده و به الف-۱ باز می‌گردیم</p> <p>الف-۱-۲. عدد تصادفی حاصل را به یک عدد صحیح تبدیل می‌کنیم که این عدد بر تعداد گره‌های جدیدالورود دلالت دارد</p> <p>(ب) اتصال ترجیحی</p> <p>ب-۱. به ازای هر گره جدید، یک عدد به شکل تصادفی یکنواخت از توزیع احتمال مفروض رشد یال‌ها اختیار میشود</p> <p>ب-۱-۱. چنانچه این عدد تصادفی از تعداد گره‌های فعلی موجود در گراف بیشتر و یا عددی منفی بود به گام ب-۱ باز می‌گردیم</p> <p>ب-۲. عدد تصادفی مزبور را به یک عدد صحیح تبدیل می‌کنیم که این عدد نشان‌دهنده تعداد یال‌هایی است که گره جدیدالورود همراه با خود دارد. البته این عدد می‌تواند برای هر یک از گره‌ها مقداری متفاوت و متمایز داشته باشد</p> <p>ب-۳. هر گره با تعداد یال محاسبه شده را به ترتیب زیر به گراف می‌افزاییم</p> <p>ب-۳-۱. به ازای هر یال از گره جدیدالورود، از بین گره‌های موجود در گراف (به استثنای آن‌هایی که در همین گام اضافه شده‌اند)، گره‌هایی متمایز را با احتمالی متناسب با درجه آن‌ها انتخاب می‌کنیم (اتصال ترجیحی خطی)</p> <p>ب-۳-۲. یک یال جدید را در گراف بین گره انتخابی و گره جدید برقرار می‌سازیم</p>	
--	--

۴- نتایج تجربی آزمون شبیه‌سازی و خواسته‌های پروژه

از شما دانشجویان عزیز انتظار میرود که برطبق الگوریتم بالا مدل BA را طوری اصلاح کنید که فرآیند ورود گره‌ها پویا باشد و نیز مدل رشد یال‌های گره‌های وارد شونده از توزیع نمایی پیروی کند. لازم است تا نتایج حاصل از شبیه‌سازی‌های انجام شده را بر روی مدل پیشنهادی اصلاح شده BA ارائه داده و سپس شکل نهایی شبکه تکامل یافته و توزیع درجه آن را تحلیل و با BA کلاسیک مشابه مقایسه کنید. هدف از این پروژه آن است که تاثیر فرآیند تصادفی ورود گره‌ها و نیز توزیع احتمال رشد یال‌ها را در گراف تولیدی مدل اصلاح شده BA بر رفتار و مشخصات شبکه نهایی مشخص سازیم. لازم به ذکر است که نتایج شبیه‌سازی را باید از متوسط‌گیری بر روی حداقل ۱۰۰ تکرار آزمون شبیه‌سازی مستقل استخراج و فاصله اطمینان شبیه‌سازی را گزارش بفرمایید.

بدین منظور و برای آسانی کار شما، دو برنامه تولید مدل BA که با پایتون کدنویسی شده و تکامل شبکه را در طی گام‌های مختلف به نمایش می‌گذارند به شما داده شده است. این دو برنامه میتوانند در خواش جزئیات پیاده‌سازی مدل BA اصلی به شما کمک کنند و نیز در پایان توزیع درجه شبکه را به شما نمایش دهند.