

شبکه پیچیده پویا

دکتر علی اکبری

تمرین 3 (شناسایی و تحلیل معیارهای ۳ شبکه)

سجاد آقانصیری – 402443009



دانشجو کارشناسی ارشد دانشکده مهندسی و علوم کامپیوتر

دانشگاه شهید بهشتی

بهار 1403

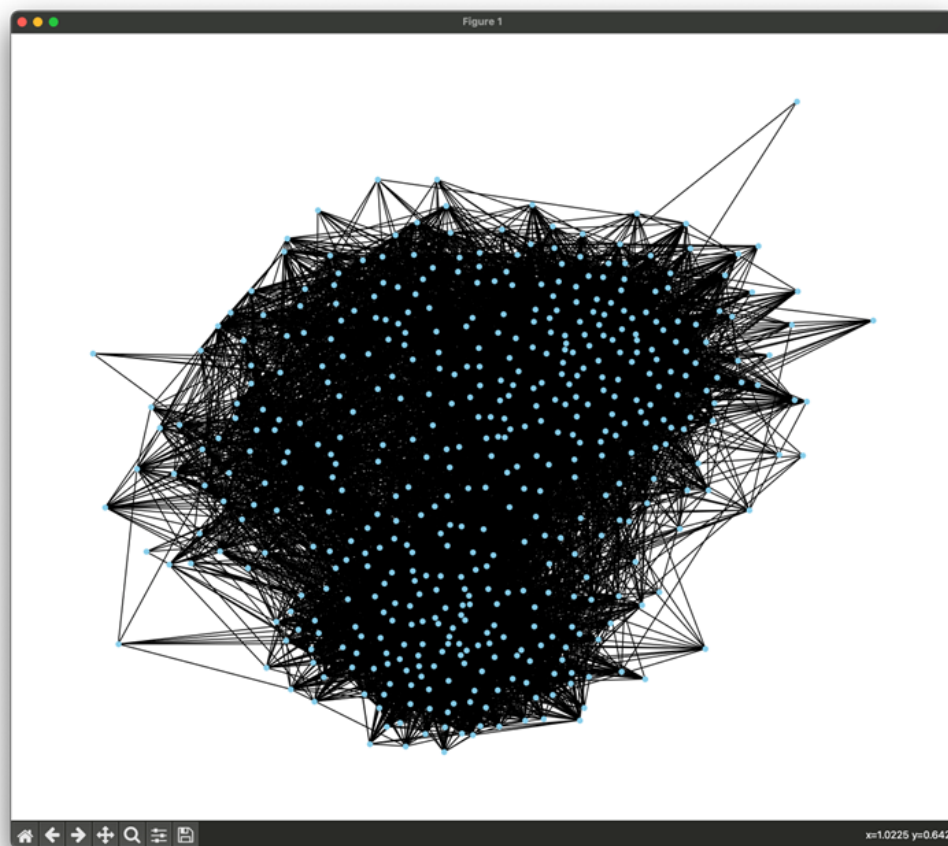
ما در این تمرین قصد داریم به عنوان شخصی که معیار های شبکه پیچیده را به بتوانیم این معیار ها را برای 3 دیتا [1],[2] تازگی یاد گرفته به کمک ابزار هایی مانند گرفته شده که فایل های دریافتی به [4],[5],[6] ست واقعی که از وبسایت های را نشان edge بود . به صورت نرمالیز 3 ستون دارد 2 سطر اول csv صورت میدهد و معمولا ستون سوم وزن آن یال را نشان میدهد .

## Dataset 1(twitter)

این دیتاست از سایت [5] گرفته شده که ستون اول و دوم یال ها را و ستون سوم وزن آن را نشان میدهد .

```
{weight': 0.002105263157894737'} 4 0
```

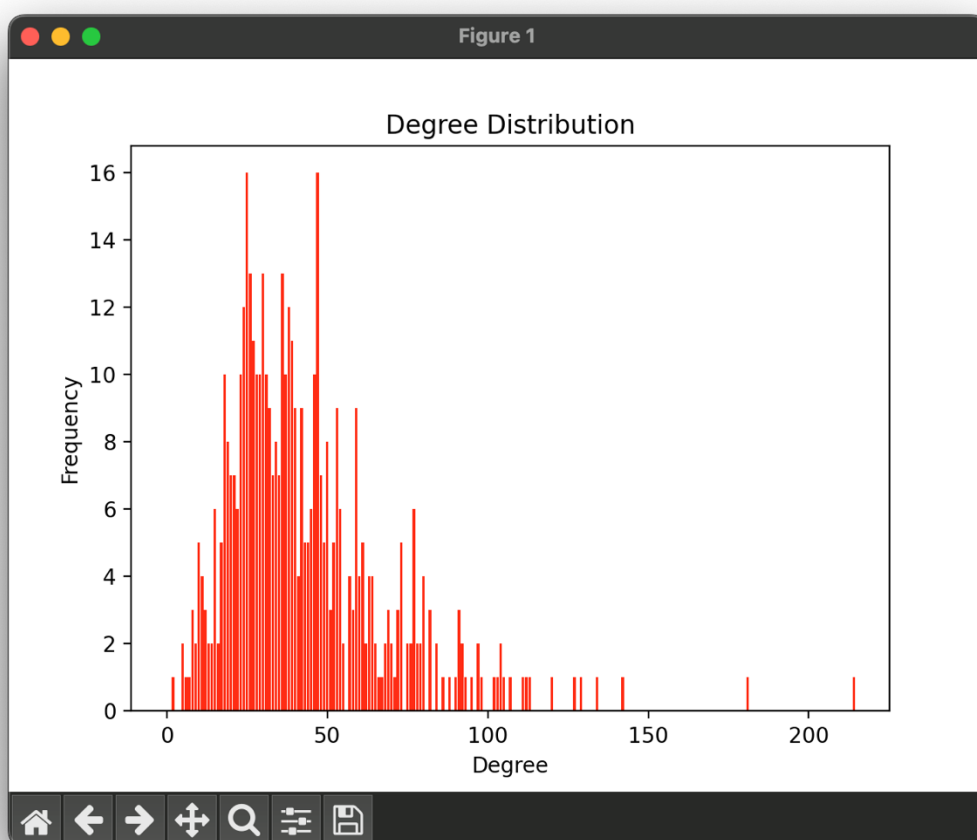
```
{weight': 0.002105263157894737'} 4 7
```



نگاه شهودی:

خب در نگاه اول بنظر میاد با یک small world network سر کار داریم شاید اونقدر چگال نیست و حدود 3 انجمن میتوان در آن دید(البته انجمن ها بسیار بزرگ هستند) .

قطر شبکه بنظر میرسد مانند تمام social network ها قطر بسیار بزرگی نداشته باشد و زیر عدد ۶ باشد. حال با نگاهی به نمودار (fig1) degree distribution متوجه میشویم [3] [1], که توزیع آن کمی شبیه به لاگ نرمال است که برای این شبکه منطقی است و هرچه از بازه حدود بین ۲۵ و ۵۰ اگر دور شویم frequency کاهش پیدا میکند.



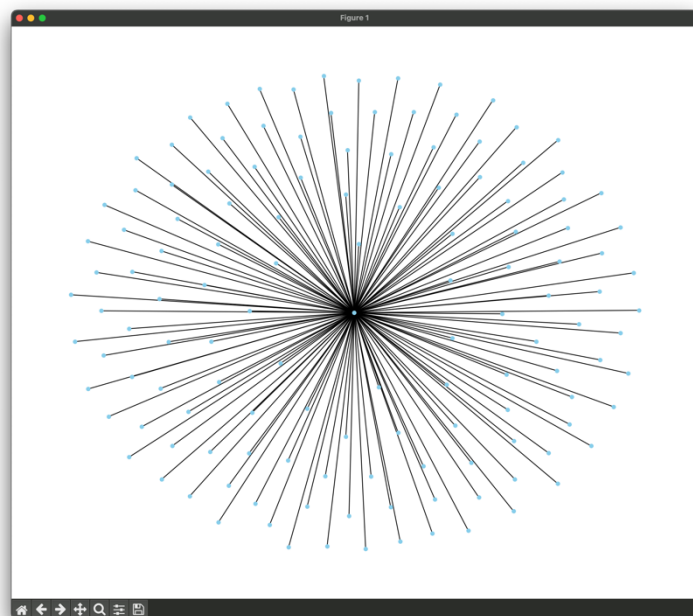
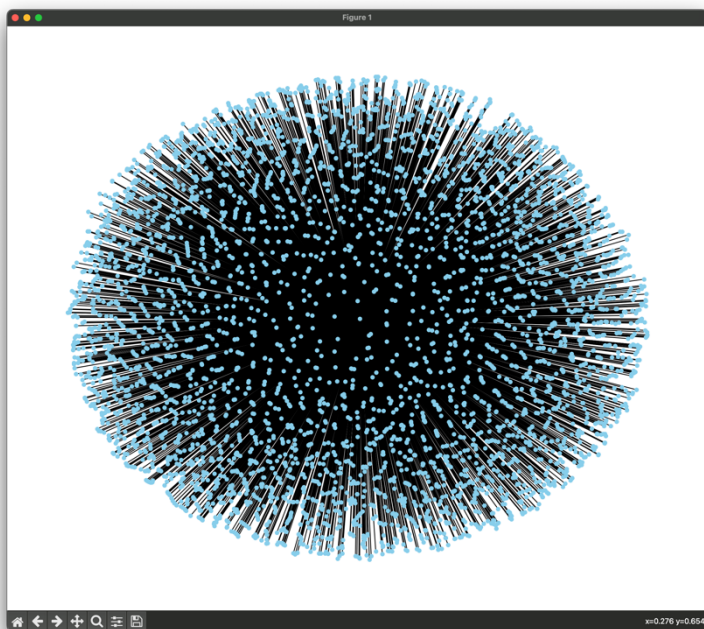
خب یه نگاهی به معیارهای بدست آمده کنیم:

475 نود و 10222 یال داریم. میانگین درجه شبکه 43.04 میباشد. همانطور که میبینیم خیلی شبکه چگال نیست خب بدیهی هم است چون این که یک نفر با همه interaction داشته باشد کم است پس گراف ما خیلی چگال نیست. قطر شبکه که هم اندازه بیشترین فاصله ۲ نود در گراف است که در اینجا ۴ است و به این معناست که شبکه ما small world است پس دارای ارتباطات خوب و مستحکم است. ولی توجه داشته باشیم که شبکه چون average clustering coefficient نسبتا بالایی دارد احتمالا انجمن های کوچک با چگالی بسیار زیاد باید بتوان در آن پیدا کرد و اینجاست که ثابت میشود که transitivity بالایی دارد که ۰.۳ است. در شبکه های پیچیده، «همگنی» یا «همبستگی درجه» به گرایش گره ها برای اتصال به دیگر گره هایی که خصوصیات مشابهی دارند، اشاره دارد. به طور خاص، همگنی درجه در مورد این است که آیا گره های با درجه بالا (آنهايي که اتصالات زیادی دارند) ترجیح می دهند به سایر گره های با درجه بالا متصل شوند، و به همین ترتیب، گره های با درجه پایین به سایر گره های با درجه پایین [chatgpt].

و همانطور که میبینیم عدد بسیار بالایی نیست پس لزوماً نودهای rich به یکدیگر وصل نیستند پس در مدل کردن این نتورک نباید از راه‌هایی که به rich get richer می‌رسد استفاده کرد.

PageRank یک الگوریتم است که توسط موتور جستجوی گوگل برای رتبه‌بندی صفحات وب در نتایج جستجوی آن استفاده می‌شود. این الگوریتم به نام لری پیج، یکی از بنیان‌گذاران گوگل، نامگذاری شده است. الگوریتم PageRank وزن عددی را به هر عنصری از مجموعه‌ای از اسناد مرتبط با پیوند، مانند شبکه جهانی وب، اختصاص می‌دهد، با هدف "اندازه‌گیری" اهمیت نسبی آن در مجموعه. الگوریتم بر اساس یک اصل بازگشتی پیچیده کار می‌کند و از ساختار پیوند به عنوان شاخصی برای ارزش صفحه فردی استفاده می‌کند. به طور خلاصه، PageRank نتیجه "رای‌گیری" بین تمام صفحات دیگر وب است، که صفحاتی که لینک‌ها را از سایت‌های با اهمیت بالا دریافت می‌کنند، رتبه بالاتری کسب می‌کنند [wikipedia].

Dataset 2(cisco) بود که البته کمی عجیب بود 😊



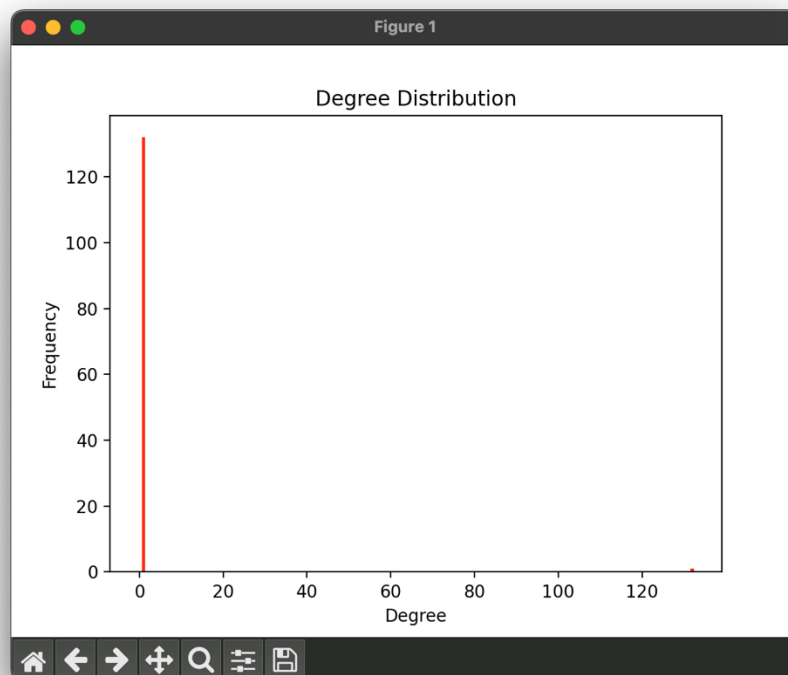
این دیتا ست عجیب ورودی های عجیبی هم دارد که با مطالعه بهتر سایت میتوان درک بهتری از آن داشت. [6] این دیتا ست ستون اول میگوید کدام گراف ۲ ستون بعدی نشان دهنده درخواست دهنده و هاب است و درواقع برای درک بهتر آن باید دیدگاهمان این باشد که شبکه به شورت جفت گره های بسیار زیاد است که هیچیک به دیگری وصل نیست و ستون آخر پروتکل است.

g21 20 23 1p6-30,3p17-8

g21 10 29 1p6-32

g21 43 31 1p6-16

همانند شبکه اول ابتدا به سوراغ جدول degree distribution میرویم



که خب واضحاً چون یک نود هاب است و بقیه به او وصل هستند تعداد نود زیادی با درجه یک داریم و یک نود که به بقیه وصل است و درجه  $n-1$  را دارد. در ادامه مطالعه معیارها را میبینیم .

قطر ۲ ، چگالی کم ، و دوست نداشتن نود های شبکه برای افزایش تعداد و ایجاد انجمن مشخص است. در این شبکه یک نود بسیار مهم داریم که همان نود هاب است که به تمام نود های دیگر وصل است پس باید در نظر گرفت با آسیب رساندن به این نود میتوان کل شبکه را نابود کرد پس باید از لحاظ امنیتی تامین شود.

### Dataset 3(US Airlines)

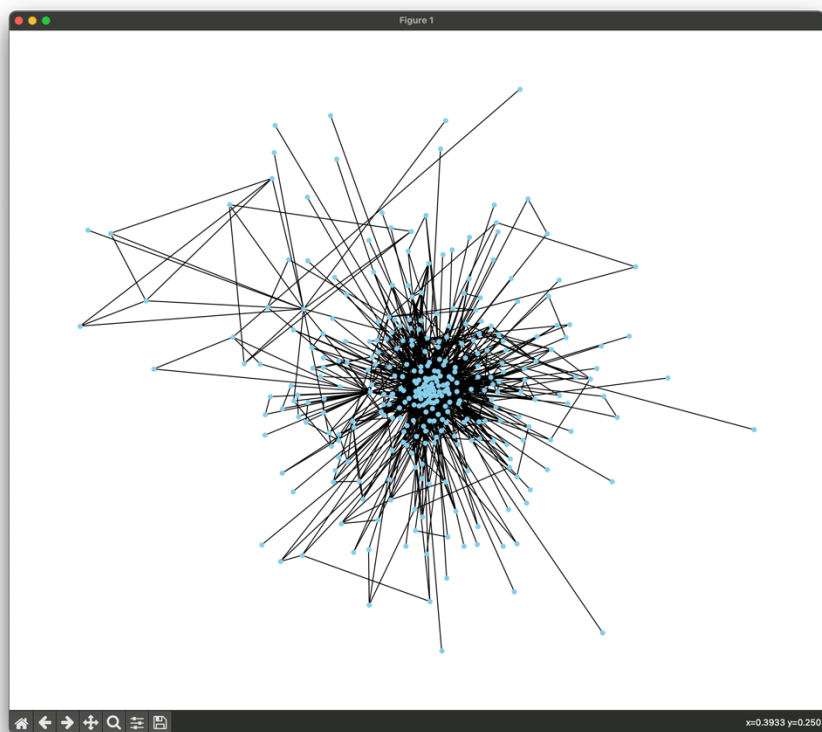
این شبکه همانطور که از اسمش واضح است خط های هوایی آمریکا را نشان میدهد و وزن میان این ارتباطات نیز مشخص و در دیتا موجود است. [4]  
نمونه دیتا :

8 1 .1026

4 2 .0515

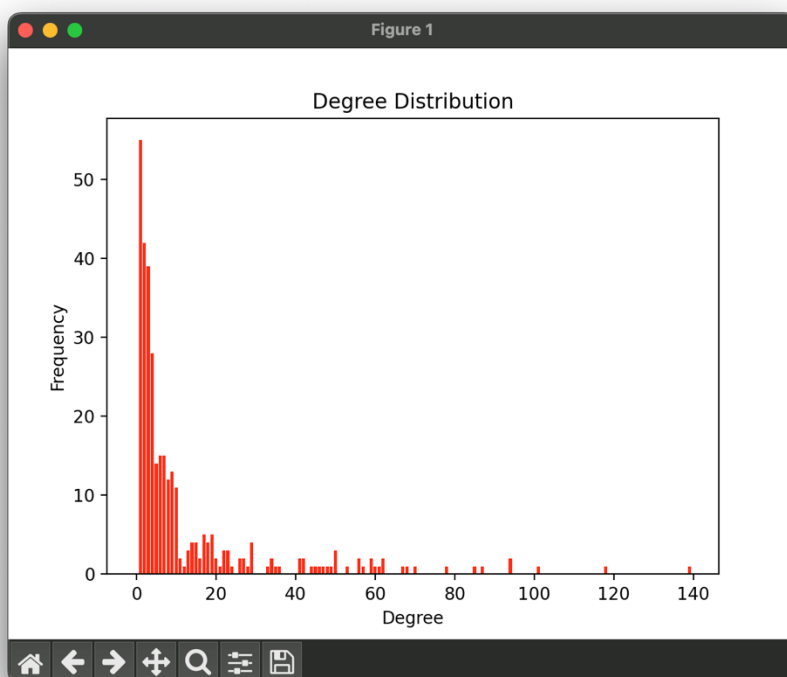
8 2 .0866

که به صورت نرمال قابل خواندن است حال به خود نتورک نگاهی بندازیم





همانطور که میبینیم یک هسته مرکزی داریم که خوشه بسیار چگال و پرتراکمی است و در حاشیه تعدادی نود های پرت که خیلی درجه بالایی ندارند. که انتظار میرود توضیح نرمالی داشته باشد.



تعداد درجه ۱ که گپ آن است و هرچه دور تر میشویم نود هایی با درجه زیاد کمتر پیدا میشوند.

```
[Running] python3 -u "/Users/sj_ag77/Desktop/Complex network Dr Aliakbari/assignment 3_sajjad aghanasiri/USAAirlines/main.py"
Number of nodes: 332
Number of edges: 2126
Metrics:

```

		Metric	Value
0		Average Degree	12.807229
1		Density	0.038693
2		Diameter	6.000000
3	Average Clustering Coefficient		0.625217
4	Transitivity		0.396392
5	Average Shortest Path Length		2.738125
6	Assortativity (Degree Correlation)		-0.207876

```

Top 5 Nodes based on Centrality Measures:
Degree Centrality: [(118, 0.4199395770392749), (261, 0.3564954682779456), (255, 0.30513595166163143), (152, 0.283987915407855), (182, 0.283987915407855)]
Betweenness Centrality: [(118, 0.20830586985886954), (8, 0.17006973867446296), (261, 0.1531955411430337), (201, 0.09422863574735517), (47, 0.09281838103560969)]
Closeness Centrality: [(118, 0.6073394495412844), (261, 0.5544388609715243), (67, 0.5399673735725938), (255, 0.5355987055016181), (201, 0.533011272141707)]
PageRank Centrality: [(118, 0.036083465742291592), (261, 0.03173406805278732), (201, 0.02308216381149434), (47, 0.02263384789929445), (67, 0.02231674679500937)]
Degree Centralization: 0.0011552835235794508
Closeness Centralization: 0.28926242318012463
Betweenness Centralization: 0.0012342559736294588
Pagerank Centralization: 0.03313769945863244
2024-04-24 13:45:31.404 Python[62378:3379780] WARNING: Secure coding is not enabled for restorable state! Enable secure coding by implementing NSApplicationDelegate.
applicationSupportsSecureRestorableState: and returning YES.

```

قطر ۶ چگالی کم، کلاسترینگ بالا (که میتواند عجیب باشد چرا باید این معیار زیاد باشد چون اگر از  $a$  به  $b$  برویم و از  $b$  به  $c$  برویم و مسیری بین  $a$  و  $c$  وجود داشته باشد چرا اصلا باید از  $b$  رد بشویم!! ☹)

با بررسی assortativity که مقدارش کم است متوجه میشویم انجمن های کوچک با چگالی زیاد در این شبکه وجود ندارد که برای راه های هوایی منطقی است.

قطر ۶ هم میتواند مشکوک باشد چون درست است که شبکه small world. باشد ولی بدان معناست که وجود دارد  $a$  که فاصله اش با  $b$  ۶ باشد که بنظر منطقیش میتوانست ۲ و یا حداکثر ۳ باشد و ۶ عجیب است.

و betweenness کم نشان میدهد که کنترل ترافیکی شبکه مستحکم است چون اگر نودی پیدا شود که betweenness بالایی داشته باشد باید نود شلوغی باشد و دارای ارزش بالاتری نیز است و خب برای راه های هوایی منطقی نیست و باید ترافیک پخش شود.

## نتیجه گیری:

در کل با داشتن و بررسی معیار های یک نتورک میتوان یک دید نسبتا خوب به شبکه داشت و یا چند شبکه را باهم مقایسه کرد که این کار فقط با داشتن عکس شبکه امکان پذیر نیست و خب این معیار ها مارا به فکر فرو میبرند که آیا کم بودن بیش از اندازه یک معیار یا بلعکس چه دلیلی می تواند داشته باشد و یا چگونه آن را بهبود بخشید و یا اثر دهی آن را حتی بالاتر برد.

## References

- [1] "Visual studio code"
- [2] "NetworkX Library"
- [3] "Math Plot Library"
- [4] "NetworkRepository.com"
- [5] "Snap.stanford.edu/data/congress-twitter"
- [6] "Snap.stanford.edu/data/cisco-networks"