Universidad Nacional Autónoma de México

Facultad de Ciencias

GRÁFICAS HÍBRIDAS LIBRES DE ESCALA: EFECTO DE MUNDO PEQUEÑO

T E S I S

QUE PARA OBTENER EL TÍTULO DE:
MATEMÁTICO

PRESENTA:
ANAHY SANTIAGO ARGUELLO

DIRECTOR DE TESIS:
PEDRO EDUARDO MIRAMONTES VIDAL

2012

1.Datos del alumno	
Apellido paterno	Santiago
Apellido materno	Arguello
Nombre	Anahy
Teléfono	24 57 85 22
Universidad	Universidad Nacional Autónoma de México
Facultad o escuela	Facultad de Ciencias
Carrera	Matemáticas
Número de cuenta	407004506
2. Datos del tutor	
Grado	Dr.
Nombre(s)	Pedro Eduardo
Apellido paterno	Miramontes
Apellido materno	Vidal
3. Datos del sinodal 1	
Grado	Dr.
Nombre(s)	Denis Pierre
Apellido paterno	Boyer
4. Datos del sinodal 2	
Grado	Dr.
Nombre(s)	Juan José
Apellido paterno	Montellano
Apellido materno	Ballesteros
5. Datos del sinodal 3	
Grado	Dr.
Nombre(s)	Manuel Jesús
Apellido paterno	Falconi
Apellido materno	Magaña
6. Datos del sinodal 4	
Grado	Mat.
Nombre(s)	Carlos Alberto
Apellido paterno	Serrato
Apellido materno	Hernández
7.Datos del trabajo escrito.	3
Título	Gráficas híbridas libres de escala: efecto de mundo pequeño
Número de páginas	63 p.
Año	2012

Índice general

Introducción

En 1967, el psicólogo Stanley Milgram [6] hizo una serie de experimentos que lo condujeron a concluir que cualesquiera dos extraños están conectados por una cadena de personas de longitud de a lo más seis. En 1999, Barabási observó que en ciertas partes del internet, dos páginas web cualesquiera están conectadas por a lo más 19 clicks una de otra. Hay muchos ejemplos más en los que se ha observado que muchas redes o gráficas, que modelan procesos reales, poseen el llamado *efecto de mundo pequeño*, con dos tratamientos propios de él, la distancia promedio pequeña para cualesquiera dos vértices de la gráfica y el efecto de agrupamiento en el cual cualquier par de nodos es más probable que sean adyacentes si comparten un vecino.

En 1999, muchos grupos de investigación concluyeron independientemente que muchas redes, como el internet, redes sociales, etc.[7,8,9,10,11,12,13,14,15, 16,17,18,19 y 20] poseen la propiedad de ser libre de escala. Una gráfica libre de escala es aquella en que la probabilidad de que un vértice tenga grado k es proporcional a $k^{-\beta}$ para alguna constante positiva $2 < \beta$. Usando el modelo de gráfica aleatoria, combinada con la propiedad de ser libre de escala con $2 < \beta < 3$, demostraremos que casi seguramente la distancia promedio es del orden de $loglog(n)$ y tiene diámetro del orden de $log(n)$.

También se han dado intentos de modelar el efecto de agrupamiento, la manera más común es agregar aristas aleatoriamente a una cuadrícula o algo similar. Algunos de estos modelos se pueden encontrar en [21,22,23 y 24].

Muchos modelos de gráficas propuestos atrapan uno solo de los aspectos del efecto de mundo pequeño; al parecer hay dificultades para relacionarlos.

En este trabajo se busca definir dos modelos de gráficas para luego "juntarlos" en una sola, de tal forma que se tenga los dos aspectos del efecto de mundo pequeño en ella, es por eso el uso del término *gráfica híbrida*. Además la exposición de tales conceptos se hacen de una manera totalmente accesible a la mayoría de los estudiantes de la carrera de matemáticas ya que no se parte de ningún concepto anterior; todo lo necesario se define y hay realmente pocos teoremas usados que no son demostrados.

El trabajo se organiza como sigue: En el Capítulo 1 se define el concepto de gráfica aleatoria con una sucesión de grados esperados dados; también se demuestra en tal capítulo lemas y teoremas que servirán como base para demostrar los teoremas centrales, que serán de gran utilidad para desarrollar el concepto de efecto de mundo pequeño en las gráficas híbridas. En el Capítulo 2 demuestro los teoremas centrales que acabo de mencionar. En el Capítulo 3 defino el otro modelo de gráfica que necesito; este modelo será definido como gráfica local, este capítulo es muy breve ya que el modelo considerado es relativamente simple. En el Capítulo 4 considero un modelo híbrido que combinará el modelo expuesto en el Capítulo 1 con el concepto de gráfica local, desarrollado en el Capítulo 3. En la primera sección del mismo se muestra que la gráfica local es fuerte, es decir, si tengo una gráfica híbrida puedo extraer, mediante un algoritmo, la gráfica local con un mínimo de error. La segunda sección tiene como propósito mostrar que la gráfica híbrida combina los dos aspectos del fenómeno de mundo pequeño, la distancia promedio pequeña y el efecto de agrupamiento.

Capítulo 1

Gráficas aleatorias con grados esperados dados

1.1. Conceptos básicos

Comenzaremos esta sección definiendo conceptos básicos de teoría de gráficas.

Definición 1.1. *Una gráfica G es un arreglo de tres conjuntos ordenados $\{V(G), E(G), \psi_G\}$, donde V(G) es un conjunto de vértices, E(G) es un conjunto, ajeno a V(G), de aristas y una función de incidencia ψ_G que asocia a cada arista con un par de vértices (no necesariamente distintos) de G. Si e es una arista y u y v son vértices tales que $\psi_G(e) = uv$, entonces se dice que e une a u y v; los vértices u y v se llaman extremos de e.*

Definición 1.2. *El grado, deg(v), de un vértice v en G es el número de aristas de G que inciden con v, cada lazo (arista que comienza y termina en el mismo vértice) cuenta como dos aristas incidentes.*

Definición 1.3. *Una gráfica G es conexa si dados calesquiera dos vértices u y v en G, existe un camino de v a u.*

Ahora definiremos algunos conceptos matemáticos que nos serán de gran utilidad en el resto de este trabajo.

Definición 1.4. *Se dice que $f : \mathbb{N} \longrightarrow \mathbb{R}$ es $f(n) = O(g(n))$ si $\exists k > 0$ y $\exists n_0$ tal que $\forall n > n_0$ $|f(n)| \leq k|g(n)|$*

Definición 1.5. *Se dice que* $f : \mathbb{N} \longrightarrow \mathbb{R}$ *es* $f(n) = \Theta(g(n))$ *si* $\exists k_1 > 0$, $\exists k_2 > 0$ *y* $\exists n_0$ *tal que* $\forall n > n_0$ $k_1|g(n)| \leq |f(n)| \leq k_2|g(n)|$

Definición 1.6. *Se dice que* $f : \mathbb{N} \longrightarrow \mathbb{R}$ *es* $f(n) = o(g(n))$ *si* $\forall \epsilon > 0$ $\exists n_0$ *tal que si* $n > n_0$ *se cumple que* $|f(n)| \leq \epsilon |g(n)|$

La siguiente definición es clave en todo el trabajo, ya que con el siguiente tipo de gráficas desarrollaremos los resultados de esta tesis.

Definición 1.7. *Consideramos una clase de gráficas aleatorias; a cada una de ellas se le asigna una sucesión de grados esperados(es la esperanza de la variable aleatoria grado)* $\mathbf{w}=(\omega_1, \omega_2, ..., \omega_n)$. *La probabilidad* p_{ij} *de que exista un arista entre un vértice* v_i *y un vértice* v_j *es* $\omega_i \omega_j \rho$ *para cualesquiera* i *y* j. *Se tiene que* ρ *es* $(\sum \omega_i)^{-1}$ *y supondremos que el* $\max_i\{\omega_i^2\} \leq \sum_k \omega_k$ *esto asegura que* $p_{ij} \leq 1$ *para toda* i *y* j.

Remarcamos el hecho de que $\max_i\{\omega_i^2\} \neq \sum_k \omega_k$ implica que la sucesión es graficable (en el sentido que satisface las condiciones necesarias y suficientes para que una sucesión represente una gráfica [1]) excepto que no necesitamos que los ω_i sean enteros. Hay que notar que este modelo permite una probabilidad diferente de cero para los lazos. El número de lazos es muy pequeño (de orden pequeño) comparado con el número total de aristas. De esta manera, los lazos tienen un efecto pequeño sobre las diversas propiedades de las gráficas formadas con este modelo, como la distancia promedio, el índice de agrupamiento, etcétera.
Denotaremos una gáfica aleatoria con una sucesión de grados esperados $\mathbf{w}$ por $G(\mathbf{w})$. Por ejemplo, la gráfica aleatoria típica $G(n, p)$ (ver [2].) de n vértices y densidad de aristas p es sólo una gráfica aleatoria con una sucesión de grados esperados (*pn, pn,..., pn*). Nuestro modelo es también diferente al modelo libre de escala.

Ahora necesitaremos algunas definiciones para algunas cantidades que asociaremos con G y $G(\mathbf{w})$.

Definición 1.8. *En una gráfica G, el volumen de un subconjunto S de vértices en G se define como vol(S)=* $\sum_{v \in S} deg(v)$, *donde deg(v) es el grado del vértice v; es decir, el volumen de S es la suma de los grados de todos los vértices en S. Para una gráfica G en G(*$\mathbf{w}$*), el grado esperado de* v_i *es exactamente* ω_i *y el volumen esperado de G es Vol(G)=*$\sum_i \omega_i$. *Para* $k \geq 2$, *definimos el momento k-ésimo del volumen esperado por* $Vol_k(S) = \sum_{v_i \in S} \omega_i^k$ *y escribimos* $Vol_k(G) = \sum_i \omega_i^k$.

Definición 1.9. *En una gráfica G, la distancia d(u, v) entre dos vértives u y v es la longitud del camino más corto que une a u con v (si éste existe). En una gráfica conexa G, la distancia promedio de G es el promedio sobre todas las distancias d(u, v) para u y v en G. Consideraremos gráfica poco densas que a menudo no son conexas. Si G no es conexa, definiremos la distancia promedio como el promedio entre todas las distancias d(u, v) para pares u y v en la cual ambos pertenecen a el mismo componente conexo.*

Definición 1.10. *El diámetro de G es la máxima distancia d(u, v), donde u y v están en el mismo componente conexo. Claramente, el diámetro es al menos tan grande como la distancia promedio.*

Todas las gráficas con las que trabajaremos típicamente tienen un único componente conexo grande, llamado el *componente gigante*, una fracción importante de aristas.

Definición 1.11. *El grado promedio de segundo orden se define como:*

$$\tilde{d} = \frac{Vol_2(G)}{Vol_1(G)} = \frac{\sum_i \omega_i^2}{\sum_i \omega_i}$$

Definición 1.12. *La sucesión de grados esperados* $\mathbf{w}$ *para una gráfica G de n vértices en G(*$\mathbf{w}$*) se dice que es fuertemente esparcida si cumple con las siguientes propiedades:*

i) El grado promedio de segundo orden satisface que $0 < \log(\tilde{d}) = o(\log(n))$.

ii) Para alguna constante $c > 0$*, todos, excepto o(n) vértices, tienen grado esperado* ω_i *que satisface* $\omega_i \geq c$ *. El grado esperado promedio d, definido como* $d = \sum_i \frac{\omega_i}{n}$ *, es estrictamente mayor que 1, es decir,* $d > 1 + \epsilon$ *para algún valor positivo* ϵ *independiente de n.*

Definición 1.13. *La sucesión de grados esperados* $\mathbf{w}$ *para una gráfica G sobre n vértices en G(*$\mathbf{w}$*) se dice admisible si la siguiente condición se cumplen, además del hecho de que* $\mathbf{w}$ *es fuertemente esparcida.*

iii) Existe un subconjunto U que satisface:

$$Vol_2(U) = (1 + o(1))Vol_2(G) \ll \frac{Vol_3(U)\log(\tilde{d})\log(\log(n))}{\tilde{d}\log(n)}$$

La primera restricción nos dice que $\tilde{d}$ está creciendo junto con n, pero de manera más lenta. Ésta crece lo suficientemente rápido como para que el componente más grande conexo tenga una fracción importante de nodos con grado esperado 1. La segunda restricción nos dice que pueden existir nodos con grado promedio significativamente mayor a d; pero no demasiados; esta condición acota superiormente el número de nodos con grado promedio grande.

Definición 1.14. *La sucesión de grados esperados* $\mathbf{w}$ *para una gráfica G de n vértices se llama especialmente admisible si i) se reemplaza por i') y iii) se reemplaza por iii');*

i') $\log(\tilde{d}) = O(\log(d))$.

iii')Existe un subconjunto U que satisface

$$Vol_3(U) = O(Vol_2(G))\frac{\tilde{d}}{\log(\tilde{d})}$$

y

$$Vol_2(U) \geq dVol_2(G)/\tilde{d}.$$

La primera restricción de la definición anterior tiene una explicación similar a la de la definición 1.13, sólo que ésta es menos fuerte, es decir, se cumple de manera más fácil. Nos dice que $\tilde{d}$ crece con la misma rapidez que d; esto nos asegura que no hay explosiones; es decir, el número de nodos con grados esperados grandes están acotados superiormente.

Definición 1.15. *Un componente conexo C, se llama ϵ-pequeño para un $\epsilon < \frac{1}{2}$ si el volumen de C es a lo más $\epsilon Vol(G)$.*

Definición 1.16. *Decimos que un componente es c-gigante si su volumen es al menos c Vol(G), para alguna constante* $c > 0$.

Para un subconjunto S de vértices, una medida típica es el número de vértices en S lo cual llamaremos el orden o tamaño de S. Para una gráfica aleatoria clásica $G(n,p)$, un componente gigante es un componente conexo que tiene al menos $c.n$ vértices, para alguna constante $c > 0$. Nuestra definición de componente gigante involucra el volumen en lugar del tamaño de la gráfica. De hecho, la definición de componente gigante que usa el tamaño de la gráfica simplemente no funciona para gráficas con una sucesión de grados esperados dados. Lo anterior es ilustrado por el siguiente ejemplo.

Ejemplo: Consideraremos la sucesión de grados $\mathbf{w}$ que consiste en n^{α} vértices con peso 2 y los otros vértices con peso 0. Donde α es una constante que satisface $\frac{1}{2} < \alpha < 1$. La gráfica aleatoria $G(\mathbf{w})$ es la unión de la gráfica aleatoria $G(n^{\alpha}, \frac{2}{n^{\alpha}})$ y algunos vértices aislados. Entonces el componente conexo más grande claramente no tiene $\Theta(n)$ vértices.

Definición 1.17. *En una gráfica G, denotamos a $\Gamma_k(u)$ al conjunto de vértices en G a distancia k del vértice u*

$$\Gamma_k(u) = \{v \in G : d(u, v) = k\}$$

Definición 1.18. *Sea S un conjunto de vértices de una gráfica G, denotamos a $\Gamma_i(S)$ al conjunto de vértices en $G \backslash S$ a distancia i de cualesquiera de los nodos de S.*

$$\Gamma_i(S) = \{v \in G \backslash S, u \in S : d(u, v) = k\}$$

Definición 1.19. *La distancia entre cualesquiera dos conjuntos S y T de vértices, denotada por $d(S, T)$, se define como la longitud del camino más corto que conecta a los conjuntos S y T*

$$d(S, T) = \text{mín}\{d(u, v) | u \in S, v \in T\}$$

1.2. Lemas y teoremas

En esta sección demostraremos lemas y teoremas que servirán de base para los resultados del siguiente capítulo.

Primero consideraremos la probabilidad de tener un componente conexo de tamaño k. Suponga que tenemos un subconjunto de vértices $S = \{v_{i_1}, v_{i_2}, ..., v_{i_k}\}$ con pesos $\omega_{i_1}, \omega_{i_2}, ..., \omega_{i_k}$.Para eso demostraremos unos lemas importantes:

Lema 1.1. *El valor esperado $E(X_k)$ del número de componentes conexas con tamaño k es a lo más*

$$E(X_k) \leq \sum_{S} \omega_{i_1}\omega_{i_2}...\omega_{i_k} Vol(S)^{k-2}\rho^{k-1}e^{-Vol(S)(1-(Vol(S)/Vol(G)))}$$

donde la suma se hace sobre todos los posibles conjuntos S de k vértices.

Prueba: La probabilidad de que S no esté conectado con ningún otro vértice del resto de la gráfica es

$$\prod_{v_i\in S, v_j\notin S} (1-\omega_i\omega_j\rho)$$
$$\approx e^{-\rho\sum_{v_i\in S, v_j\notin S}\omega_i\omega_j}$$
$$= e^{-\rho Vol(S)(Vol(G)-Vol(S))} \tag{1.1}$$

Si S es un componente conexo, la gráfica inducida por S contiene al menos un árbol generador T. La probabilidad de que contenga un árbol generador T fijo es

$$Pr(T) = \prod_{(v_{i_j}, v_{i_l})\in E(T)} \omega_{i_j}\omega_{i_l}\rho.$$

Entonces la probabilidad de que la gráfica inducida por S contenga un árbol generador es

$$\sum_T Pr(T) = \sum_T \prod_{(v_{i_j}, v_{i_l})\in E(T)} \omega_{i_j}\omega_{i_l}\rho.$$

donde T corre sobre todos los árboles generadores posibles sobre S.

Por una generalización del teorema de el árbol-matriz, el cual se puede ver enunciado y probado en [3], la suma de arriba es igual al determinante de cualquier matriz subprincipal de $k-1$ por $k-1$ de la matriz $D-A$, donde A es la matriz

$$A = \begin{bmatrix} 0 & \omega_{i_1}\omega_{i_2}\rho & \cdots & \omega_{i_1}\omega_{i_k}\rho \\ \omega_{i_2}\omega_{i_1}\rho & 0 & \cdots & \omega_{i_2}\omega_{i_k}\rho \\ \vdots & \vdots & \ddots & \vdots \\ \omega_{i_k}\omega_{i_1} & \omega_{i_k}\omega_{i_2} & \cdots & 0 \end{bmatrix}$$

y D es la matriz diagonal $diag(\omega_{i_1}(Vol(S)-\omega_{i_1})\rho, ..., \omega_{i_k}(Vol(S)-\omega_{i_k}-\omega_{i_k})\rho)$. Calculando el determinante, concluimos que

$$\sum_T P(T) = \omega_{i_1}\omega_{i_2}...\omega_{i_k} Vol(S)^{k-2}\rho^{k-1} \tag{1.2}$$

Combinando (1.1) y (1.2) tenemos el resultado.

□

Lema 1.2. *Para una $\epsilon < \frac{1}{2}$, el valor esperado $E(Y_k)$ del número de componentes conexas ϵ-pequeñas de tamaño k es a lo más*

$$E(Y_k) \leq \sum_S \omega_{i_1}\omega_{i_2}...\omega_{i_k} Vol(S)^{k-2}\rho^{k-1}e^{-Vol(S)(1-\epsilon)}$$

donde la suma es sobre todos los conjuntos de vértices S de tamaño k con $Vol(S) < Vol(G)$.

Prueba: La prueba se sigue inmediatamente del lema anterior.

$\square$

Algunos de los resultados principales de este capítulo son los siguientes tres teoremas:

Teorema 1.1. *Para cualquier $\epsilon < 1/2$ y $d > \frac{4}{e(1-\epsilon)^2} \approx (1+2\epsilon)1{,}4715...$, en una gráfica aleatoria $G(\mathbf{w})$ con grado promedio d, casi seguramente cada componente conexo que tiene volumen ϵ-pequeño tiene tamaño a lo más $\frac{\log(n)}{1+\log(d)-\log(4)+2\log(1-\epsilon)}$. La cota superior $\frac{\log(n)}{1+\log(d)-\log(4)}$ para componentes pequeños es asintóticamente la mejor para d grandes.*

Prueba: Suponga que G es una gráfica en $G(\mathbf{w})$ con un grado promedio esperado $d > 1+\delta$. Queremos mostrar que el número esperado $E(Y_k)$ de componentes conexas ϵ-pequeñas de tamaño k es chico.

Del Lema 1.2 podemos extraer una cota superior,$f(k)$.

$$f(k) = \sum_S \omega_{i_1}\omega_{i_2}...\omega_{i_k} Vol(S)^{k-2}\rho^{k-1}e^{-Vol(S)(1-\epsilon)}$$

Usando el hecho de que la función $x^{2k-2}e^{-x(1-\epsilon)}$ alcanza su máximo en $x =$

$(2k-2)/(1-\epsilon)$, tenemos que

$$\begin{aligned}
f(k) &= \sum_S \omega_{i_1}\omega_{i_2}...\omega_{i_k} Vol(S)^{k-2}\rho^{k-1}e^{-Vol(S)(1-\epsilon)} \\
&\leq \sum_S \frac{\rho^{k-1}}{k^k} Vol(S)^{2k-2}e^{-Vol(S)(1-\epsilon)} \\
&\leq \sum_S \frac{\rho^{k-1}}{k^k}(\frac{2k-2}{1-\epsilon})^{2k-2}e^{-(2k-2)} \\
&\leq \frac{n^k}{k}\frac{\rho^{k-1}}{k^k}(\frac{2k-2}{1-\epsilon})^{2k-2}e^{2k-2} \\
&\leq \frac{1}{4\rho(k-1)^2}(n\rho)^k(\frac{2}{1-\epsilon})^{2k}e^{-k} \\
&\leq \frac{1}{4\rho(k-1)^2}(\frac{4}{de(1-\epsilon)^2})^k.
\end{aligned}$$

La desigualdad de arriba es útil cuando $d > \frac{4}{e(1-\epsilon)^2}$ lo cual es proporcionado por el teorema que estamos demostrando. Si k satisface que $\frac{\log(n)}{1+\log(d)-\log(4)-2\epsilon} < k < \frac{2\log(n)}{1+\log(d)-\log(4)-2\epsilon}$, entonces

$$f(k) \leq \frac{1}{4n\rho(k-1)^2} = o(\frac{1}{\log(n)}).$$

Cuando k satisface que $\frac{2\log(n)}{1+\log(d)-\log(4)-2\epsilon} \leq k \leq n$, tenemos que

$$f(k) \leq \frac{1}{4n^2\rho(k-1)^2} = o(\frac{1}{n\log(n)}).$$

Para $k_0 = \frac{\log(n)}{1+\log(d)-\log(4)-2\epsilon}$, la probabilidad de que un componente ϵ-pequeño tenga tamaño $k > k_0$ es a lo más

$$\sum_{k>k_0} f(k) \leq \frac{\log(n)}{1+\log(d)-\log(4)-2\epsilon}.o(\frac{1}{\log(n)}) + n.o(\frac{1}{n\log(n)}) = o(1).$$

Entonces casi seguramente, el tamaño de los componentes ϵ-pequeños es a lo más $k_0 = \frac{\log(n)}{1+\log(d)-\log(4)-2\epsilon}$. Hemos demostrado el teorema.

Para mostrar que la cota superior que se dice en el teorema es asintóticamente la mejor posible para d grandes consideraremos el siguiente ejemplo.

Ejemplo: Consideraremos una gráfica aleatoria con los siguientes pesos como sucesión de grados esperados. Asumiremos que $d > 10$.

Hay $n^{\frac{2}{3}}$ vértices con peso $(d-1)n^{\frac{1}{3}}+1$. Cada uno de los $n - n^{\frac{2}{3}}$ vértices tienen peso 1. El grado promedio es exactamente d.

Sea S_1 el conjunto de vértices con peso 1, y S_2 el conjunto de vértices con peso $(d-1)n^{\frac{1}{3}}+1$. Sea G_i la gráfica inducida de G sobre S_i, para $i = 1, 2$. La gráfica G_2 es una gráfica aleatoria clásica $G(N, p)$, con $N = n^{\frac{2}{3}}$ y $Np = n^{\frac{2}{3}}((d-1)n^{\frac{1}{3}}+1)^2/(nd) = \Theta(\sqrt{N})$. Casi seguramente G_2 es conexo. De hecho, G_2 está contenido en el componente gigante de G. Sea c el número de vértices que no están el el componente gigante. Decimos que c está acotado por 0.

Para probar lo anterior, consideraremos un proceso de derivación especial. Primero tomaremos en cuenta todas las aristas de G_2. Entonces examinaremos la frontera de S_2 en S_1, la 2-frontera de S_2, y así sucesivamente, de manera que eventualmente expondremos todos los vértices del componente gigante de G. Para cualquier vértice $u \in S_1$, la probabilidad de que u esté en $\Gamma(S_2)$ es

$$1 - (1 - \frac{(d-1)n^{1/3}+1}{nd})^{n^{2/3}} \approx 1 - e^{-1+\frac{1}{d}}$$

El tamaño de $\Gamma(S_2)$ puede ser aproximado por la distribución binomial con $N = n - n^{2/3}$ y $p = 1 - e^{-1+\frac{1}{d}}$. Entonces, con probabilidad alta, el tamaño es alrededor de $(1 - e^{-1+\frac{1}{d}}n)$. Estimaremos el tamaño de $\Gamma_i(S_2)$ para $i > 1$ por inducción. Suponga que $|\Gamma_i(S_i)|$ está muy concentrada en $a_i n$ para alguna constante a_i para $i \geq 2$. Sea $c_i = 1 - \sum_{k=1}^{i} a_i$. Para cualquier vértice u que no está en $\bigcup_{j \leq i} \Gamma_i(S_i)$, la probabilidad de que $u \in \Gamma_{i+1}(S_2)$ es

$$1 - (1 - \frac{1}{nd})^{a_i n} \approx 1 - e^{-\frac{a_i}{d}}.$$

El tamaño de $\Gamma_{i+1}(S_2)$ puede ser aproximado por la distribución binomial con $N = c_i n$ y $p = 1 - e^{-\frac{a_i}{d}}$. Por definición de a_i, tenemos que

$$\begin{aligned} a_{i+1} &= c_i(1 - e^{-\frac{a_i}{d}}). \\ c_{i+1} &= c_i - a_{i+1} = c_i e^{-\frac{a_i}{d}}. \end{aligned}$$

Entonces

$$c_{i+1} = c_1 \prod_{k=1}^{i} e^{\frac{-a_i}{d}} = (1 - e^{-1+\frac{1}{d}})e^{-\frac{1-c_i}{d}}.$$

De la expresión anterior de recurrencia para c_i, vemos que el límite $c = lim_{i\to\infty} c_i$ existe y satisface que

$$c = (1 - e^{-1+\frac{1}{d}})e^{-\frac{1-c}{d}}.$$

Es fácil ver que la ecuación anterior tiene una solución única para $c \in [0,1]$, para $d > 1$ y la solución para c se incrementa como una función de d. Como hemos escogido a $d > 10$, c está acota por 0. Lo que aseguramos está probado.

El tamaño del segundo componente más grande puede ser estimado como sigue. Después de remover el componente gigante de G, la gráfica resultante es una gráfica aleatoria clásica $G(t,p)$ con $t = cn$ y $p = \frac{1}{nd} = \frac{c}{dt}$. Por [2], el componente más grande de $G(t, \frac{c}{dt})$ con $d < 1$ tiene tamaño a lo más

$$\frac{\log(n) - 5/2\log(\log(n))}{\frac{c}{d} - 1 - \log(\frac{c}{d})} = \frac{(1+o(1))\log(n)}{\log(d) - \log(c) - 1 + \frac{c}{d}}.$$

La constante $\frac{1}{\log(d)-\log(c)-1+\frac{c}{d}}$ está asintóticamente cerca a $\frac{1}{1+\log(d)-\log(4)}$ cuando d es grande y ϵ es arbitrariamente pequeña. Esto completa la prueba del teorema.

□

El siguiente lema nos será útil para probar el siguiente teorema, la prueba del mismo puede encontrarse en [4].

Lema 1.3. *Sean $X_1, ..., X_n$ variables aleatorias independientes con*

$$Pr(X_i = 1) = p_i;$$
$$Pr(X_i = 0) = 1 - p_i.$$

Para $X = \sum_{i=1}^n a_i X_i$, tenemos que $E(X) = \sum_{i=1}^n a_i p_i$ y definimos a $\nu = \sum_{i=1}^n a_i^2 p_i$. Entonces tenemos que

$$Pr(X < E(X) - \lambda) \leq e^{\frac{-\lambda^2}{2\nu}} \tag{1.3}$$

$$Pr(X > E(X) + \lambda) \leq e^{-\frac{\lambda^2}{2(\nu + a\lambda/3)}} \tag{1.4}$$

donde $a = \text{máx}\{a_1, a_2, ..., a_n\}$.

Teorema 1.2. *Para cualquier $\epsilon < 1/2$ y $\frac{1}{1-\epsilon} < d < \frac{2}{1-\epsilon}$, en una gráfica aleatoria $G(\mathbf{w})$ con grado promedio d, casi seguramente cada componente conexo con volumen ϵ-pequeño tiene a lo más $\frac{\log(n)}{d-1-\log(d)-\epsilon d}$. La cota superior $\frac{\log(n)}{d-1-\log(d)}$ para componentes pequeños es asintóticamente la mejor para d grandes.*

Prueba: Los métodos para probar el lema anterior no funcionan para demostrar este teorema por lo que recurriremos a otros para acotar superiormente a $f(k)$, la cota superior de $E(Y_k)$. Para eso partiremos en dos a la función $f(k)$.

$$f(k) = f_1(k) + f_2(k)$$

donde

$$f_1(k) = \sum_{Vol(S)<dk} \omega_{i_1}\omega_{i_2}...\omega_{i_k} Vol(S)^{k-2}\rho^{k-1}e^{-Vol(S)(1-\epsilon)};$$
$$f_2(k) = \sum_{Vol(S)\geq dk} \omega_{i_1}\omega_{i_2}...\omega_{i_k} Vol(S)^{k-2}\rho^{k-1}e^{-Vol(S)(1-\epsilon)}.$$

Para acotar a $f_1(k)$ notamos que $x^{2k-2}e^{-x(1-\epsilon)}$ es una función creciente cuando $x < (2k-2)/(1-\epsilon)$. Entonces tenemos que

$$Vol(S)^{2k-2}e^{-Vol(S)} \leq (dk)^{2k-2}e^{-dk(1-\epsilon)}$$

Como $Vol(S) < dk < (2k-2)/(1-\epsilon)$, esto implica que

$$\begin{aligned}
f_1(k) &= \sum_{Vol(S)<dk} \omega_{i_1}\omega_{i_2}...\omega_{i_k} Vol(S)^{k-2}\rho^{k-1}e^{-Vol(S)(1-\epsilon)} \\
&\leq \sum_{Vol(S)<dk} \frac{\rho^{k-1}}{k^k} Vol(S)^{2k-2}e^{-Vol(S)(1-\epsilon)} \\
&\leq \sum_{Vol(S)<dk} \frac{\rho^{k-1}}{k^k}(dk)^{2k-2}e^{-dk(1-\epsilon)} \\
&\leq \binom{n}{k}\frac{\rho^{k-1}}{k^k}(dk)^{2k-2}e^{-dk(1-\epsilon)} \\
&\leq \frac{n^k}{k}\frac{\rho^{k-1}}{k^k}(dk)^{2k-2}e^{-dk(1-\epsilon)} \\
&\leq \frac{1}{d^2k^2\rho}(n\rho)^k d^{2k} e^{(d(1-\epsilon)-1)k} \\
&= \frac{n}{dk^2}(\frac{d}{e^{d(1-\epsilon)-1}})^k
\end{aligned}$$

Ahora consideraremos a $f_2(k)$. Como $x^{2k-2}e^{x(1-\epsilon)}$ es una función decreciente cuando $x > (2k-2)/(1-\epsilon)$, tenemos que

$$Vol(S)^{k-2}e^{-Vol(S)(1-\epsilon)} \leq (dk)^{k-2}e^{-dk(1-\epsilon)}$$

Usando que $Vol(S) \geq dk \geq \frac{k-2}{1-\epsilon}$. Entonces, tenemos que

$$\begin{aligned}
f_2(k) &= \sum_{Vol(S)\geq dk} \omega_{i_1}\omega_{i_2}...\omega_{i_k} Vol(S)^{k-2}\rho^{k-1}e^{-Vol(S)(1-\epsilon)} \\
&\leq \sum_{Vol(S)\geq dk} \omega_{i_1}\omega_{i_2}...\omega_{i_k}\rho^{k-1}(dk)^{k-2}e^{-dk(1-\epsilon)} \\
&\leq \sum_{S} \omega_{i_1}\omega_{i_2}...\omega_{i_k}\rho^{k-1}(dk)^{k-2}e^{-dk(1-\epsilon)} \\
&< \frac{Vol(G)^k}{k}\rho^{k-1}(dk)^{k-2}e^{-dk(1-\epsilon)} \\
&\leq \frac{1}{d^2k^2\rho}d^k e^{-(d(1-\epsilon)-1)k} \\
&\leq \frac{n}{dk^2}(\frac{d}{e^{(d(1-\epsilon)-1)}})^k
\end{aligned}$$

Junto, tenemos que

$$f(k) = f_1(k) + f_2(k) \leq \frac{2n}{dk^2}(\frac{de}{e^{d(1-\epsilon)-1}})^k.$$

Si $\frac{\log(n)}{d(1-\epsilon)-1-\log(d)}\epsilon < k < \frac{2\log(n)}{d(1-\epsilon)-1-\log(d)}$, tenemos que

$$f(k) \leq \frac{2}{dk^2} = O(\frac{1}{\log^2(n)}).$$

Si $\frac{2\log(n)}{d(1-\epsilon)-1-\log(d)} \leq k \leq n$, entonces

$$f(k) \leq \frac{2}{ndk^2} = O(\frac{1}{n\log^2(n)}.$$

Sea $k_1 = \frac{\log(n)}{d(1-\epsilon)-1-\log(d)}$, la probabilidad de tener un componente ϵ-pequeño de tamaño $k > k_1$ es a lo más

$$\sum_{k>k_1} f(k) \leq \frac{\log(n)}{d(1-\epsilon)-1-\log(d)}.O(\frac{1}{\log^2(n)}) + n.O(\frac{1}{n\log^2(n)} = o(1).$$

Entonces casi seguramente el tamaño de un componente ϵ-pequeño es a lo más $k_1 = \frac{\log(n)}{d(1-\epsilon)-1-\log(d)}$.

Para ver que la cota superior es la mejor posible consideraremos el siguiente ejemplo. En la gráfica aleatoria $G(n, \frac{d}{n})$ con $d < 1$, el componente más grande tiene tamaño de alrededor $\frac{\log(n)-5/2\log(\log(n))}{d-1-\log(d)}$,ver [2], como se desea.

$\square$

Antes de probar el siguiente teorema, probaremos algunos lemas.

Lema 1.4. *Suponga que una gráfica aleatoria G en $G(\mathbf{w})$ tiene grado promedio $d > 1+\delta$, y contiene un subconjunto conexo que tiene más de C.log(n) vértices, donde C= $\{\frac{2}{\delta-\log(\delta)}, 10\}$. Entonces casi seguramente existe un componente gigante en G.*

Prueba: Para demostrar lo anterior utilizaremos los teoremas 1.1 y 1.2. Supongamos que la gráfica en cuestión sigue con las hipótesis del lema pero que no tiene un componente gigante, entonces, todos sus componentes deben ser pequeños; por lo tanto se cumplirían los teoremas 1.1 o 1.2 dependiendo el valor de d.

Caso 1: $1+\delta < d \leq 2$

Se sigue que:

$$C \geq \frac{2}{\delta - \log(\delta)} > \frac{1}{d-1-\log(d)-\epsilon_1 d}$$

Para alguna ϵ_1.

Lo anterior contradice el teorema 1.2, por lo que se concluye que si existe un componente gigante en G.

Caso 2: $2 < d$

Se sigue que:

$$C \geq 10 > \frac{1}{1+\log(d)-\log(4)+2\log(1-\epsilon_2)}$$

Para alguna ϵ_2.

Lo anterior contradice el teorema 1.1, por lo que se concluye que si existe un componente gigante en G.

$\square$

Lema 1.5. *Una subgráfica inducida H en $G \in G(\mathbf{w})$ es una gráfica aleatoria con una sucesión de grados esperados $\mathbf{w}'$ que consiste en $\omega_i' = \omega_i Vol(H)\rho$ para $v_i \in H$*

Prueba: Esto sigue fácilmente del hecho de que el grado esperado de v_i en H es justamente

$$\sum_{j\in V(H)} \omega_i\omega_j\rho = \omega_i Vol(H)\rho$$

□

Lema 1.6. *Suponga que en una gráfica aleatoria $G \in G(\mathbf{w})$, existe M (independiente de n) tal que $\omega_i \leq M\ \forall i$, y el grado promedio esperado $d \geq 1+\delta$, donde δ es una constante positiva. Entonces seguramente G tiene un único componente gigante.*

Prueba: Haremos un proceso de construcción; primero escogeremos un vértice u con peso mayor a 1 y llevaremos a cabo una búsqueda de su componente conexo. Sea X_k la suma de todos los pesos de los vértices a una distancia k de u. Suponga que $X_k \leq \epsilon Vol(G)$ para $\epsilon < 1/2$, es decir, todos estos vértices forman un componente ϵ-pequeño; además de que $\delta > 2\epsilon$ Para cualquier v_j, la probabilidad de que diste de u en k+1 es $X_k\omega_j\rho$. Entonces el valor esperado de X_{k+1} es $\sum_{v_j noinvolucrados} X_k\omega_j\rho \geq (1-\epsilon)X_d\tilde{d}$. Por el lema 1.4, tenemos que

$$Pr(X_{k+1} < (1-\epsilon)X_k\tilde{d} - \lambda) \leq e^{-\lambda^2/2\nu}$$

donde $\nu = \sum_j X_k\omega_j^3\rho \leq MX_k\tilde{d}$. Escogiendo a $\lambda = X_k(\tilde{d}-1-\epsilon)/2$, tenemos que

$$Pr(X_{k+1} < \frac{1}{2}(\tilde{d}-1-\epsilon)X_k) \leq e^{-\frac{(\tilde{d}-1-\epsilon)^2}{8M\tilde{d}}X_k}$$

Para cada k, X_k se incrementa por un factor de $\frac{1}{2}(\tilde{d}-1-\epsilon) > 1$ con una probabilidad de error de c^X , donde $c = e^{-\frac{(\tilde{d}-1)^2}{8M\tilde{d}}}$. Como $\sum_j c^j$ converge, existe una constante t_0 que satisface $\sum_{j\geq t_0} c^j < 1-\epsilon$ para una constante positiva ϵ. Con una probabilidad positiva constante, X_k se incrementa al menos por un factor de $\frac{1}{2}(\tilde{d}+1) > 1$ si $X > t_0$.

Como t_0 es una constante absoluta, el evento $X_1 > t_0$ pasa con una probabilidad positiva constante. Si este proceso se detiene rápido (es decir, la componente es ϵ-pequeña), entonces comenzamos de nuevo escogiendo otro vértice con peso mayor a 1; hay suficientes de estos vértices, al menos hay $\frac{d-1}{M}n$. Después de a lo más $\Theta(\log(n))$ intentos, el componente gigante se revelará.

□

Ahora demostraremos a otro de los principales lemas de este capítulo.

Lema 1.7. *Suponga que G es una gráfica aleatoria en $G(\mathbf{w})$ con una susesión de grados esperados dada $\mathbf{w}$. Si el grado promedio esperado d es estrictamente mayor a 1, entonces lo siguiente se cumple:*

(1) Casi seguramente G tiene un único componente gigante. Aún más, el volumen del componente gigante es al menos $(1 - \frac{2}{\sqrt[2]{de}} + o(1))Vol(G)$ si $d \geq \frac{4}{e} = 1{,}4715...$, y es al menos $(1 - \frac{1+\log(d)}{d} + o(1))Vol(G)$ si $d < 2$.

(2) El segundo componenete más grande tiene una talla casi seguramente de $O(\frac{\log(n)}{\log(d)})$.

Prueba: Sea $y < 1$ una constante que satisface $(1-y)^2 d > 1 + \delta/2$ (por ejemplo, escoge a $y = 1 - \delta/4$). Ordenamos los vértices de tal forma que $\omega_1 \leq \omega_2 \leq ... \leq \omega_n$. Sea i_0 el índice que denota al entero más grande que satisface:

$$\sum_{i \geq i_0} \omega_i \geq yVol(G) \tag{1.5}$$

Supongamos que $\omega_{i_0} > 2/y$, utilizando la subgráfica H_1 inducida por los últimos $n - i_0$ vértices veremos, por el lema 1.5 que sus grados son mayores a 2:

$$\omega_i' = \omega_i Vol(H_1)\omega\rho > \frac{2}{y}Vol(H_1)\rho \geq \frac{2}{y}y\frac{Vol(G)}{Vol(G)} = 2$$

Esta subgráfica inducida contiene una gráfica Erdos-Renyi $G(n-i_0, \frac{2}{n-i_0})$ entonces ésta contiene un componente de $c_1(n - i_0)$ vértices para alguna constante c_1.

Si suponemos que $\omega_{i_0} \leq \frac{2}{y}$ consideramos la gráfica inducida por los primeros i_0 vértices, H_2. Por el lema 1.5, y (1.5) se tiene que el $Vol(H_2) = (1-y)^2 dn$, entonces tiene grado promedio de al menos $1 + \frac{\delta}{2}$. Aún más, todos los vértices de esta subgráfica están acotados por $\frac{y}{2}$, utilizando el lema 1.6 se concluye que tiene un componente gigante de medida $c_2 i_0$ único.

De definición de gráfica aleatoria con grados esperados dados se tiene que:

$$\begin{aligned}
\text{máx}\{\omega_i^2\} &< \sum_i \omega_i = Vol(G) \\
\text{máx}\{\omega_i^2\} &< Vol(G) \\
\text{máx}\{\omega_i\} &< \sqrt{Vol(G)}
\end{aligned}$$

Entonces H_1 y H_2 tienen al menos $\sqrt{n}$ vértices. Utilizando el lema 1.4 se concluye que G tiene al menos un componente gigante. Ahora demostraré la unicidad del conponente gigante. Para eso supongamos que existen al menos dos componentes gigantes S_1 y S_2; entonces, por definición, existen $c_1, c_2 > 0$ tal que:

$$
\begin{aligned}
Vol(S_1) &= c_1 Vol(G) \\
Vol(S_2) &= c_2 Vol(G)
\end{aligned}
$$

La probabilidad de que S_1 y S_2 no estén conectados por una arista es a lo más:

$$
\begin{aligned}
\prod_{u\in S_1, v\in S_2} e^{-\omega_u\omega_v\rho} &= e^{-\sum_{u\in S_1, v\in S_2}\omega_u\omega_v\rho} \\
&= e^{-Vol(S_1)Vol(S_2)\rho} \\
&= e^{-c_1c_2Vol(G)^2\rho} \\
&= e^{-c_1c_2Vol(G)} \\
&\leq e^{-c_1c_2n} \\
&\leq e^{-c_1c_2\log(n)} \\
&= n^{-c_1c_2}
\end{aligned}
$$

Pero $n^{-c_1c_2} = o(1)$; entonces la probabilidad de que las componentes gigantes estén conectadas es al menos 1- o(1); por lo tanto el componente gigante es casi seguramente único.

Ahora probaremos la primera parte de la afirmación (1). Para eso supongamos que ésta no se cumple, es decir, sea $d \geq e$, entonces el volumen de el componente gigante es menor a $(1 - \frac{2}{\sqrt{de}} + o(1))Vol(G)$, por lo tanto existe $\epsilon < (1 - \frac{2}{\sqrt{de}} + o(1))$ tal que el volumen del comonente gigante G_1 es:

$$Vol(G_1) = \epsilon Vol(G)$$

Podemos aplicar el teorema 1.1, entonces el volumen del componente gigante es a lo más $\frac{\log(n)}{1+\log(d)-\log(4)+2\log(1-\epsilon)}$, por lo tanto, existe al menos un vértice con esperanza ω mayor o igual al promedio:

$$\omega \geq \frac{\epsilon Vol(G)}{\frac{\log(n)}{1+\log(d)-\log(4)+2\log(1-\epsilon)}} \geq c_\epsilon \frac{Vol(G)}{\log(n)}$$

Entonces

$$\begin{aligned}
&\omega^2 \geq c_\epsilon^2 \frac{Vol(G)^2}{(\log(n))^2}\\
&\omega^2\rho \geq c_\epsilon^2 \frac{Vol(G)}{(\log(n))^2} > 1\\
&\omega^2 \geq \frac{1}{\rho} = Vol(G)\\
&\max\{\omega^2\} \geq Vol(G)!
\end{aligned}$$

Ya que por la definición de gráfica aleatoria con grados esperados dados

$$\max\{\omega^2\} < Vol(G).$$

Por lo tanto, la primera parte de (1) está probada.

La otra parte de la afirmación (1) también la probaré por contradicción. Suponga que $1 < d < e$, entonces el volumen del componente gigante, G_2 es menor a $(1 - \frac{1+\log(d)}{d} + o(1))Vol(G)$, es decir, existe $\epsilon' < 1 - \frac{1+\log(d)}{d} + o(1)$ tal que

$$Vol(G_2) = \epsilon' Vol(G)$$

Por el teorema 1.2, el volumen del componente gigante es a lo más $\frac{\log(n)}{d-1-\log(d)-\epsilon' d}$. Entonces existe al menos un vértice mayor o igual al promedio,

$$\begin{aligned}
&\omega \geq \frac{\epsilon' Vol(G)}{\frac{\log(n)}{d-1-\log(d)-\epsilon d}} = c_{\epsilon'} \frac{Vol(G)}{\log(n)}\\
&\omega^2 \geq c_{\epsilon'} 2 \frac{Vol(G)^2}{(\log(n))^2}\\
&\omega^2\rho \geq c_{\epsilon'}^2 \frac{Vol(G)}{(\log(n))^2} > 1\\
&\omega^2 \geq \frac{1}{\rho} = Vol(G)\\
&\max\{\omega^2\} \geq Vol(G)!
\end{aligned}$$

por definición de gráfica aleatoria con grados esperados dados esta es una contradicción; por lo tanto (1) se cumple.

Ahora demostraré que el tamaño del segundo componente más grande es $O(\frac{\log(n)}{\log(d)})$, esto se sigue de los teoremas 1.1 y 1.2 ya que si $d \geq e$ entonces; por el teorema 1.1 el segundo componente más grande,G_3, tiene a lo más tamaño $\frac{\log(n)}{1+\log(d)-\log(4)}$, es decir,

$$Vol(G_3) \leq \frac{\log(n)}{1+\log(d)-\log(4)} = \frac{\log(n)}{\log(d)+\log(\frac{1}{4})} < \frac{\log(n)}{|log(\frac{1}{4})|\log(d)} = k\frac{\log(n)}{\log(d)}$$

Por lo tanto $Vol(G_3) = O(\frac{\log(n)}{\log(d)})$

Si $d \leq e$, por el teorema 1.2, se tiene que el componente más grande tiene tamaño a lo más:

$$Vol(G_3) \leq \frac{logn}{d-1-logd}$$

Pero, para $1 < d < e$, se tiene que

$$0{,}34 < \frac{\log(d)}{d-1-\log(d)} < 1$$

Por lo tanto; para $M > 1$ se tiene que

$$\frac{\log(d)}{d-1-\log(d)} < M$$

Entonces

$$Vol(G_3) < \frac{\log(n)}{d-1-\log(d)} < M\frac{\log(n)}{\log(d)}$$

En conclusión, $Vol(G_3) = O(\frac{\log(n)}{\log(d)})$

$\square$

También probaremos algunos teoremas y lemas relacionados con las distancias y vecindades en $G(\mathbf{w})$. Estos hechos serán útiles para las pruebas de teoremas propuestos más adelante.

Lema 1.8. *En una gráfica aleatoria G en $G(\mathbf{w})$ con una sucesión de grados esperados $\mathbf{w}$= $(\omega_1, \omega_2, ..., \omega_n)$, para cualquier par fijo de vértices (u,v), la distancia d(u,v) entre u y v es mayor que $\{\frac{\log(Vol(G))-c}{\log(\tilde{d})}\}$ con probabilidad de al menos 1 - $\frac{\omega_\upsilon \omega_\nu}{\tilde{d}(\tilde{d}-1)}e^{-c}$.*

Prueba: Sea $k = (\frac{\log(Vol(G))-c}{\log(\tilde{d}})$, de lo anterior se tiene que

$$\begin{aligned} &k\log(\tilde{d}) = \log(Vol(G)) - c \\ &\log(\tilde{d})^k = \log(Vol(G)) - c \\ &\tilde{d}^k = Vol(G)e^{-c} \end{aligned}$$

Por otro lado cualquier sucesión de vértices $\pi = (u = v_0, v_1, ..., v_{j-1}, v_j = v)$; la probabilidad de que π no sea un camino de G es:

$$\text{P}(\pi \text{ no sea un camino})= 1 - \text{P}(\pi \text{ sea un camino})$$

pero tenemos que:

$$\text{P}(\pi \text{ es un camino})= \omega_u\omega_{i_1}\rho\omega_{i_1}\omega_{i_2}\rho...\omega_{i_{j-1}}\omega_v\rho = \omega_u\omega_v\omega_{i_1}^2...\omega_{i_{j-1}}^2\rho^j$$

entonces

$$\text{P}(\pi \text{ no es un camino})=1- \omega_u\omega_v\omega_{i_1}^2...\omega_{i_{j-1}}^2\rho^j \text{ donde } \rho = \frac{1}{Vol(G)}$$

Ahora tomemos en cuenta la desigualdad FKG que nos dice lo siguiente (una prueba de esta desigualdad se puede encontrar en [5]):

Si A y B son eventos decrecientes entonces P($A \cap B$) $\geq$ P(A) P(B)

Lo anterior lo aplicamos a los eventos A= π no sea un camino de longitud j y B= que la longitud de π sea menor o igual que k, evidentemente

$$\text{P}(\text{d}(u,v) \geq \text{k})=\text{P}(\text{A}\cap \text{B})$$

Pero

$$\text{P(A)}= \prod_{i_1,...,i_{j-1}}(1 - \omega_u\omega_v\omega_{i_1}^2...\omega_{i_{j-1}}^2\rho^j)$$
$$\text{P(B)}= \prod_{j=1}^{k}(1 - \omega_u\omega_v\omega_{i_1}^2...\omega_{i_{j-1}}^2\rho^j)$$

Entonces, por la desigualdad FGK se concluye que:

$$\begin{aligned}
P(d(u,v) \geq k) \quad & \geq \prod_{j=1}^{k}\prod_{i_1,...,i_{j-1}}(1 - \omega_u\omega_v\omega_{i_1}^2...\omega_{i_{j-1}}^2\rho^j) \\
& \approx \prod_{j=1}^{k} e^{-\omega_u\omega_v\rho^j\sum_{\omega_1,...,\omega_{j-1}}\omega_1^2...\omega_{j-1}^2} \\
& \approx e^{-\omega_u\omega_v\sum_{j=1}^{k-1}\rho^j(\sum_{i=1}^{n}\omega_i^2)^{j-1}} \\
& \approx e^{-\omega_u\omega_v\rho((\sum_i\omega_i^2\rho)^{k-1}-1)/(\sum_i\omega_i^2\rho-1)} \\
& \geq e^{-\omega_u\omega_v e^{-c}/\tilde{d}(\tilde{d}-1)} \\
& \geq 1 - \frac{\omega_u\omega_v}{\tilde{d}(\tilde{d}-1)}e^{-c}
\end{aligned}$$

por la definición de k el lema queda demostrado.

$\square$

Lema 1.9. *En una gráfica aleatoria $G \in G(\mathbf{w})$, para cualesquiera dos subconjuntos S y T de vértices, tenemos que:*

$$Vol(\Gamma(S) \cap T) \geq (1-2\epsilon)Vol(S)\tfrac{Vol_2(T)}{Vol(G)}$$

con probabilidad de al menos 1-e^{-c} donde $\Gamma(S) = \{v : v \sim u \in S$ y $v \notin S\}$, y el Vol(S) satisface:

$$\frac{2cVol_3(T)Vol(G)}{\epsilon^2 Vol_2^2(T)} \leq Vol(S) \leq \frac{\epsilon Vol_2(T)Vol(G)}{Vol_3(T)} \tag{1.6}$$

Prueba: Sea X_j la variable aleatoria que nos indica que un vértice $v_j \in T$ está en $\Gamma(S)$, es decir,

$$X_j = \begin{cases} 1 & v_j \in \Gamma(S) \\ 0 & v_j \notin \Gamma(S) \end{cases}$$

Además sabemos que

$$P(X_j = 1) = 1 - \textstyle\prod_{v_i \in S}(1 - \omega_i\omega_j\rho)$$

Tomando en cuenta que $N =\mid S \mid$ y $C = \{1, ..., N\}$

$$1 - \prod_{v_i \in S}(1 - \omega_i\omega_j\rho) = 1 - 1 + \omega_j\rho\sum_{i=1}\omega_i - \sum_{k,l \in C k \neq l}\omega_k\omega_l\omega_j^2\rho^2 +$$
$$+ \sum_{k,l,m \in C k \neq l \neq m}\omega_k\omega_l\omega_m\omega_j^3\rho^3 + ... +$$
$$+(-1)^N \sum_{n_1,..,n_N \in C, n_1 \neq ... \neq n_N}\omega_{n_1}...\omega_{n_N}\omega_j^N\rho^N$$

Entonces

$$1 - \prod_{v_i \in S}(1 - \omega_i\omega_j\rho) = \omega_j\rho Vol(S) - \sum_{k,l \in C k \neq l}\omega_k\omega_l\omega_j^2\rho^2 + ... +$$
$$+(-1)^N \sum_{n_1,..,n_N \in C, n_1 \neq ... \neq n_N}\omega_{n_1}...\omega_{n_N}\omega_j^N\rho^N$$

Pero

$$\sum_{k,l \in C k \neq l}\omega_k\omega_l\omega_j^2\rho^2 \leq Vol_2(S)\omega_j^2\rho^2$$

y

0 $\sum_{k,l,m \in C k \neq l \neq m} \omega_k \omega_l \omega_m \omega_j^3 \rho^3 + ... + (-1)^N \sum_{n_1,..,n_N \in C, n_1 \neq ... \neq n_N} \omega_{n_1} ... \omega_{n_N} \omega_j^N \rho^N \geq$

por lo tanto,

$$P(X_j = 1) = 1 - \prod_{v_i \in S}(1 - \omega_i \omega_j \rho) \geq \omega_j \rho Vol(S) - Vol(S)^2 \omega_j^2 \rho^2$$

Por el lema 1.3:

$$P(X \geq E(X) - \lambda) = 1 - P(X \leq E(X) - \lambda) \geq 1 - e^{\lambda^2/2\nu}$$

tomando a $\lambda = \sqrt{2cVol(S)Vol_3(T)\rho}$, se tiene que

$$P(X > E(X) - \lambda) \geq 1 - e^{\frac{-2cVol(S)Vol_3(T)\rho}{2\nu}}$$

donde

$$\nu \geq \sum_{i=1}^{n} \omega_i^2 (Vol(S)\omega_j \rho) = Vol(S)\rho \sum_{i=1}^{n} \omega_i^3 = Vol(S)\rho Vol_3(T)$$

Entonces

$$P(X > E(X) - \lambda) \geq 1 - e^{-c} \tag{1.7}$$

Por otro lado tenemos que

$$Vol(\Gamma(S) \cap T) = \sum_{j \in T} \omega_j X_j \tag{1.8}$$

Entonces

$$\begin{aligned} E(Vol(\Gamma(S) \cap T)) \quad &= \sum_{j \in T} \omega_j (1 - \prod_{v_i \in S}(1 - \omega_i \omega_j \rho)) \\ &\geq \sum_{j \in T} \omega_j ((\omega_j \rho Vol(S)) - Vol(S)^2 \omega_j^2 \rho^2) \\ &= \sum_{j \in T} (\omega_j^2 \rho Vol(S) - Vol(S)^2 \omega_j^3 \rho^2) \\ &= \rho Vol(S) Vol_2(T) - Vol(S)^2 Vol_3(T) \rho^2 \qquad (1{,}9) \end{aligned}$$

Por (1.7), (1.8) y (1.9) se tiene que con probabilidad de al menos $1 - e^{-c}$

$$Vol(\Gamma(S)\cap T) = \sum_{v_j \in T} \omega_j X_j \geq Vol(S)Vol_2(T)\rho - Vol(S)Vol_3(T)\rho - \sqrt{-2cVol(S)Vol_3(T)\rho} \tag{1.9}$$

Ademas, de la desigualdad de la izquierda de (1.7)

$$\frac{2cVol_3(T)Vol(G)}{\epsilon^2 Vol_2(T)^2} \leq Vol(S)$$

$$\frac{2cVol_3(T)Vol(S)}{Vol_2(T)^2 Vol(S)^2} \leq \frac{2cVol_3(T)Vol(S)Vol(G)}{Vol_2(T)^2 Vol(S)^2} \leq \epsilon^2$$

$$\sqrt{2cVol_3(T)Vol(S)} \leq \epsilon Vol_2(T)Vol(S) \tag{1.10}$$

Por la desigualdad de la derecha de la expresión (1.7)

$$Vol(S) \leq \epsilon \frac{Vol_2(T)Vol(G)}{Vol_3(T)}$$

$$\begin{aligned} Vol_3(T)Vol(S)^2 \quad & \leq \epsilon Vol_2(T)Vol(G)Vol(S)\rho^2 \\ & \leq \epsilon Vol_2 Vol(S)\rho \qquad (1,12) \end{aligned}$$

Sustituyendo (1.11) y (1.12) en (1.10)

$$Vol(\Gamma(S) \cap T) \geq (1 - 2\epsilon)Vol(S)Vol_2(T)\rho$$

□

Lema 1.10. *Para cualesquiera dos subconjuntos ajenos S y T de vértices con $Vol(S)Vol(T) > cVol(G)$, tenemos que:*

$$Pr(d(S,T) > 1) < e^{-c}$$

donde d(S,T) denota la distancia entre S y T.

Prueba: Para cualesquiera vértices $v_i \in S$ y $v_j \in T$, la probabilidad de que $v_i v_j$ no es una arista es

$$1 - \omega_i \omega_j \rho$$

Como las aristas se eligen de manera independiente

$$\begin{aligned} P(d(S,T) > 1) &= \prod_{v_i \in S v_j \in T} (1 - \omega_i \omega_j) \\ &\leq e^{-Vol(S)Vol(T)\rho} \\ &< e^{-c} \end{aligned}$$

□

Lema 1.11. *Suponga que G es una gráfica aleatoria sobre n vértices tal que* $\min\{\omega_i\omega_j\rho\} \geq \frac{c\log(n)}{n}$*,con una constante $c \leq 2$. G tiene grado promedio d estrictamente mayor a 1. Entonces para cualquier vértice v fijo en el componente gigante, si $\tau = o(\sqrt[2]{n})$, entonces existe un índice $i_0 \leq c_0\tau$. Y con una probabilidad de al menos 1 - $o(n^{-1})$, tenemos que*

$$Vol(\Gamma_{i_0}(u)) \geq \tau$$

donde denotamos a $\Gamma_i(S) = \Gamma(\Gamma_{i-1}(S))$ para $i > 1$ y $\Gamma_1(S) = \Gamma(S)$.

Prueba: Sea $\tau = \frac{M}{c}$ donde M es una constante mayor a 1. Evidentemente

$$\tau = \frac{M}{c} = o(\sqrt{n})$$

Por otro lado tomemos a $k = \lfloor \frac{1}{c} \rfloor - 1$. Como u está en el componente gigante, $Vol(\Gamma_k(u)) \geq 1$. Existe un camino $u_0, u_1, ..., u_k$ que satisface la condición de que $u_j \in \Gamma_j(u)$ para $1 \leq j \leq k$. Escribiremos $u_0 = u$.

Sea $f(u_j)$ una función que denota el número de vértices y, donde $u_j y$ es una arista pero y no es ninguno de los $u_0, u_1, ..., u_k$ vértices.

Calcularemos la probabilidad de que $f(u_j) \leq \frac{\tau}{d}$ como

$$
\begin{aligned}
Pr(f(u_j) \leq \tfrac{\tau}{d}) \quad & \leq \sum_{l=0}^{\frac{\tau}{d}} \binom{n-k-1}{l} \prod_{i_1,\ldots,i_{n-k-l-1}} \prod_{i_1,\ldots,i_l} (\omega_j \omega_{i_j} \rho)(1-\omega_j \omega_{i_k}) \\
& \leq \sum_{l=0}^{\frac{\tau}{d}} \frac{n^l}{l!} (\frac{c \log(n)}{n})^l e^{-\omega_j \rho \sum \omega_{i_k}} \\
& \leq \sum_{l=0}^{\frac{\tau}{d}} \frac{(c \log(n))^l}{l!} e^{-(n-k-l-1)\frac{c \log(n)}{n}} \\
& \leq (c \log(n))^{\frac{\tau}{d}} e^{-(n-k-1-\frac{\tau}{d})\frac{c \log(n)}{n}} \sum_{l=0}^{\frac{\tau}{d}} \frac{1}{l!} \\
& \leq (c \log(n))^{\frac{\tau}{d}} e^{-(1-\frac{(k+1+\frac{\tau}{d})}{n})c \log(n)} e \\
& = o(n^{-c+\epsilon})
\end{aligned}
$$

Para cualquier $\epsilon > 0$ pequeña- Las $f(u_j)$ son variables aleatorias independientes. La probabilidad de que $f(u_j) \leq \tau$ para $0 \leq j \leq k$ es a lo más

$$o((n^{-a+\epsilon})^{k+1}) = o(n^{-1})$$

si ϵ es suficientemente pequeño.

Con probabilidad de al menos $1-o(n^{-1})$ existe un índice $1 \leq i_0 \leq k+1 \leq \tau$ que satisface que $f(u_{i_0-1}) \geq \frac{\tau}{d}$. Entonces $Vol(\Gamma_{i_0}(u)) \geq \tau$.

$\square$

Capítulo 2

Teoremas relevantes para G(w)

Ahora propondremos el siguiente teorema; central en este capítulo.

Teorema 2.1. *Para una gráfica aleatoria G con una sucesión de grados esperados admisible $(\omega_1, ..., \omega_n)$ la distancia promedio es casi seguramente $(1 + o(1))\frac{\log(n)}{\log(\tilde{d})}$.*

Prueba: Del lema 1.7 sabemos que con una alta probabilidad el componente gigante de la gráfica tiene un volumen de al menos $O(Vol(G))$. Del mismo lema sabemos que las otras componentes tiene un volumen de a lo más $O(\frac{\log(n)}{\log(d)})$. Como $\frac{\log(n)}{\log(d)}$ es mucho menor a $Vol(G)$ se concluye que la distancia promedio está determinada básicamente por el componente gigante.

Observamos que para cualquier vértice u en el componente gigante, aplicando el lema 1.11 con $i_0 \leq C.\epsilon.\frac{\log(n)}{\log(\tilde{d})}$; entonces $\Gamma_{i_0}(u)$ satisface que

$$Vol(\Gamma_{i_0}(u)) \geq \epsilon\frac{\log(n)}{\log(\tilde{d})}$$

con probabilidad $1 - o(1)$.

Ahora utilizaremos el lema 1.9, con $T = G$, $S = \Gamma_{i_0}(u)$ y $c = 2\log(\log(n))$. Esto es posible ya que, por ser G admisible

$$\frac{2cVol_3Vol(G)}{\epsilon^2Vol_2^2(G)} \leq \frac{2c\tilde{d}\log(n)}{\log(\tilde{d})\log(\log(n))Vol_2(G)} = \frac{4\tilde{d}\log(n)}{\log(\tilde{d})Vol_2(G)} =$$

$$= \frac{4\log(n)}{Vol(G)\log(\tilde{d})} < \frac{\epsilon\log(n)}{\log(\tilde{d})} \leq Vol(\Gamma_{i_0}(u))$$

Siempre y cuando

$$Vol(\Gamma_{i_0}(u)) \leq \epsilon\frac{Vol_2(G)}{Vol_3(G)}Vol(G)$$

el lema 1.9 nos dice que

$$Vol(\Gamma_{i_0}(u)) \geq \epsilon(1-2\epsilon)\tilde{d}\frac{\log(n)}{\log(\tilde{d})}$$

con probabilidad de $1-e^{-c}$.

Aplicandolo i_1 veces con $i_1 = \log(\frac{\sqrt{cVol(G)}\log(\tilde{d})}{\epsilon\log(n)}).\frac{1}{2\log((1-2\epsilon)\tilde{d})}$ se tiene que

$$Vol(\Gamma_{i_0+i_1}(u)) \geq \epsilon((1-2\epsilon)\tilde{d})^{i_1}\frac{\log(n)}{\log(\tilde{d})} = \sqrt{cVol(G)}$$

con probabilidad $1-i_1e^{-c} = 1-o(1)$.

Teniendo $i_0+i_1 = (1+o(1))\frac{\log(n)}{\log(\tilde{d})}$.

Por otro lado, sea v un vértice en el componente gigante; por un análisis análogo existen enteros i'_0 e i'_1 que satisface que $i'_0+i'_1 = (1+o(1))\frac{\log(n)}{\log(\tilde{d})}$ tal que $Vol(\Gamma_{i'_0+i'_1}(v)) \geq \sqrt{cVol(G)}$ con probabilidad $1-o(1)$.

Por el lema 1.10, aplicado a los conjuntos $\Gamma_{i_0+i_1}(u)$ y $\Gamma_{i'_0+i'_1}(v)$ se tiene que con probabilidad $1-o(1)$ existe un camino entre u y v con longitud $(1+o(1))\frac{\log(n)}{\log(\tilde{d})}$. Entonces, casi seguramente, la distancia promedio de una gráfica con una sucesión de grados esperados admisible es $(1+o(1))\frac{\log(n)}{\log(\tilde{d})}$.

□

Corolario 2.1. *Si $np \geq c > 1$ para alguna constante c, entonces casi seguramente, la distancia promedio de $G(n,p)$ es $(1+o(1))\frac{\log(n)}{\log(np)}$, proveyendo a $\frac{\log(n)}{\log(np)}$ se va a infinito como $n \longrightarrow \infty$.*

Prueba: *Esto se sigue tomando a $\omega_i = np$ y U como el conjunto de todos los vértices. Es fácil verificar que en el caso en el que* **w** *es admisible el teorema 2.1 aplica.*

□

Teorema 2.2. *Para una gráfica aleatoria con una secesión de grados especialmente admisible $(\omega_1, ..., \omega_n)$ el diámetro es casi seguramente $\Theta(\log(n)/\log(\tilde{d}))$.*

Prueba: Aplicando el lema 1.8, escogiendo a $c = \epsilon\log(n)$, para una $\epsilon > 0$ fija, se tiene que con probabilidad de al menos $1-\frac{\omega_u\omega_v}{\tilde{d}(\tilde{d}-1)}n^{\epsilon}$ la distancia entre u y v satisface que:

$$d(u,v) > \frac{\log(Vol(G)) - \epsilon\log(n)}{\log(\tilde{d})} \tag{2.1}$$

Pero $Vol(G) = dn$ entonces $\log(Vol(G)) = \log(d) + \log(n)$. Como $\frac{\log(d)}{\log(n)} = o(1)$ se tiene que $\frac{\log(d)}{\log(n)} = o(1)$ entonces $\log(d) = o(1)\log(n)$. Por lo tanto $\log(Vol(G)) = (1 + o(1))\log(n)$.

Sustituyendo lo anterior en (2.1)

$$d(u, v) > \frac{(1 + o(1))\log(n) - \epsilon\log(n)}{log(\tilde{d})} = (1 - \epsilon + o(1))\frac{\log(n)}{\log(\tilde{d})}$$

con probabilidad $1 - o(1)$, o sea, casi seguramente, la distancia promedio de G es al menos $(1 + o(1))\frac{\log(n)}{\log(\tilde{d})}$.

Para la cota superior hacemos un análisis análogo al del teorema anterior; lo único que hay que verificar es la condición de cota inferior para la aplicación del lema 1.9; esto se verifica ya que G es especialmente admisible

$$\frac{2cVol_3Vol(G)}{\epsilon^2 Vol_2^2(G)} \leq \frac{2cM\tilde{d}\log(n)}{\log(\tilde{d})\log(\log(n))Vol_2(G)} = \frac{4M\tilde{d}\log(n)}{\log(\tilde{d})Vol_2(G)} =$$
$$= \frac{4M\log(n)}{Vol(G)\log(\tilde{d})} < \frac{\epsilon\log(n)}{\log(\tilde{d})} \leq Vol(\Gamma_{i_0}(u))$$

Por lo tanto se tiene una cota superior, concluyendo que el diámetro es $\Theta(\frac{\log(n)}{\log(\tilde{d})})$

□

Corolario 2.2. *Si $np = c > 1$ para alguna constante c, entonces casi seguramenteel diámetro de $G(n, p)$ es $\Theta(\log(n))$*

Prueba: Se sigue del teorema 2.2

□

Definición 2.1. *Para gráficas aleatorias con una sucesión de grados esperados dados que satisfacen una distribución libre de escala con exponente β, asumiré que los grados esperados son de la forma $\omega_i = ci^{-\frac{1}{\beta-1}}$ para i que satisface que $i_0 \leq i < n + i_0$.*

Por la definición anterior:

$$Vol(G) = \sum_{i=i_0}^{n+i_0} \omega_i = \sum_{i=i_0}^{n+i_0} ci^{-\frac{1}{\beta-1}} = c\sum_{i=i_0}^{n+i_0} i^{-\frac{1}{\beta-1}} \approx c\int_{i_0}^{n+i_0} i^{-\frac{1}{\beta-1}} \approx \int_0^n i^{-\frac{1}{\beta-1}} \approx c\frac{\beta-1}{\beta-2}n^{1-\frac{1}{\beta-1}} \tag{2.2}$$

Entonces:

$d = \frac{Vol(G)}{n} \approx c\frac{\beta-1}{\beta-2}n^{-\frac{1}{\beta-1}}$
Por lo tanto

$$c = \frac{\beta-2}{\beta-1}dn^{\frac{1}{\beta-1}} \tag{2.3}$$

También sabemos que:

$$\omega_{i_0} = m$$

donde m es el grado esperado máximo e i_0 es el vértice donde se comienza el conteo. Entonces:

$$m = ci_0^{-\frac{1}{\beta-1}} \tag{2.4}$$

Sustituyendo (2.3) en (2.4) :

$$m = \frac{\beta-2}{\beta-1}dn^{\frac{1}{\beta-1}}i_0^{-\frac{1}{\beta-1}}$$
$$i_0^{-\frac{1}{\beta-1}} = \frac{m}{d}\frac{\beta-1}{\beta-2}n^{-\frac{1}{\beta-1}}$$
$$i_0^{\frac{1}{\beta-1}} = \frac{d}{m}\frac{\beta-2}{\beta-1}n^{\frac{1}{\beta-1}}$$

$$i_0 = (\frac{d}{m}\frac{\beta-2}{\beta-1})^{\beta-1}n \tag{2.5}$$

Por lo tanto, c depende del grado promedio e i_0 depende del grado máximo esperado m.

Ahora calcularemos a $\tilde{d}$ si $\beta > 3$. Por definición:

$$Vol_2(G) = \sum_{i=i_0}^{n+i_0} \omega_i^2 = \sum_{i=i_0}^{n+i_0} c^2 i^{-\frac{2}{\beta-1}} = c^2 \sum_{i=i_0}^{n+i_0} i^{-\frac{2}{\beta-1}}$$

$$\approx c^2 \int_{i_0}^{n+i_0} i^{-\frac{2}{\beta-1}} \approx c^2 \int_0^n i^{-\frac{2}{\beta-1}} = c^2\frac{\beta-1}{\beta-3}n^{1-\frac{2}{\beta-1}} \tag{2.6}$$

Entonces utilizando (2.2) y (2.6) se tiene que:

$$\tilde{d} = \frac{Vol_2(G)}{Vol(G)} = \frac{c^2\frac{\beta-1}{\beta-3}n^{-\frac{2}{\beta-1}}}{c\frac{\beta-1}{\beta-2}n^{1-\frac{1}{\beta-1}}} = c\frac{\beta-2}{\beta-3}n^{-\frac{1}{\beta-1}}$$

$$= \frac{(\beta-2)^2}{(\beta-1)(\beta-3)} dn^{-\frac{1}{\beta-1}} = \frac{(\beta-2)^2}{(\beta-1)(\beta-3)} d \tag{2.7}$$

Si $\beta = 3$ se tiene que:

$$Vol_2(G) = \sum_{i=i_0}^{n+i_0} \omega_i^2 = c^2 \sum_{i=i_0}^{n+i_0} i^{-\frac{2}{\beta-1}} = c^2 \sum_{i=i_0}^{n+i_0} i^{-1} \approx c^2 \int_{i_0}^{n+i_0} \frac{1}{i} = c^2 \int_{1}^{n+1} \frac{1}{i} = c^2 Ln(n+1)$$

Entonces:

$$\tilde{d} = \frac{Vol_2(G)}{Vol(G)} = \frac{c^2 Ln(n+1)}{cn^{\frac{1}{2}}} = \frac{c}{n^{\frac{1}{2}}} Ln(n+1) = \frac{1}{2} \frac{dn^{\frac{1}{2}}}{n^{\frac{1}{2}}} Ln(n+1) =$$

$$\frac{1}{2} dLn(n+1) = \frac{1}{2} dLn \frac{2m}{d} i_0^{\frac{1}{2}} + 1 = \frac{1}{2} dLn \frac{2m}{d} (i_0^{\frac{1}{2}} + \frac{d}{2m}) \tag{2.8}$$

Si $2 < \beta < 3$ se encuentra que

$$Vol_2(G) = \sum_{i=i_0}^{n+i_0} \omega_i^2 = c^2 \sum_{i=i_0}^{n+i_0} i^{-\frac{2}{\beta-1}} \approx c^2 \int_{i_0}^{n+i_0} i^{-\frac{2}{\beta-1}} \approx c^2 \int_{0}^{n} i^{-\frac{2}{\beta-1}} = c^2 \frac{\beta-1}{3-\beta} n^{1-\frac{2}{\beta-1}}$$

Entonces

$$\tilde{d} = \frac{Vol_2(G)}{Vol(G)} = \frac{c^2 \frac{\beta-1}{\beta-3} n^{-\frac{2}{\beta-1}}}{c\frac{\beta-1}{\beta-2} n^{1-\frac{1}{\beta-1}}} = c\frac{\beta-2}{3-\beta} n^{-\frac{1}{\beta-1}} = \frac{(\beta-2)^2}{(3-\beta)(\beta-1)} d \tag{2.9}$$

En conclusión

$$\tilde{d} = \begin{cases} \frac{(\beta-2)^2}{(\beta-1)(\beta-3)} d & \beta > 3 \\ \frac{1}{2} dLn(\frac{2m}{d}(i_0^{\frac{1}{2}} + \frac{d}{2m})) & \beta = 3 \\ \frac{(\beta-2)^2}{(3-\beta)(\beta-1)} d & 2 < \beta < 3 \end{cases}$$

De lo anterior podemos concluir que las gráficas libres de escala con exponentes β en cada uno de los casos de la ecuación anterior son diferentes entre si.

El desarrollo que acabamos de hacer nos da pie a la demostración de otro de los teoremas centrales de este trabajo.

Teorema 2.3. *Para una gráfica libre de escala con exponente $\beta > 3$ y grado promedio d estrictamente mayor a 1, casi seguramente la distancia promedio es $(1 + o(1))\frac{\log(n)}{\log(\tilde{d})}$ y el diámetro es $\Theta(\log(n))$.*

Prueba: Basta demostrar que este tipo de gráficas son admisibles y especialmente admisibles y aplicar los teoremas 2.1 y 2.2.

Primero demostraremos que este tipo de gráficas cumplen la propiedad i) de la definición de gráfica admisible.

P.D. $\log(\tilde{d}) = o(\log(n))$

Se sabe que $\tilde{d} = \frac{(\beta-2)^2}{(\beta-1)(\beta-3)} = c\frac{(\beta-2)}{(\beta-3)}n^{-\frac{1}{\beta-1}}$

Si $n > e^{\frac{logc\frac{\beta-2}{\beta-3}}{\epsilon+\frac{1}{\beta-1}}}$ se tiene que:

$$\log(n) > \log e^{\frac{\log(c)\frac{\beta-2}{\beta-3}}{\epsilon+\frac{1}{\beta-1}}} = \frac{\log(c)\frac{\beta-2}{\beta-3}}{\epsilon+\frac{1}{\beta-1}}$$
$$(\epsilon + \frac{1}{\beta-1})\log(n) > \log(c)\frac{\beta-2}{\beta-3}$$
$$\epsilon\log(n) > \log(c)\frac{\beta-2}{\beta-3} - \frac{1}{\beta-1}\log(n)$$
$$\epsilon\log(n) > \log(c)\frac{\beta-2}{\beta-3}n^{-\frac{1}{\beta-1}} = \log(\tilde{d})$$

Por lo tanto $\log(\tilde{d}) = o(\log(n))$.

Ahora demostraremos la propiedad i') para este tipo de gráficas.

P.D. $\log(\tilde{d}) = O(\log(d))$

Sabemos que

$$\tilde{d} = \frac{(\beta-2)^2}{(\beta-1)(\beta-3)}d$$

Entonces

$$\log\tilde{d} = \log(\frac{(\beta-2)^2}{(\beta-1)(\beta-3)})d = \log(\frac{(\beta-2)^2}{(\beta-1)(\beta-3)}) + \log(d)$$

Pero $\frac{(\beta-2)^2}{(\beta-1)(\beta-3)}$ es una función continua bien definida en $[4,\infty)$ y $lim_{\beta\to\infty}\frac{(\beta-2)^2}{(\beta-1)(\beta-3)} = 1$, por lo tanto, por la regla de L'Hopital, se tiene que $\frac{(\beta-2)^2}{(\beta-1)(\beta-3)}$ está acotada. Supongamos que c_0 es la cota, entonces:

$$\log(\frac{(\beta-2)^2}{(\beta-1)(\beta-3)}) + \log(d) < \log(c_0) + \log(d)$$

Como $\log(d)$ es una función continua y el $lim_{x\to\infty}\log(x) = \infty$ se puede afirmar que existe d_0 tal que si $d > d_0$ se tiene que $\log(c_0) < \log(d)$, por lo tanto:

$$\log(c_0) + \log(d) < 2\log(d)$$

Por lo tanto

$$\log(\tilde{d}) < 2\log(d)$$

Concluimos que $\log(\tilde{d}) = O(\log(d))$. Falta demostrar la propiedad ii) para este tipo de gráficas.

P.D. Para alguna constante $c' > 0$, todos, excepto $o(n)$ vértices tienen un grado esperado que satisface $\omega_i \geq c'$. Como

$$\begin{gathered}\omega_i = ci^{-\frac{1}{\beta-1}} \; i_0 \leq i \leq n + i_0 \\ \text{mín}\{\omega_i\} = c(n+i_0)^{\frac{-1}{\beta-1}} \\ \text{si } c' = c(n+i_0)^{\frac{-1}{\beta-1}}\end{gathered}$$

se cumple que

$$\omega_i \geq c' \text{ para todo } i_0 \leq i \leq n + i_0$$

otra parte de la propiedad ii) que debemos demostrar es la siguiente

P.D. El grado promedio esperado es de la forma $d > 1 + \epsilon$

Pero esto se cumple por hipótesis.

Sólo resta desmostrar la existencia de un conjunto U de vértices que cumplen las propiedades iii) y iii'). Para eso consideraremos un conjunto U_y que consiste en todos los vértices con peso menor o igual a $d\frac{\beta-2}{\beta-1}y$ donde y es un número fijo. Antes de verificar las propiedades iii) y iii') desarrollaremos algunas expresiones. Primero calcularemos el $Vol_2(U_y)$ para eso necesitamos determinar el índice en el que el peso es igual a $d\frac{\beta-2}{\beta-1}y$.

$$\begin{aligned}&\omega_i = d\frac{\beta-2}{\beta-1}y \\ &ci^{-\frac{1}{\beta-1}} = d\frac{\beta-2}{\beta-1}y \\ &\frac{\beta-2}{\beta-1}dn^{\frac{1}{\beta-1}}i^{-\frac{1}{\beta-1}} = d\frac{\beta-2}{\beta-1}y \\ &i^{\frac{1}{\beta-1}} = n^{\frac{1}{\beta-1}}y^{-1} \\ &i = ny^{-(\beta-1)}\end{aligned}$$

Por lo tanto

$$Vol_2(U_y) = \sum_{i=ny^{-(\beta-1)}}^{n} \omega_i^2 = c^2 \sum_{ny^{-(\beta-1)}}^{n} i^{-\frac{2}{\beta-1}} \approx c^2 \int_{ny^{-(\beta-1)}}^{n} i^{-\frac{2}{\beta-1}} =$$
$$= c^2 \frac{\beta-1}{\beta-3} i^{1+\frac{1}{\beta-1}}|_{ny^{-(\beta-1)}}^{n} = \frac{(\beta-2)^2}{(\beta-1)(\beta-3)} d^2 n^{\frac{2}{\beta-1}} \frac{\beta-1}{\beta-3}(n^{1+\frac{2}{\beta-1}} - (ny^{-(\beta-1)})^{1-\frac{2}{\beta-1}}) =$$
$$= \frac{(\beta-2)^2}{(\beta-1)(\beta-3)} d^2 n - \frac{(\beta-2)^2}{(\beta-1)(\beta-3)} d^2 n y^{-(\beta-3)} =$$
$$= d^2 \frac{(\beta-2)^2}{(\beta-1)(\beta-3)} n(1 - y^{-(\beta-3)}).$$

Ahora calcularemos a $Vol_3(Uy)$ si $\beta > 4$

$$Vol_3(U_y) = \sum_{i=ny^{-(\beta-1)}} \omega_i^3 = \sum_{i=ny^{-(\beta-1)}} c^3 i^{-\frac{3}{\beta-1}} \approx$$
$$\approx c^3 \int_{ny^{-(\beta-1)}}^{n} = c^3 \frac{\beta-1}{\beta-4} i^{1-\frac{3}{\beta-1}}|_{ny^{-(\beta-1)}}^{n} =$$
$$= \frac{(\beta-2)^2}{(\beta-1)^3} d^3 n^{\frac{3}{\beta-1}} \frac{\beta-1}{\beta-4}(n^{1-\frac{3}{\beta-1}} - n^{1-\frac{3}{\beta-1}} y^{-(\beta-2)(1-\frac{3}{\beta-1})}) =$$
$$= \frac{(\beta-2)^3}{(\beta-1)^2(\beta-4)} d^3 n(1 - y^{-(\beta-4)})$$

Si $\beta = 4$ se tiene que:

$$Vol_3(U_y) = \sum_{ny^{-(\beta-1)}}^{n} \omega_i^3 = c^3 \sum_{ny^{-(\beta-1)}}^{n} i^{-1} \approx c^3 \int_{ny^{-(\beta-1)}}^{n} \frac{1}{i} =$$
$$= c^3 Lni|_{ny^{-(\beta-1)}}^{n} = \frac{8}{27} d^3 n Lny^3$$

Si $3 < \beta < 4$ se tiene el mismo resultado que para $\beta > 3$; excepto que en lugar de poner $(\beta - 4)$ se reemplaza por $(4 - \beta)$; es decir,

$$Vol_3(U_y) = \frac{(\beta-2)^2}{(\beta-1)^2(4-\beta)} d^3 n(1 - \lceil y^{-(4-\beta)} \rceil)$$

Por lo tanto concluimos que

$$Vol(U_y) = \begin{cases} d^3\frac{(\beta-2)^3}{(\beta-1)^2(\beta-4)}(1-y^{-(\beta-4)}) & \beta > 4 \\ \frac{8}{27}d^3nLn(y^3) & \beta = 4 \\ d^3\frac{(\beta-2)^3}{(\beta-1)(4-\beta)}n(y^{-(4-\beta)}) & 3 < \beta < 4 \end{cases} \tag{2.10}$$

Ahora consideraremos tres casos para demostrar las propiedades iii) o iii').

Caso 1: $\beta > 4$. Escogemos a $y = n^{\frac{1}{4}}$ y $U = U_y$, entonces

$$\begin{aligned} Vol_2(U) &= d^2\frac{(\beta-2)^2}{(\beta-1)(\beta-3)}n(1+o(1)) = (1+o(1))Vol_2(G) \\ \frac{Vol_3(U)}{Vol_2(U)} &= (1+o(1))d\frac{(\beta-1)(\beta-3)}{9} = O(\frac{\tilde{d}}{log(\tilde{d})}). \end{aligned}$$

Entonces la sucesión de grados libre de escala es admisible y especialmente admisible.

Caso 2: $\beta = 4$. Para probar que la sucesión es admisible escogemos a $y = e^{\sqrt{\frac{log(n)}{log(d)loglog(n)}}}$. Entonces $U = U_y$ satisface que

$$\begin{aligned} Vol_2(U) &= d^2\frac{(\beta-2)^2}{(\beta-1)(\beta-3)}n(1+o(1)) = (1+o(1))Vol_2(G) \\ \frac{Vol_3(U)}{Vol_2(U)} &= (1+o(1))d\frac{2}{9}log(y) = o(\frac{\tilde{d}}{log(\tilde{d})}\frac{log(n)}{loglog(n)}. \end{aligned}$$

Entonces, la sucesión de grados libre de escala con $\beta = 4$ es admisible.

Para probar la condición de admisibilidad especial escogemos a $y = 4$. Entonces $U = U_y$ satisface que

$$\begin{aligned} Vol_2(U) &= d^2\frac{(\beta-2)^2}{(\beta-1)(\beta-3)}n(1-\frac{1}{4}+o(1)) \\ &= (\frac{3}{4}+o(1))Vol_2(G) \\ &\approx dVol_1(G). \end{aligned}$$

$$\begin{aligned} \frac{Vol_3(U)}{Vol_2(U)} &= (1+o(1))d\frac{8}{27}d\log(4) \\ &= O(\frac{\tilde{d}}{\log(\tilde{d})}). \end{aligned}$$

Como el grado promedio está acotado superiormente por una constante se concluye que la sucesión es especialmente admisible.

Caso 3: $3 < \beta < 4$. Para probar que la sucesión es admisible escogemos a $y = \frac{\log(n)}{\log(d)\log(\log(n))}$. Entonces, $U = U_y$ satisface que

$$Vol_2(U) = d^2 \frac{(\beta-2)^2 + o(1)}{(\beta-1)(\beta-3)} n = (1+o(1))Vol_2(G)$$

$$\begin{aligned} \frac{Vol_3(U)}{Vol_2(U)} &= (1+o(1))d\frac{(\beta-2)(\beta-3)}{(\beta-1)(4-\beta)} y^{1/(4-\beta)} \\ &= O(\frac{\tilde{d}}{\log(\tilde{d})}) \end{aligned}$$

Por lo tanto, la sucesión es admisible.

Para probar la condición de ser especialmente admisible escogemos a $y = (\beta-2)^{\frac{2}{(\beta-3)}}$. Entonces $U = U_y$ satisface que

$$\begin{aligned} Vol_2(U) &= d^2 \frac{(\beta-2)^2}{(\beta-1)(\beta-3)} n(1 - \frac{1}{(\beta-2)^2} + o(1)). \\ &= (d+o(1))Vol(G) \end{aligned}$$

$$\begin{aligned} \frac{Vol_3(U)}{Vol_2(U)} &= (1+o(1))d\frac{(\beta-2)(\beta-3)}{(\beta-1)(4-\beta)} y^{1/(4-\beta)}. \\ &= O(\frac{\tilde{d}}{\log(\tilde{d})}). \end{aligned}$$

Entonces la sucesión es especialmente admisible y la prueba está completa.

□

El siguiente teorema es el último que demostraremos en este capítulo, éste y el anterior son fundamentales para llegar a las conclusiones del capítulo 4.

Definición 2.2. *El núcleo de una gráfica aleatoria libre de escala con exponente β es el conjunto S_t tal que*

$$S_t = \{v_i \| \omega_i \geq t = n^{\frac{1}{\log(\log(n))}}\}$$

Teorema 2.4. *Suponga que una gráfica aleatoria libre de escala con exponente β tiene grado promedio mayor estricto que 1 y un grado máximo m que satisface que $o(\log(m)) = \log(n)/\log(\log(n))$. Si $2 < \beta < 3$, casi seguramente el diámetro es $\Theta(\log(n))$ y la distancia promedio es a lo más del $O(\log(\log(n)))$.*

Para el caso en el que $\beta = 3$, la gráfica aleatoria libre de escala tiene diámetro casi seguramente de $\Theta(\log(n))$ y la distancia promedio es $\Theta(\log(n)/\log(\log(m)))$.

Prueba: Para demostrar el teorema para el caso en el que $2 < \beta < 3$ probaremos una serie de afirmaciones.

Afirmación 2.1. *El diámetro del núcleo es casi seguramente del $O(\log(\log(n)))$*

Lo anterior se sigue del hecho de que el núcleo contiene una gráfica Erdos-Renyi $G(n', p)$ donde $n' = ct^{\beta-1}n$ y $p = \frac{t^2}{Vol(G)}$, esto último ya que

$$p_{ij} = \frac{\omega_i \omega_j}{Vol(G)} \geq \frac{t^2}{Vol(G)} = p$$

De [2] la subgráfica es casi seguramente conexa. Usando un resultado en [25] el diámetro de la subgráfica es a lo más $\frac{\log(n')}{\log(pn')}$. Pero

$$\frac{\log(n')}{\log(pn')} = \frac{\log(cnt^{\beta-1})}{\log(\frac{t^2}{Vol(G)}ct^{\beta-1}n)} = \frac{\log(ct^{\beta-1}) + \log(n)}{\log(\frac{t^{\beta-3}cn}{Vol(G)})} = \frac{\log(ct^{\beta-1}) + \log(n)}{(\beta-3)\log(t) + \log(\frac{cn}{Vol(G)})}$$

$$\leq \frac{\log(ct^{\beta-1}) + \log(n)}{(\beta-3)\log(t)} = (1 + \frac{\log(ct^{\beta-1})}{\log(n)})(\frac{\log(n)}{(\beta-3)\log(t)}).$$

Pero tenemos que $\frac{\log(ct^{\beta-1})}{\log(n)} = o(1)$ entonces el diámetro de la gráfica $G(n', p)$ está acotado por

$$D(G(n', p)) \leq (1 + o(1))(\tfrac{\log(n)}{(\beta-3)\log(t)}).$$

Además, para $k \geq (1 + o(1))(\frac{1}{(\beta-3)})$ se tiene que

$$(1 + o(1))(\frac{1}{\beta-3}) \leq k$$

$$(1 + o(1))(\frac{1}{(\beta-3)})\frac{\log(n)}{\frac{\log(log(n))\log(n)}{\log(log(n))}} \leq k$$

$$(1+o(1))(\frac{1}{(\beta-3)})\frac{\log(n)}{\log(log(n))\log(t)} \leq k$$

$$D(G(n',p)) \leq (1+o(1))(\frac{1}{(\beta-3)})\frac{\log(n)}{\log(t)} \leq k\log(log(n))$$

Podemos concluir que $D(G(n',p)) = O(\log(log(n)))$.

Afirmación 2.2. *Casi todos los vértices con grado al menos* $\log(n)$ *están casi seguramente dentro de una distancia de* $O(\log(log(n)))$ *del núcleo*

Para ver lo anterior, comenzamos con un vértice u_0 con grado esperado $\omega_0 \geq \log(n)$. Aplicamos el lema 1.10 con $S = u_0$, $T = u_1$ y $c = \frac{\log(n)}{Vol(G)}$ donde u_0 es el vértice con grado esperado ω_0 y u_1 es el vértice con grado esperado $\omega_1 \geq (\frac{\omega_0}{\log(n)})^{\frac{1}{(\beta-2)^s}}$ (s es una constante que definiremos más adelante), concluimos que con probabilidad de al menos $1-e^{-c}$ u_0 es vecino de u_1. Repitiendo este proceso encontramos un camino $u_0, u_1, ..., u_s$ teniendo a u_s como el vértice con grado $\omega_s \geq (\frac{\omega_0}{log(n)})^{\frac{1}{(\beta-2)}}$ con probabilidad de al menos $(1-e^{-c})^s$ donde s es una constante que satisface que $\log(\omega_s) \geq \frac{\log(n)}{\log(\log(n))} \geq \log(\frac{\omega_0}{\log(n)})^{\frac{1}{(\beta-2)^s}}$.

Ahora sólo falta probar que $s = O(\log(log(n)))$. Lo anterior se deriva del hecho que

$$\frac{\log(n)}{\log(\log(n))} \geq \log((\frac{\omega_0}{\log(n)})^{\frac{1}{(\beta-2)^s}})$$

$$\frac{\log(n)}{\log(\log(n))} \geq (\frac{1}{(\beta-2)^s})\log(\frac{\omega_0}{\log(n)}) \geq (\frac{1}{(\beta-2)^s})$$

$$\frac{\log(n)}{\log(\log(n))} \geq (\frac{1}{(\beta-2)})^s$$

$$\log(\log(n)) - \log(\log(\log(n))) \geq s\log(\frac{1}{(\beta-2)})$$

Pero $\log(\log(n)) \geq \log(\log(n)) - \log(\log(log(n)))$; por lo tanto

$$\log(\log(n)) \geq s\log(\frac{1}{(\beta-2)})$$

$$M\log(\log(n)) \geq s$$

donde $M = (\log(\frac{1}{(\beta-2)}))^{-1}$.

Concluimos que $s = O(\log(\log(n)))$

Afirmación 2.3. *Para vada vértice u en el componente gigante, con probabilidad* $1 - o(1)$, *v está dentro de una distancia de* $O(\log(\log(n)))$ *de un vértice de grado al menos* $\log(n)$

Para probar la afirmación anterior notamos que de (1.15) y de la definición de grado esperado en una gráfica aleatoria libre de escala se concluye que el grado esperado mínimo $\omega_{min} = d\frac{(\beta-2)}{(\beta-1)}$ con $d > 1$.

Sea S la i-ésima vecindad de u, y sea $T = (\omega_{min}, a)$ el conjunto de vértices con pesos entre ω_{min} y $a\omega_{min}$ donde a es una constante que será escogida más tarde (es una constante relativamente grande).

Tenemos

$$Vol(T) = \sum_{\omega_{min}}^{a\omega_{min}} \omega \approx \int_{\omega_{min}}^{a\omega_{min}} \omega dw = c\int_{i_i}^{i_f} i^{-\frac{1}{(\beta-1)}} di$$

donde $i_i = (\frac{c}{d}\frac{(\beta-1)}{\beta-2})^{(\beta-1)}$ e $i_f = (\frac{c}{da}\frac{(\beta-1)}{(\beta-2)})^{(\beta-1)}$.

Entonces

$$Vol(T) \approx c(\frac{\beta-1}{\beta-2})(i_f^{\frac{(\beta-2)}{(\beta-1)}} - i_i^{\frac{(\beta-2)}{(\beta-1)}}) = c\frac{(\beta-1)}{(\beta-2)}((\frac{c}{ad}\frac{(\beta-1)}{(\beta-2)})^{\beta-2} - (\frac{c}{d}\frac{(\beta-1)}{(\beta-2)})^{\beta-2})$$
$$= dn^{\frac{1}{\beta-1}}((a^{-1}n^{\frac{1}{\beta-1}})^{\beta-2} - n^{\frac{\beta-2}{\beta-1}}) = dn(a^{2-\beta} - 1)$$

Por otro lado

$$Vol_2(T) = \sum_{\omega_{min}}^{a\omega_{min}} \omega^2 \approx \int_{\omega_{min}}^{a\omega_{min}} \omega^2 dw = c^2\int_{i_i}^{i_f} i^{-\frac{2}{\beta-1}} di = c^2(\frac{\beta-1}{\beta-3}) i^{\frac{\beta-3}{\beta-1}}|_{i_i}^{i_f}$$
$$= c^2(\frac{\beta-1}{\beta-3})(i_f^{\frac{\beta-3}{\beta-1}} - i_i^{\frac{\beta-3}{\beta-1}}) = c^2(\frac{\beta-1}{\beta-3})((\frac{c}{ad}\frac{(\beta-1)}{(\beta-2)})^{\beta-3} - (\frac{c}{d}\frac{(\beta-1)}{(\beta-2)}^{\beta-3}))$$
$$= \frac{(\beta-2)^2}{(\beta-1)^2} d^2 n^{\frac{2}{\beta-1}}(\frac{\beta-1}{\beta-3})((a^{-1}n^{\frac{1}{\beta-1}})^{\beta-3} - n^{\frac{\beta-3}{\beta-1}}) =$$
$$= \frac{(\beta-2)^2}{(\beta-3)(\beta-1)} d^2 n(a^{3-\beta} - 1)$$

Por último

$$Vol_3(T) = \sum_{\omega_{min}}^{a\omega_{min}} \omega^3 \approx \int_{\omega_{min}}^{a\omega_{min}} \omega^3 d\omega = c^3\int_{i_i}^{i_f} i^{-\frac{3}{\beta-1}} di = c^3(\frac{\beta-1}{\beta-4}) i^{\frac{\beta-4}{\beta-1}}|_{i_i}^{i_f}$$

$$= c^3(\frac{\beta-1}{\beta-4})(i_f^{\frac{\beta-4}{\beta-1}} - i_i^{\frac{\beta-4}{\beta-1}}) = c^3(\frac{\beta-1}{\beta-4})((\frac{c}{ad}\frac{(\beta-1)}{(\beta-2)})^{\beta-4} - (\frac{c}{d}\frac{(\beta-1)}{(\beta-2)})^{\beta-4})$$
$$= \frac{(\beta-2)^3}{(\beta-1)^3}d^3 n^{\frac{3}{\beta-1}}(\frac{\beta-1}{\beta-4})((a^{-1}n^{\frac{1}{\beta-1}})^{\beta-4} - n^{\frac{\beta-4}{\beta-1}})$$
$$= \frac{(\beta-2)^3}{(\beta-1)^2(\beta-4)}d^3 n(a^{4-\beta}-1)$$

Para poder aplicar el lema 1.9 el $Vol(\Gamma(S))$ debe satisfacer

$$Vol(\Gamma(S)) \geq \frac{2c}{\epsilon}\frac{Vol_3(T)}{Vol_2^2(T)}Vol(G)$$
$$= \frac{2c}{\epsilon}(\frac{(\beta-2)^3 d^3 n}{(\beta-1)^2(\beta-4)}(a^{4-\beta}-1))((\frac{\beta-2}{\beta-1})dn^{\frac{1}{\beta-1}}(\frac{\beta-1}{\beta-2})n^{\frac{\beta-2}{\beta-1}})(\frac{(\beta-3)^2(\beta-1)^2}{(\beta-2)^4 d^4 n^2(a^{3-\beta}-1)})$$
$$= \frac{2c}{\epsilon}(\frac{(\beta-2)^3}{(\beta-1)^2(\beta-4)}(a^{\beta-4}-1))(\frac{(\beta-3)^2(\beta-1)^2}{(\beta-2)^4}(a^{3-\beta}-1)^2))$$
$$= \frac{2c}{\epsilon}\frac{(\beta-3)^2}{(\beta-2)(\beta-4)}\frac{(a^{4-\beta}-1)}{(a^{3-\beta}-1)} \approx \frac{2c}{\epsilon}\frac{(\beta-3)^2}{(\beta-2)(\beta-4)}a^{\beta-2}$$

y

$$Vol(\Gamma(S)) \leq \epsilon\frac{Vol_2(T)}{Vol_3(T)}$$
$$= \epsilon(\frac{(\beta-2)^2}{(\beta-1)(\beta-3)}d^2 n(a^{3-\beta}-1))(dn)(\frac{(\beta-1)^2(\beta-4)}{(\beta-2)^3 d^3 n(a^{4-\beta}-1)})$$
$$= \epsilon\frac{(\beta-1)(\beta-4)}{(\beta-2)(\beta-3)}n\frac{(a^{3-\beta}-1)}{(a^{4-\beta}-1)} \approx \epsilon\frac{(\beta-1)(\beta-4)}{(\beta-2)(\beta-3)a}n$$

Las dos ecuaciones anteriores son fáciles de satisfacer escogiendo valores apropiados de a, c y ϵ. Por ejemplo $a = c = logloglog(n)$ y $\epsilon = \frac{1}{4}$. Sea $\tau = logloglog(n)$, aplicando el lema 1.11 existe c_0 e $i_0 \leq c_0\tau$ tal que

$$Vol(\Gamma_{i_0}(u)) \geq \tau$$

con probabilidad $1 - o(n^{-1}) = 1 - o(1)$. Por el lema 1.9 con probabilidad $1 - e^{-c} = 1 - \frac{1}{loglog(n)}$, el volumen de $\Gamma_i(u)$ para $i > i_0$ crecerá a razón de

$$(1-2\epsilon)\frac{Vol_2(T)}{Vol(G)} = (\frac{1}{2})(\frac{(\beta-2)^2}{(\beta-3)(\beta-1)}d^2 n(a^{3-\beta}-1))(\frac{1}{dn}) \approx \frac{d(\beta-2)^2}{2(\beta-1)(\beta-3)}(a^{3-\beta})$$

si $\Gamma_i(u)$ tiene un volumen no muy grande, entonces, después de a lo más

$\frac{2loglog(n)}{(3-\beta)log(a)} = o(loglog(n))$ pasos llegaremos a un vértice con grado esperado de al menos $\log^2(n)$.

Aplicando el lema 1.10 esto implica que con un paso adicional podemos llegar a un vértice con grado al menos de $\log(n)$ con probabilidad de al menos $1 - e^{-\log^2(n)}$. El número total de pasos es

$$c_o\tau + o(\log(\log(n))) + 1 = o(\log(\log(n)))$$

La probabilidad de que u no esté conectado por un camino de longitud $o(\log(\log(n)))$ a un vértice con grado al menos de $\log(n)$ es

$$o(1) + o(\log(\log(n)))\frac{1}{\log(\log(n))} + e^{-\log^2(n)} = o(1)$$

Afirmación 2.4. *Para cada vértice u en el componente gigante, con probabilidad de* $1-o(1)$*, u está dentro de una distancia del* $O(log(n))$ *de un vértice con grado al menos del* $O(\log(n))$*. Entonces, con probabilidad de* $1-o(1)$*, el diámetro es* $\Theta(\log(n))$*.*

Para probar lo anterior hacemos un procedimiento análogo a la demostración de la afirmación pasada. Aplicamos el lema 1.10 escogiendo a $a = 100$, $c = 3\log(n)$, $\epsilon = \frac{1}{4}$ y $\tau = (96)(\frac{(\beta-3)^3}{(\beta-2)(4-\beta)}100^{\beta-2}\log(n))$. Con el mismo argumento concluimos que la probabilidad de que u no esté conectado por un camino de longitud del $O(\log(n))$ a un vértice de grado $O(\log(n))$ es

$$c_0\tau^{\frac{3}{2}} + O(\log(\log(n)))e^{-3\log(n)} + e^{-\log^2(n)} = o(\frac{1}{n^2})$$

El número total de pasos es a lo más

$$c_0\tau + O(\log(\log(n))) + 1 = O(\log(n))$$

Por lo tanto, con probabilidad $1 - o(n^{-2})$ se cumple la primera parte de la afirmación.

De la primera parte de la afirmación hemos derivado una cota superior para el diámetro de la gráfica; $O(\log(n))$.

Ahora estableceremos una cota inferior de orden de $\log(n)$. Recordemos que $\omega_{min} = d\frac{(\beta-2)}{(\beta-1)}$; consideraremos todos los vértices con grado menor a d. Existen $(1-(\frac{\beta-2}{\beta-1})^{\beta-1})n$ de estos vértices. Para cada vértice u, la probabilidad de que u tenga un único vecino con peso menor a d es:

$$\begin{aligned}\textstyle\sum_{\omega_v<d} \quad & \omega_u\omega_v\rho\prod_{j\neq v}(1-\omega_u\omega_j\rho) \\ & \approx \omega_u Vol(S(\omega_{min},\frac{\beta-1}{\beta-2}))\rho e^{-\omega_u} \\ & \approx (1-(\frac{\beta-2}{\beta-1})^{\beta-1})\omega_u e^{-\omega_u}\end{aligned}$$

Note que la probabilidad anterior es más grande que una constante e. Entonces con probabilidad $e.\log(n)$ encontramos un camino de longitud $\log(n)$ contenido en $S(\omega_{min},(\frac{\beta-1}{\beta-2}))$. Por lo tanto, el diámetro de la gráfica aleatoria G propuesta en el teorema es $\Theta(\log(n))$.

Combinando la afirmaciones 1, 2, y 3 derivamos una cota superior para la distancia promedio del $O(\log(\log(n)))$, por lo que el primer párrafo del teorema queda demostrado. (Se encuentra una mejor cota superior para la distancia promedio en [26] donde se hace un análisis mucho más cuidadoso y se concluye que tal cota es $(2+o(1))\frac{\log(\log(n))}{\log(\frac{1}{\beta-2})}$).

Ahora demostraremos el teorema para el caso en el que $\beta=3$.

Sea T el conjunto de vértices con pesos menores a t. Entonces tenemos que

$$\begin{aligned} Vol(T) &= \sum_{i=n(\frac{d}{2t})^2}^{n}\frac{d}{2}(\frac{i}{n})^{-\frac{1}{2}} \\ &\approx n\int_{(\frac{d}{2t})^2}^{n}\frac{d}{2}i^{-\frac{1}{2}}di \\ &= nd(1-\frac{d}{2t})\end{aligned}$$

$$\begin{aligned} Vol_2(T) &= \sum_{n(\frac{d}{2t})^2}^{n}\frac{d^2}{4}(\frac{i}{n})^{-1} \\ &\approx n\int_{(\frac{d}{2t})^2}^{n}\frac{d^2}{4}i^{-1}di \\ &= \frac{nd^2}{2}log\frac{2t}{d}\end{aligned}$$

$$\begin{aligned} Vol_3(T) &= \sum_{n(\frac{d}{2t})^2}^{n} \frac{d^3}{8}(\frac{i}{n})^{-\frac{3}{2}} \\ &\approx n\int_{(\frac{d}{2t})^2}^{n} \frac{d^3}{8} i^{-\frac{3}{2}} di \\ &= \frac{nd^3}{4}(\frac{2t}{d}-1) \\ &= \frac{nd^2}{2}(t-\frac{d}{2}) \end{aligned}$$

Entonces, para que se satisfaga el lema 1.9 se tiene que

$$\frac{2cVol_3(T)Vol(G)}{\epsilon^2 Vol_2^2(T)} \approx \frac{2c\frac{nd^2}{2}(t-\frac{d}{2})nd}{\epsilon^2(\frac{nd^2}{2}log(\frac{2t}{d})^2)} \approx \frac{2c(\frac{2t}{d}-1)}{\epsilon^2 log^2(\frac{2t}{d})}$$

Ahora tomaremos en cuenta la siguiente afirmación que es una consecuencia inmediata del lema 1.9

Afirmación 2.5. *Suponga que una gráfica aleatoria libre de escala con exponente $\beta = 3$ tiene grado promedio d. Para cualquier $\epsilon < \frac{1}{2}$, $c > 0$, y cualquier conjunto de vértices S que satisface que $Vol(S) > \frac{2c}{\epsilon^2}\frac{\frac{2t}{d}}{log^2(\frac{2t}{d})}$ y $Vol(S) \leq n^{\frac{2}{3}}$, satisface que*

$$Vol(\Gamma(S)) > (1-2\epsilon)\frac{d}{2}log(\frac{2t}{d})Vol(S)$$

con probabilidad de al menos $1 - e^{-c}$

Afirmación 2.6. *Para un vértice u en el componente gigante, con probabilidad de al menos $1 - \frac{1}{\log^2(n)}$, el volumen de $\Gamma_{i_1}(u)$ es al menos $\log^6(\log(n))$ para alguna $i_1 = O(\log^6(\log(n))$.*

Para probar lo anterior usamos el lema 1.11 con la elección de $\tau = \log^6(\log(n))$. Entonces, con probabilidad de al menos $1-\frac{1}{\log^2(n)}$ existe una constante C y un índice i_0 que satisface que $i_0 \leq C\log^6(\log(n))$ y $Vol(\Gamma_{i_0}(u)) \geq \log^6(\log(n))$.

Afirmación 2.7. *Con probabilidad de al menos $1-o(\frac{1}{\log^2(n)})$, un subconjunto S con $Vol(S) \geq \log^6(\log(n))$ tiene $Vol(\Gamma_i(S)) > m$ si $i > \frac{\log(m)}{\log(\log(m))}$*

Para demostrar la afirmación anterior aplicamos la afirmación 1.5 repetidamente. En cada paso escogemos a $c = \log^2 \log(n)$, $\epsilon = \frac{1}{\log(\log(n))}$ y $t = \frac{d\epsilon^2 a_i}{4c} \log^2(\frac{\epsilon^2 a_i}{2c})$ donde a_i se define recursivamente como sigue. Primero definimos a $a_0 \geq \log^6(\log(n))$. Para $i \geq 1$ definimos a $a_{i+1} = \frac{d}{10} a_i \log(a_i)$. Vemos que $a_{i+1} > a_i$ y $a_i \geq \log^6(\log(n))$. Probaremos por inducción que $Vol(\Gamma_i(S)) \geq a_i$. Por la afirmación 1.6 se vale para $i = 0$. Suponga que se vale para i. Verificamos las condiciones de la afirmación 1.5

$$\begin{aligned} \frac{\frac{2t}{d}}{\log^2(\frac{2t}{d})} &= \frac{\epsilon^2 a_i}{2c} \frac{\log^2(\frac{\epsilon^2 a_i}{2c})}{(\log(\frac{\epsilon^2 a_i}{2c}) + 2\log(\log(\frac{\epsilon^2 a_i}{2c})))^2} \\ &\leq \frac{\epsilon^2 a_i}{2c} \\ &\leq \frac{\epsilon^2 Vol(\Gamma_i(S))}{2c} \end{aligned}$$

Entonces podemos aplicar la afirmación 1.5

$$\begin{aligned} Vol(\Gamma_{i+1}(S)) &\geq (1-2\epsilon) a_i \frac{d}{2} log(\frac{2t}{d}) \\ &\geq \frac{d}{10} a_i \log(a_i) \\ &= a_{i+1} \end{aligned}$$

Ahora probaremos de manera inductiva que $a_{i+1} \geq (i+s)^{i+s}$ para $s = e^{\frac{10e}{d}}$. Asumiremos que $a_0 = \log^6(log(n)) \geq s^s$ ya que s está acotada. Para $i+1$, tenemos

$$\begin{aligned} a_{i+1} &\geq \frac{d}{10} a_i \log(a_i) \\ &\geq \frac{d}{10} (i+s)^{i+s} (i+s) \log(i+s) \\ &\approx (i+1+s)^{i+1+s} \frac{d}{10e} \log(i+s) \\ &> (i+1+s)^{i+1+s} \end{aligned}$$

Entonces hemos probado que $a_i \geq (i+s)^{i+s}$. Sea $i = \frac{\log(m)}{\log(\log(m)) - \log(\log(log(m)))} - s = (1+o(1)) \frac{\log(m)}{\log(\log(m))}$. Entonces

$$\begin{aligned} a_i &\geq (i+s)^{i+s} \\ &\geq m \end{aligned}$$

Afirmación 2.8. *Con probabilidad de al menos* $1 - o(\frac{1}{log^2(n)})$*, un subconjunto* S *con* $Vol(S) \geq m$ *satisface que* $Vol(\Gamma_i(S)) > \sqrt{n}\log(n)$ *si* $i > \frac{(1+o(1))(\log(\sqrt{n})-\log(m))}{\log(\frac{d}{2}\log(m))}$.

Aplicaremos la afirmación 1.5 escogiendo a $c = \log^2(\log(n))$, $\epsilon = \frac{1}{\log(\log(n))}$ y $t = m$. Las condiciones para poder aplicar la afirmación 1.5 son verificadas a continuación

$$\begin{aligned} Vol(S) &\geq m \\ &\frac{2c}{\epsilon^2}\frac{\frac{2m}{d}}{\log(\frac{2m}{d})} \end{aligned}$$

Aquí usaremos el hecho de que $m > n^\delta$. Con probabilidad $1 - O(\frac{1}{\log^2(\log(n))})$ el volumen de $\Gamma_i(S)$ crece a razón de $(1-2\epsilon)\tilde{d}$ como i se incrementa. La afirmación queda probada.

Por el lema 1.10, casi seguramente la distancia entre dos conjuntos con peso mayor a $\sqrt{n}\log(n)$ es a lo más 1. De las afirmaciones 1.5, 1.6, 1.7 y 1.8, casi seguramente la distancia entre dos vértices u y v en el componente gigante será

$$2(O(\log^6(\log(n)))) + (1+o(1))\frac{\log(m)}{\log(\log(m))} +$$

$$\frac{(1+o(1))(\log(\sqrt{n})-\log(m))}{\log(\frac{d}{2}\log(m))} = \Theta(\frac{\log(n)}{\log(\log(m))})$$

Hemos demostrado una parte del caso $\beta = 3$.

Ahora derivaremos la cota superior para el diámetro, necesitaremos las siguientes afirmaciones.

Afirmación 2.9. *Para cualquier vértice* u *en el componente gigante, con probabilidad de al menos* $1-\frac{1}{n^3}$*, el volumen de* $\Gamma_{i_0}(u)$ *es al menos de* $8e^4\log(n)$ *para alguna* $i_0 = O(\log(n))$.

Escogemos a $\tau = 8e^4\log(n)$ y aplicamos el lema 1.11. Entonces existe una constante C y un índice $i_0 \leq 8Ce^4\log(n)$ y $Vol(\Gamma_{i_0}) \geq 8e^4\log(n)$ con probabilidad de al menos $1 - n^{-3}$

Afirmación 2.10. *Con probabilidad de al menos* $1 - n^{-3}$*, cualquier subconjunto* S *con* $Vol(S) \geq 8e^4 \log(n)$ *satisface que* $Vol_i(S) > \sqrt{n}\log(n)$ *si* $i > \frac{\log(n)}{2\log(d)}$.

Para probar la afirmación 1.5 escogiendo a $c = 4\log(n)$, $\epsilon = \frac{1}{4}$ y $t = \frac{e^4 d}{2}$. Notamos que

$$\frac{2c}{\epsilon^2}\frac{\frac{2t}{d}}{\log^2(\frac{2t}{d})} = 8e^4\log(n)$$

Por la afirmación 1.5 con probabilidad $1 - n^{-4}$ en cada paso, el volumen de la i-ésima vecindad de S crece a razón de

$$(1 - 2\epsilon)\frac{d}{2}\log(\frac{2t}{d}) = d$$

si el volumen de $\Gamma_i(S)$ es $O(\sqrt{n})$. La afirmación está probada.

Por las afirmaciones anteriores, con probabilidad de al menos $1 - \frac{1}{n}$, para cada par de vértices u y v en el componente gigante , la distancia entre u y v es a lo más

$$2(O(\log(n)) + \frac{\log(n)}{2\log(d)}) + 1 = O(\log(n))$$

La cota inferior para el diámetro se calcula de manera idéntica que para el caso $2 < \beta < 3$.

Por lo tanto, cuando $\beta = 3$ el diámetro es del $O(\log(n))$.

El teorema está probado.

$\square$

Capítulo 3

Gráficas locales

Este capítulo es breve ya que sólo se define el concepto de gráfica local. No por la brevedad es menos importante ya que estas definiciones y teorema aquí demostrado serán de suma importancia para definir y entender el concepto de gráfica híbrida.

Definición 3.1. *Una gráfica local es altamente conexa. Para ser precisos, usaré dos parámetros para describir la conectividad local. Para cualesquiera dos enteros fijos $k \geq 2$ y $l \geq 2$, una gráfica L es llamada localmente (k,l)-conexa si para cualquier arista uv, existe al menos l caminos ajenos por aristas (es decir, ningún par de caminos tienen una arista en común) con longitud a lo más k que unen a u y a v (incluyendo la arista uv).*

Por definición, la union de dos gráficas localmente (k,l)-conexas es localmente (k,l)-conexa. La máxima subgráfica (k,l)-conexa, H, es la union de todas las subgráficas localmente (k,l)-conexas de G. Entonces, para cualquier gráfica G, la máxima subgráfica localmente conexa es única. Remarco que una gráfica (k,l)-conexa no necesariamente es conexa. Por ejemplo, la unión disjunta de dos gráficas (k,l)-conexa es todavía (k,l)-conexa.

Ahora mostraremos un algoritmo simple para encontrar la máxima subgráfica localmente (k,l)-conexa.

Algoritmo(k,l)**:**

Para cada arista $e = uv$, checar si hay l caminos disjuntos por aristas con longitud a lo más k que unen a u y v en la gráfica G. Si no, elimina la arista e de G. Se sigue con este procedimiento hasta que ningún arista tenga que ser removido.

Teorema 3.1. *Para cualquier gráfica G, el algoritmo(k,l), encuentra la única subgráfica máxima localmente (k,l)-conexa sin importar el orden en que se escogieron las aristas.*

Prueba: Sea H' una gráfica producto del algoritmo (k, l), donde el orden de las aristas que se fue removiendo es arbitrario. Supongamos que ésta es máxima, entonces es suficiente mostrar que $H = H'$.

Sea $G = H_0 \supset H_1 \supset H_2 \supset ... \supset H_r = H'$ la sucesión de las subgráficas intermedias que produce en cada paso el algoritmo (k, l). Probaré que $H \subset H_i$ para todo $i = 0, 1, 2, ..., r$ por inducción sobre i. Es fácil verlo para $i = 0$ ya que $H \subset H_0 = G$. Ahora supondremos que $H \subset H_i$. Para $i+1$, sea $e_{i+1} = uv$, la arista que se remueve en el paso $i + 1$. Es suficiente mostrar que uv no es una arista de H. De otra manera hay l caminos disjuntos de H uniendo a u y v. Como $H \subset H_i$, existen caminos en H que también son caminos en H_i. De acuerdo con el algoritmo, este no puede ser removido, lo cual es una contradicción. Entonces, tenemos que $H \subset H_{i+1}$ y $H \subset H_r = H'$.

En la otra dirección, como H' es localmente (k, l)-conexa, tenemos que $H' \subset H$. H es la máxima gráfica con esta propiedad. La prueba está completa. $\square$

Capítulo 4

Gráficas híbridas libres de escala

4.1. El modelo de gráfica híbrida libre de escala

Si una gráfica es libre de escala, entonces el grado promedio y la conectividad (es decir, la distribución de sus componentes conexas) estarán completamente determinadas por el exponente de la ley de potencia . Sin embargo, para la mayoría de las gráficas reales, la ley de potencia se sigue sólo para cierto rango de grados, a saber, para grados que no son muy pequeños y no muy grandes. Consideraremos el siguiente modelo con la condición de que la mayoría de los ejemplos de gráficas masivas satisfacen una ley de potencias con exponente $\beta > 2$

Modelo $M(n, \beta, d, m)$ donde

•n es el número de vértices,

•$\beta > 2$ es la potencia de la ley de potencia,

•d es el grado promedio esperado,

•m es el grado esperado máximo (o la cota superior para el rango de grados que obedecen la ley de potencia) y $m^2 = o(nd)$.

Consideraremos, como ya lo hicimos en el capítulo 2, que el vértice $i - i_0 + 1$, v_i, tiene grado esperado:

$$\omega_i = ci^{-\frac{1}{\beta-1}}$$

para $i_0 \leq i < n + i_0$. Donde, como vimos en el capítulo 2, c depende de el grado promedio d e i_0 depende del máximo grado esperado m. Es fácil ver que el número de vértices de grado esperado dado entre k y $k+1$ es del orden de $c'k^{-\beta}$ donde $c' = c^{\beta-1}(\beta-1)$ como se requiere para la ley de potencias. Consideraremos las siguientes expresiones que ya demostré en el capítulo 2:

$$c = \tfrac{\beta-2}{\beta-1} d n^{\frac{1}{\beta-1}}$$
$$i_0 = n\left(\tfrac{d(\beta-2)}{m(\beta-1)}\right)^{\beta-1}$$

Por otro lado, sea $f(x) = \frac{\beta-2}{\beta-1} d x^{-\frac{1}{\beta-1}}$. Los grados esperados (o pesos) son $f(\frac{i}{n})$, $i_0 \leq i \leq n$.

También se considerará un modelo alternativo $M'(n, \beta, d, m)$, en el cual a cada vértice x se le asigna un peso $f(y)$, donde y es un número real que se escoge de manera uniforme del rango (i_0, n). Se puede demostrar fácilmente que los dos modelos son equivalentes (para $i_0 \ll n$) en el sentido de que cada propiedad que se vale para una gráfica aleatoria en M casi seguramente se vale para la misma gráfica en M' y viceversa.

Una gráfica híbrida consiste en dos partes, una gráfica global y una gráfica local. El conjunto de aristas de una gráfica híbrida es la unión ajena de conjuntos de aristas de la gráfica global y de la gráfica local. El conjunto de nodos se comparte entre las dos gráficas. Los parámetros relacionados son:

•β, el exponente de la ley de potencia.

•d, el grado promedio

•m, el grado esperado máximo (o la cota superior del rango de grados que obedecen la ley de potencias)

•L, la gráfica local.

Es importante tomar en cuenta que para una gráfica dada, todos los parámetros son sencillos de calcular y estimar. Entonces, es muy fácil construir una simulación de una gráfica con los parámetros dados.

La gráfica híbrida $H(n, \beta, d, m, L)$:

La gráfica local L es localmente (k, l)-conexa con grados acotados. El vértice v_i de H tiene peso ω_i donde $\omega_1, \omega_2, ..., \omega_n$, satisfacen una distribución de ley de potencias con exponente $\beta > 2$ usando el modelo $M'(n, \beta, d, m)$. También se asumirá que $d \geq 1$.

Para cualesquiera dos puntos u y v, la probabilidad de que exista una arista entre ellos se denota por $\mathrm{p}(u, v)$, que se define como

$$p(u,v) = \begin{cases} 1 & \text{si } uv \text{ es una arista de } L \\ \omega_u \omega_v \rho & \text{en otro caso} \end{cases}$$

Se verá que la gráfica local es fuerte en el sentido de que puede ser extraida y recuperada de la gráfica híbrida casi en su totalidad.

Teorema 4.1. *Para cualesquiera constantes fijas $M, k \geq 2$, y $l \geq 3$, suponga que L es una gráfica (k,l)-conexa con grados acotados por M. Sea L' una subgráfica máxima (k,l)-conexa en la gráfica híbrida $H(n, \beta, d, m, L)$ con $m = o(n^{\frac{1}{2}-\frac{1}{l}})$. Entonces L' satisface que:*

1) $L \subset L'$. El número esperado de aristas en $L' \backslash L$ es pequeo, es decir, $e(L') - e(L) = O(m)$.

2) El grado de los vértices en L' están casi seguramente acotado por arriba por $M + \lceil \frac{l}{2} \rceil - 1$.

3) El diámetro $D(L')$ de L' es a lo más $D(L)$.

Prueba: De las definiciones concluimos que $L \subset L'$. Para demostrar la propiedad 1) consideraremos aristas en L' que no están en L, a éstas las llamaremos aristas sobrevivientes. Llamaremos a la distancia entre dos vértices en L la distancia local, denotada por d_L. La vecindad de un vértice en L se le llama vecindad local. La i-ésima vecindad de v consiste en todos los vértices dentro de una distancia local no mayor a i de v.

Ahora probaremos el siguiente enunciado para probar fácilmente 1).

Casi seguramente todas las aristas sobrevivientes uv tienen vértices extremos con distancias locales $d_L(u, v)$ a lo más k.

Para probar lo anterior supongamos lo contrario, es decir, que u y v son dos vértices con $d_L(u, v) > k$, entonces cualquier trayectoria de longitud a lo más k en L' de u a v tiene que contener al menos una arista sobreviviente (si esto no ocurriera entonces $d_L(u, v) = k$). Como esta arista uv sobrevive después de que el algoritmo terminó, existen al menos l aristas en L' de la i-ésima vecindad local de u a la j-ésima vecindad local de v, con $i + j = k - 1$. Como los grados locales están acotados por M, el número de vértices en la i-ésima vecindad local de u es a lo más

$$\sum_{s=0}^{i} M^s = \frac{M^{i+1}}{M-1} \leq 2M^i$$

De manera similar, el número de vértices en la j-ésima vecindad local de v es a lo más $2M^j$. Hay a lo más $2M^i x 2Mj = 4Mk - 1$ pares de tales vértices. Para cada par la probabilidad de ser escogidos aleatoriamente para la gráfica híbrida es menor o igual a $m^2\rho$. Entonces la probabilidad de que uv sobreviva

es a lo más
$$\binom{4M^{k-1}}{l}(m^2\rho)^l = o(\frac{1}{n^2})$$
Esta probabilidad es muy pequeña, por lo tanto, casi seguramente, todas las aristas que sobreviven tienen $d_L(u,v) \leq k$.

Por lo tanto la proposición es cierta.

Ahora consideraremos el número esperado de aristas sobrevivientes, las cuales tienen puntos extremos con distancias locales en L a lo más k. Escogemos un vértice u. Hay a lo más $2M^k$ vértices con distancia local a lo más k de u. El número esperado de aristas sobrevivientes uv es a lo más $\sum_{u\in H} 2M^k\omega_u m\rho = 2M^k m = O(m)$.

Como los vértices en L tienen grado a los más M; para calcular la cota superior de los grados de los vértices en L' se tiene que primero calcular la cota superior de el grado de un vértice que es extremo de una arista sobreviviente. Supongamos que tal vértice es u y calcularemos el número de aristas uv sobrevivientes. Como seguramente v está a una distancia local de u de a lo más k (quitando la arista uv), existen a lo más $2M^k$ v's posibles. La probabilidad de que existan igual o más de $\lceil \frac{l}{2} \rceil$ v's con uv sobreviviente es a lo más
$$\binom{2M^k}{\lceil \frac{l}{2} \rceil}(m^2\rho)^{\lceil \frac{l}{2} \rceil} = o(\frac{1}{n})$$
Por lo tanto existen a lo más $\lceil \frac{l}{2} \rceil - 1$ v's con esta propiedad, entonces si u está en L tiene grado a lo más $M + \lceil \frac{l}{2} \rceil - 1$.

El enunciado 2) queda demostrado.

Sea $g(n)$ una función que crece muy lentamente y cuando n crece se aproxima a infinito. existen a lo más $4M^{2g(n)}$ pares de vértices con una distancia local a lo más $g(n)$ de un vértice dado u. La probabilidad de que existan l aristas sobrevivientes con distancia local de a lo más $g(n)$ de un vértice dado u es a lo más
$$\binom{4M^{2g(n)}}{l}(m^2\rho)^l = o(\frac{1}{n^2})$$
para alguna función que crece lentamente $g(n) = o(log\frac{n^{\frac{1}{2}-\frac{1}{l}}}{m})$. Casi seguramente, para todo vértice u, existen a lo más $l-1$ aristas sobrevivientes con distancia local a lo más k de u.

Sea (u,v) un par de vértices con $d_L(u,v) = D(L)$. La distancia entre u y v en la gráfica híbrida puede ser reducida (de la distancia local de u y v) por aristas sobrevivientes. Cada arista sobreviviente puede reducir la distancia de u y v en a los más $k-1$. Por lo tanto $D(L') \leq D(L)$.

El enunciado 3) queda demostrado.

$\square$

4.2. El fenómeno de mundo pequeño en gráficas híbridas libres de escala

La mayoría de las gráficas locales tienen diámetros grandes y distancias promedio también numerosa. Sin embargo, con los *hiperlinks* adicionales (es decir, aristas de la gráfica global libre de escala), la distancia promedio de la gráfica híbrida puede reducirce significativamente.

En una gráfica híbrida H, G denota a la gráfica global libre de escala como se definió en la sección anterior. Sea $\mathbf{w} = (\omega_1, \omega_2, ..., \omega_n)$ la sucesión de grados esperados de G. Diremos que el vértice v_i tiene peso ω_i y diremos que para un subconjunto S de vértices, tenemos que $Vol(S) = \sum_{v_i \in S} \omega_i$ y $Vol(G) = \sum \omega_i$. También para $k \geq 1$, tenemos que $Vol_k(S) = \sum_{v_i \in S} \omega_i^k$. En particular, el grado promedio de segundo orden $\tilde{d}$ es justamente $Vol_2(G)/Vol(G)$. Los siguientes teoremas se siguen inmediatamente de los teoremas demostrados en el capítulo 1.

Teorema 4.2. *Para una gráfica híbrida $H(n, \beta, d, m, L)$ con $\beta > 3$, casi seguramente, la distancia promedio es $(1+o(1))\frac{log(n)}{log(\tilde{d})}$ y el diámetro es $O(log(n))$*

Prueba: Se sigue inmediatamente del teorema 1.5, nada más que el diámetro disminuye en magnitud por las aristas agregadas por la gráfica local, por lo que el diámetro no necesariamente sigue la cota inferior pero sigue manteniendo la cota superior por lo que éste será del $O(log(n))$

$\square$

Teorema 4.3. *Para una gráfica híbrida $H(n, \beta, d, m, L)$ con $2 < \beta < 3$, casi seguramente la distancia promedio es $O(loglog(n))$ y el diámetro es $O(log(n))$.*

Para una gráfica híbrida $H(n, \beta, d, m, L)$ con $\beta = 3$, casi seguramente la distancia promedio es $O(log(n)/loglog(n))$ y el diámetro es $O(logn)$.

Prueba: Se sigue inmediatamente del teorema 1.6, pero pasa lo del teorema anterior, las cotas superiores para el diámetro y las distancias promedio siguen siendo válidas pero no necesariamente las cotas inferiores.

$\square$

Para el rango $2 < \beta < 3$, las gráficas libres de escala están presentes en muchas redes reales. Podemos reducir el diámetro si se dan algunas condiciones adicionales.

Definición 4.1. *Una gráfica local se dice que tiene dimensión isoperimétrica δ si para cada vértice v en L y cada entero $k < (loglogn)^{\frac{1}{\delta}}$, hay al menos k^{δ} vértices en L a distancia k de v.*

Teorema 4.4. *En una gráfica híbrida $H(n, \beta, d, m, L)$ con $2 < \beta < 3$ suponga que la gráfica local tiene dimensión isoperimétrica δ. Entonces casi seguramente el diámetro es $O((logn)^{\frac{1}{\delta}})$*

Prueba: Sabemos que en la gráfica global el diámetro es del $O(log(n))$, pero si tomamos en cuenta la distancia local, tenemos que cualesquiera dos vértices están a distancia a lo más $(loglog(n))^{\frac{1}{\delta}}$ entonces la distancia de estos dos vértices estará acotada por $(log(n))^{\frac{1}{\delta}}$. Como $(log(n))^{\frac{1}{\delta}} < log(n)$ se concluye que el diámetro se verá reducido con la introducción de la gráfica local, por lo que el diámetro será del $O((log(n))^{\frac{1}{\delta}})$.

$\square$

Conclusiones

En este trabajo se hizo el esfuerzo de conjuntar varios trabajos hechos anteriormente, ampliando las demostraciones y en algunos casos dando demostraciones propias. Ésto con el objetivo de tener una visión general y accesible a cualquier estudiante de la carrera de matemáticas sobre el tema tratado; no se necesita tener conocimientos previos de teoría de gráficas, ya que todas las definiciones y elementos para las demostraciones son proporcionados (exceptuando unos pocos).

En el capítulo 2 de este trabajo se pudo ver, en esencia, que las gráficas aleatorias con una sucesión de grados esperados libres de escala capturan uno de los aspectos fundamentales del efecto de mundo pequeño, la distancia promedio del orden de $log(n)$ si $\beta > 3$, de orden de $loglog(n)$ si $2 < \beta < 3$ y del orden de $\frac{log(n)}{loglog(n)}$ para $\beta = 3$, lo que significa que hay una trancisión en el caso en el que $\beta = 3$. El efecto de agrupamiento no es modelado satisfactoriamente por estas gráficas.

Por esa razón se inspira el capítulo 4, el hecho de que se añada una gráfica local que atrapa el efecto de agrupamiento nos garantiza que tal modelo captura el efecto de mundo pequeño en su totalidad.

Una de las preguntas que me hago y que desgraciadamente no se aborda en este trabajo es la adición de gráficas locales con otras características; tendríamos resultados más ricos?. Para cada gráfica real con efecto de mundo pequeño y libre de escala se tiene una gráfica local distinta que la modela?.

Evidentemente se deja como distintas líneas de trabajo el hecho de experimentar incluso con el concepto de gráfica híbrida con distintos modelos de gráfica local y global.

Un artículo sumamente interesante para ampliar el conocimiento del tema es [27] donde se encuentra un tipo de generalización del concepto de gráfica híbrida en el sentido en que se desarrollan resultados tomando en cuenta gráficas locales arbitrarias y gráficas globales arbitrarias, mostrando así que la unión de un componente global con una distancia promedio grande y un componente local con una distancia promedio grande nos da como resultado una gráfica híbrida con una distancia promedio de $O(log(n))$; mostrando además el efecto de agrupamiento, teniéndose así el efecto de mundo pequeño.

Bibliografía

[1] P. Erdös y T. Gallai, elöírt fokú pontokkal (Graphs with points of precribed degrees, en Húngaro), *Mat. Lapok* **11** (1961), 264-274.

[2] P. Erdös y A. Rényi, On random graphs I, *Publ. Math. Debrecen* **6** (1959), 290-291.

[3] D. West, Introduction to Graph Theory, Prentice Hall (1996).

[4] Fan Chung y Linyuan Lu, Connected components in random graphs with given expected degree sequences, *Internet publications* (2002), 5-7.

[5] www.math.ucsd.edu/r̃onspubs/82_09_fkg_inequality.pdf

[6] S. Milgram, The small world problem, Psichology Today **2** (1967), 60-67.

[7] L. A. Adamic y B. A. Huberman, Growth dynamics of the World Wide Web, *Nature* **401** (1999), 131.

[8] W. Aiello, F. Chung y L. Lu, A random graph model for massive graphs, *Proceedings of the Thirty-Second Annual ACM Symposium on Theory of Computing*, (2000), 171-180.

[9] R. B. R. Azevedo y A. M. Leroi, A power law for cells, *Proc. Natl. Acad. Sci. USA* **98** (2001), 5699-5704.

[10] A. L. Barábasi y R. Albert, Emergence of scaling in random networks, *Science* **286** (1999), 509-512.

[11] A. Barábasi, R. Albert, H. Jeong, Scale-free caracteristics of random networks: the topology of world wide web, *Physica A* **272** (1999), 173-187.

[12] A. Broder, R. Kumar, F. Maghoul, P. Raghavan, S. Rajagopalan, R. Stata, A. Tompkins y J. Wiener, Graph Structure in the Web, *Proceedings of the WWWW9 Conference* edición impresa en *Computer Networks* **33** (2000), 309-321.

[13] K. Calvert, M. Doar y E. Zegura, Modeling Internet Topology. *IEEE Communications Magazine* **35** (1997), 160-163.

[14] C. Cooper y A. Frieze, On a general model of web graphs, *Random Structures and Algorithms* **22** (2003), 311-335.

[15] M. Faloutsos, P. Faloutsos y C. Faloutsos, On power-law relationships of the Internet topology, *Proceeding of the ACM SIGCOM Conference* (1999).

[16] S. Jain y S. Krishna, A model for the emergence of cooperation, interdependence, and structure in evolving networks, *Pro. Natl. Acad. Sci. USA* **98** (2001), 543-547.

[17] J. Kleinberg, S. R. Kumar, P. Raghavan, S. Rajagopalan y A. Tomkins, The web as a graph, Measurements, models and methods, *Proceedings of the international Conference on Combinatorics and Computing* (1999).

[18] S. R. Kumar, P. Raghavan, S. Rajagopalan y A. Tomkins. Extracting large-scale knowledge bases from the web, *Proceedings of the 25th VLDB Conference* (1999).

[19] M. E. J. Newman, The structure of scientific collaborations networks, *Proc. Natl. Acad. Sci. USA* **98** (2001), 404-409.

[20] E. Zegura, K. Calvert y M. Danahoo, A quantitative comparison of graph based models for internet topology. *IEEE/ACM Transactions on Networking* **6** (1997), 770-783.

[21] D. J. Watts, Small Worlds- the Dynamics of Networks between Order and Randomness, (1999).

[22] D. J. Watts y S. H. Strogats, Collective dynamics of small world networks, *Nature* **393**, 440-442.

[23] A. Fabrikant, E. Koutsoupias y C. H. Papadimitriu, Heuristically optimized trade offs: a new paradigm for power laws in the internet, *Automata Languages and Programming* (2002), 110-122.

[24] J. Kleinberg, The small world phenomenon: An algorithmic perspective, *Proc. 32nd. ACM Symposium on Theory of Computing* (2000).

[25] Fan Chung y Linyuan Lu, The diameter of random Sparse Graphs,*Advances in Applied Math* **26** (2001) 257-279.

[26] Linyuan Lu, Probabilistic Methods in Massive Graphs and Internet Computing,*Disponible en http://math.ucsd.edu/l̃lu/thesis.pdf* (2002)

[27] Katharina A. Lehmann, Hendrik D. Post y Michael Kaufmann, Hybrid graphs as a framework for small-world effect, *Disponible en la red de manera gratuita* (2006)