UNIVERSIDAD DE TALCA
FACULTAD DE INGENIERÍA
ESCUELA DE INGENIERÍA CIVIL EN COMPUTACIÓN

Generación de Grafos inducidos por la Ley de Potencia

DIEGO ITURRIAGA PEÑA

Profesor Guía: RODRIGO PAREDES MORALEDA

Profesor Co-Guía: RENZO ANGLES ROJAS

Memoria para optar al título de
Ingeniero Civil en Computación

Curicó – Chile
Julio, 2020

CONSTANCIA

La Dirección del Sistema de Bibliotecas a través de su encargado Biblioteca Campus Curicó certifica que el autor del siguiente trabajo de titulación ha firmado su autorización para la reproducción en forma total o parcial e ilimitada del mismo.

UNIVERSIDAD DE TALCA
DIRECCIÓN
SISTEMA DE BIBLIOTECAS

SISTEMA DE BIBLIOTECAS

Curicó, 2022

"Sic Parvis Magna"
La grandeza nace de pequeños comienzos.
Lema de Sir Francis Drake.

AGRADECIMIENTOS

Si hace unos 10 años me hubieran preguntado si planeaba entrar a la universidad, seguramente me hubiera reído diciendo "no lo se, no creo", porque realmente lo veía como algo muy difícil; después de todo mi familia no tenía los medios económicos ni de cerca. Mi padre es un agricultor que ganaba y que sigue ganando el mínimo, y mi madre es una dueña de casa que también cosía y reparaba ropa de lo cual obtenían un ingreso, el cual no era mucho, pero era algo de dinero extra. Con esto no quiero decir que he tenido una mala vida, sino todo lo contrario. Una vida en donde mis padres me han dado lo justo y lo necesario. Pero sobretodo, rescato las lecciones que me han dado, en particular si quieres algo, si tienes una meta, un objetivo al que realmente quieres llegar, lo único que importa es que te esfuerces, que te esfuerces mucho, cada día, cada hora, cada segundo que puedas por lo que quieres alcanzar. Recuerdo a mi madre tardes completas cosiendo ropa hasta las tantas de la mañana para terminar, y en un par de horas levantándose para llevar a mi hermana y a mí al colegio, versus mi padre que se levantaba a las seis de la mañana para tomar su bicicleta y recorrer cerca de diez kilómetros para llegar al lugar donde trabajaba. Se esforzaban para ayudarse mutuamente, se esforzaban para ayudar a mi hermana y a mí constantemente. Puede sonar simple el solo tener que esforzarse, pero realmente eso no garantiza que tengas éxito o que alcances la meta a la que quieres llegar. Porque además debes tener la voluntad de seguir y no rendirte a pesar de que falles, a pesar de que sea difícil, debes seguir intentándolo. No se si habrá sido directamente o indirectamente, pero el verlos a ambos todos este tiempo, también fue una forma de motivarme avanzar y a no rendirme a pesar de la adversidad. Por todo eso no puedo estar más agradecido con ellos.

También debo agradecer a mis amigos por estar siempre dándonos apoyo mutuamente, desde tardes en las que nos quedábamos estudiando hasta aquellos días en los que nos quedábamos para compartir y reírnos un rato. Gracias a mis dos profesores guías por preocuparse constantemente del desarrollo de este trabajo de memoria desde extensas reuniones hasta pláticas o conversaciones para reírse y hacer muy agradable el trabajar con ellos. Gracias a todas las personas que hacen de la universidad un buen lugar donde estar y gracias a ti por leer esto.

TABLA DE CONTENIDOS

ÍNDICE DE FIGURAS

página

ÍNDICE DE TABLAS

página

RESUMEN

Los grafos y la ley de potencia están estrechamente relacionados en el mundo real. Por ejemplo en una red social, la relación de amigos o seguidores globales de un individuo obedecen la distribución de ley de potencia. Este tipo de redes son muy estudiadas en la actualidad, pero no se proporciona acceso a estas, porque los datos son un recurso crítico y sensible. Esto ha llevado a la creación de métodos que generan grafos sintéticos que reproducen las características de grafos reales. Entre dichos métodos se encuentran R-MAT y R^{3}MAT, donde el primero tiene problemas del uso de memoria y al segundo le toma mucho tiempo generar la distribución de aristas por su funcionamiento recursivo. Por esto, el problema que se enfrenta en esta memoria es mejorar el tiempo de producción de un grafo sintético manteniendo las propiedades de la ley de potencia. Para esto, se usa una metodología de investigación cuantitativa de cuatro fases. Es muy relevante en esta memoria la función de probabilidad de la ley de potencia, $P(x = k) = C \cdot k^{-\tau}$ que en el contexto de grafos indica la fracción de nodos que deben tener k vecinos para que esta estructura mantenga la distribución de la ley de potencia. Para la generación de grafos sintéticos, en los métodos de esta memoria se utiliza directamente la función de probabilidad de la ley de potencia para aproximar la cantidad de nodos que deben existir por el número de vecinos, con el fin de mantener dicha propiedad en las estructuras generadas. Como resultado, se han obtenido tres métodos que generan grafos con distribución de la ley de potencia con los cuales se han generado grafos de hasta 250 millones de nodos llegando a generar más de tres mil millones de aristas en algunos casos. Mediante las evaluaciones experimentales que se realizan para comparar uno de los métodos propuestos con otros algoritmos existentes, se comprobó una reducción del tiempo de generación con respecto a estos métodos alternativos. También se probo que efectivamente los grafos generados mantienen la propiedad de la ley de potencia por medio de la visualización de la distribución del número de nodos que se genera para los valores de k vecinos posibles. Finalmente, se logró obtener un método que reduce el tiempo de generación y que produce un grafo que mantiene la propiedad de distribución de la ley de potencia.

1. Introducción

Actualmente, los grafos tienen diversas aplicaciones en áreas como Big Data [1] debido a que permiten representar estructuras interconectadas y datos sin esquemas. Esta escala masiva de datos abruma fácilmente a la memoria principal y los recursos de cómputo. Dentro de las aplicaciones de datos masivos se tiene el análisis de redes sociales [2], donde se utilizan los grafos subyacentes para la representación de la red y así poder detectar comunidades, lo que supone uno de los problemas más importantes de esta área. En el ámbito de recuperación de información [3], se utiliza el PageRank para determinar la importancia de los datos. Calcular dicho valor de forma eficiente es un área de investigación contemporánea, en donde se proponen estructuras de grafos dirigidos acíclicos y descomposición de las redes representadas por estas estructuras.

El desafío al que se enfrentan los métodos de generación de grafos es la reproducción de grafos muy grandes que posean las propiedades existentes en redes reales. Entre dichas propiedades se encuentran: distribución de grados, correlaciones de datos, comunidad, el grado de agrupamiento y clasificación.

Uno de los principales problemas al desarrollar aplicaciones basadas en grafos, es la disponibilidad de conjuntos de datos grandes y representativos debido a que los datos son un recurso sensible y crítico para las organizaciones. La falta de acceso a este tipo de información ha motivado el desarrollo de herramientas de software para generar grafos con diferentes metodologías.

1.1. Contexto

Los grafos en la actualidad son muy estudiados por las diversas aplicaciones que estos poseen en diferentes áreas, pero la falta de acceso a grandes volúmenes de información a motivado el desarrollo de métodos que generan grafos sintéticos, que son representaciones de grafos reales que cumplen con las propiedades que estos poseen. Ahora bien, el problema al que se enfrentan estos métodos es la generación de grafos grandes, lo que hace referencia a grafos que superan los 100 millones de nodos y que a su vez cumplan las propiedades que se presentan en los grafos naturales.

Un caso interesante de estudio es cuando la aridad de los nodos sigue la ley de potencia. Esta ley resulta ser una función de densidad de probabilidades que es usada para variables aleatorias que toman un conjunto discreto de valores. La ley de potencia establece una relación entre dos magnitudes y en el contexto de los grafos, esta permite establecer la cantidad de nodos que deben existir para cada grado de aridad que se pueda generar de acuerdo al tamaño del grafo. Dentro de los métodos para la generación de grafos que siguen la ley de potencia se encuentra R-MAT. Dado el hecho de que R-MAT funciona con una matriz de adyacencia, surge un problema básico que es ¿cómo generar un grafo grande cuya matriz de adyacencia no cabe en la memoria principal? Si bien, existen afirmaciones que señalan que R-MAT se paraleliza fácilmente con lo que se podrían generar grafos grandes, independiente de esto, administrar un requisito cuadrático de memoria no es una tarea factible incluso para una cantidad tan modesta como los es 100 000 nodos[1].

1.2. Problema

El tener la necesidad de generar grafos tan grandes produce problemas de administración de memoria, dependiendo de la estructura de datos que utilice un método. Además, toma mucho tiempo crear un grafo de esta escala.

Debido al problema de escalabilidad que posee R-MAT al estar limitado a generar un grafo del tamaño que soporte la cantidad de memoria principal, se ha desarrollado un método de generación que utiliza un arreglo [4] a cambio de la matriz que supone el algoritmo original. Este cambio en la estructura aumenta la escalabilidad, manteniendo el tiempo de ejecución del proceso de generación o incluso en algunos casos mejorando dicho tiempo. Esto se logra debido a que el método se basa en la creación de un arreglo auxiliar que contiene la distribución de la aridad de los nodos,

[1] Cabe destacar que en esta memoria por compatibilidad con software de matemáticas se utiliza un espacio breve como separador de miles y el punto (.) como separador de decimales, por ejemplo 1 000,0.

en donde cada posición de esta estructura es el identificador de cada nodo. Este método por lo general toma el mismo tiempo de ejecución que R-MAT, sin embargo, sólo necesita $O(n)$ espacio de memoria.

De acuerdo a lo anterior, el método basado en una matriz presenta problemas de memoria al tener que generar la estructura de datos (matriz de adyacencia) completa. Por su parte el método basado en arreglo, que logró solucionar el problema de memoria de la matriz, tarda mucho tiempo en generar grafos grandes. Esto se debe a que simula el proceso de subdividir recursivamente la matriz para generar una arista. Como debe hacer demasiadas particiones con el fin de mantener la distribución de aridades que sigue la ley de potencia, resulta costoso en términos de tiempo.

1.3. Hipótesis

La pregunta de investigación que se pretende abordar en este trabajo consiste en determinar si es posible disminuir el tiempo de producción del arreglo de aridades, y en caso de ser afirmativo las consecuencias de este procedimiento:

Hipótesis 1

- A partir de la función de probabilidad de la ley de potencia se induce la producción del arreglo de aridades.

Hipótesis 2

- El uso de la función de probabilidad de la ley de potencia permite reducir el tiempo de producción del arreglo de aridades.

Hipótesis 3

- Al inducir la creación del grafo por medio de la función de probabilidad de la ley de potencia, este mantiene las propiedades que se esperan.

1.4. Objetivos

En esta sección se detalla tanto el objetivo general como los objetivos específicos a considerar para el desarrollo de esta memoria.

1.4.1. Objetivo General

- Desarrollar y evaluar algoritmos para la generación de grafos basándose en la función de densidad de probabilidad de la ley de potencia.

1.4.2. Objetivos Específicos

- Diseñar uno o más algoritmos de generación de grafos utilizando la predicción de la función de densidad de probabilidad de la ley de potencia.

- Verificar que la generación de grafos de uno o más algoritmos cumple con la propiedad de la ley de potencia.

- Comparar el rendimiento del o los algoritmos propuestos con uno o más algoritmos ya existentes.

1.5. Propuesta de Solución

Como solución a los problemas planteados, se propone hacer uso de la probabilidad de la ley de potencia para generar la aridad de cada nodo directamente.

De acuerdo a la ley de potencia, la probabilidad de que un nodo tenga k vecinos es $P(x = k) = Ck^{-\tau}$, donde C es un factor de normalización, tal que la suma de todas las probabilidades sea 1. Entonces la cantidad de nodos que tienen k vecinos en un grafo es $nCk^{-\tau}$, donde n es el número total de nodos del grafo. Se le llama vecinos a los nodos que están conectados directamente a un nodo y en el caso de un grafo dirigido son los nodos objetivos de la o las aristas que salen de un nodo en particular. Por ejemplo, al momento de crear un grafo de $n = 10\,000$ nodos, se calcula la probabilidad de todas las posibles aridades. Suponga que $\tau = 2$ y $C = 0.67$, entonces para los nodos con aridad $x = 1$ se tiene la fórmula $P(x = 1) = C \cdot 1^{-\tau} = 0.67$, lo que corresponde a la probabilidad de que un nodo tenga un vecino.

Ahora, como la cantidad de nodos que se desea es 10 000, se multiplica la probabilidad por este valor, por lo que se tiene que $nC \cdot 1^{-\tau}$ es la fracción de los 10 000 nodos que tiene un vecino. Ahora, utilizando el valor que se obtuvo para la probabilidad de tener un vecino, se tiene que la cantidad de nodos que tienen un vecino está dada por $nC \cdot 1^{-\tau} = 6\,667$. Con $k = 2$, se tiene $P(x = 2) = C \cdot 2^{-\tau} = 0.16667$ que es la probabilidad de tener 2 vecinos, entonces evaluando $nC \cdot 2^{-\tau}$ se obtiene

$1\,666.833 \approx 1,667$ que es la cantidad de nodos que tienen 2 vecinos. Esto continua sucesivamente hasta $k = 80$, donde se tiene que $P(x = 80) = C \cdot 80^{-\tau} = 0.00010$ es la probabilidad de tener 80 vecinos, entonces: $nC \cdot 80^{-\tau} = 1.04177 \approx 1$, lo que indica la cantidad de nodos que tienen 80 vecinos. Para valores de $k > 80$, la estimación de la cantidad de nodos que tenga k vecinos es menor que 1.

Si se repite el proceso anterior desde 1 hasta $n-1$ (que es la cantidad de posibles aridades que puede tener un nodo en un grafo de n nodos) se tiene la cantidad de nodos asociada a cada aridad. El Cuadro 1.1 muestra la distribución de aridades.

Cuadro 1.1: Distribución de aridades.

k (Aridad)	Cantidad de nodos
80	1
79	1
$\vdots$	$\vdots$
44	3
$\vdots$	$\vdots$
20	17
$\vdots$	$\vdots$
2	1 667
1	6 667

A medida que se generan las aridades, se incorporan al arreglo de aridades tal como se muestra en el Cuadro 1.2.

Cuadro 1.2: Arreglo de aridades.

80	79	...	44	44	44	...	20	20	20	...	2	2	...
0	1	...	26	27	28	...	80	81	82	...	1854	1855	...

De acuerdo con el Cuadro 1.2, se agrega inicialmente las aridades con menor frecuencia de nodos existentes, como es el caso de 80 en donde sólo existe 1 nodo con esa aridad, mientras que el caso de 44 existen 3 nodos, por lo que en el arreglo se distribuye en 3 celdas la aridad 44 y así sucesivamente. Por lo mismo, el cálculo de estas inicia desde el número más alto posible hasta el más bajo (de $n-1$ hasta 1) y se agregan directamente al arreglo. Para efectos del ejemplo, el Cuadro 1.1 sólo muestra aquellas aridades que poseen al menos un nodo, de tal forma que se omiten los valores de k para los cuales no existe nodo alguno. Por lo que cabe señalar que se dan los casos en que existen 0 nodos con aridad k, para valores de $k > 80$.

Este proceso supone un ahorro en el tiempo en la distribución de aristas generando el arreglo de aridades directamente desde la probabilidad de la ley de potencia.

1.6. Alcances

En el marco del desarrollo de esta memoria, se establecen los siguientes alcances que delimitan diferentes aspectos de ésta:

- Los algoritmos están diseñados y desarrollados para operar en una sola máquina con un solo procesador.
- Los algoritmos utilizados en las comparaciones deben tener características similares a los propuestos.
- El desarrollo de los algoritmos y su implementación no comprende el desarrollo de una interfaz gráfica, sino que son ejecutados por consola por medio de instrucciones con parámetros específicos.
- El formato del resultado de los algoritmos es una lista de aristas de la forma (u, v), donde u es el nodo de origen y v el nodo objetivo. Estas indican los nodos conectados en forma de texto plano, por lo que no se espera la generación visual de los grafos con nodos en forma de círculos y aristas en forma de flechas o líneas.

1.7. Estructura del Documento

En el capítulo de *Antecedentes*, se muestra un breve descripción de funciones de probabilidad, teoría de grafos, estructuras de datos y el trabajo relacionado con algoritmos de generación de grafos sintéticos, particularmente relacionados a R-MAT.

En el capítulo de *Metodología*, se describe la planificación de actividades realizadas durante la investigación, los criterios empleados para el diseño de los algoritmos desarrollados y la planificación de las pruebas preliminares.

En el capítulo de *Generación de Grafos acorde a la Ley de Potencia* se describen los diferentes algoritmos que se proponen para lograr los objetivos de esta memoria y también se plantean experimentos preliminares para detallar el análisis que se hace de cada método durante el desarrollo e investigación.

En el capítulo de *Evaluación Experimental*, se muestran las mediciones efectuadas para los diferentes algoritmos implementados y se entrega evidencia empírica que valida el cumplimiento y comportamiento esperado por cada uno de los métodos que se evalúa.

En el capítulo de *Conclusiones* se realiza una revisión global de los resultados con el fin de analizar si se ha cumplido lo sostenido en la Hipótesis y sus consecuencias. Por ultimo, se habla del trabajo futuro al que puede llegar esta memoria.

2. Antecedentes

2.1. Funciones de Probabilidad

En teoría de probabilidades, una función de probabilidad [5] establece una relación entre cada punto de su espacio muestral (conjunto de todos los posibles resultados de un experimento aleatorio) con la probabilidad de que ocurra.

En otras palabras, si se tiene un espacio muestral correspondiente a una variable X que tiene los puntos $x_1, x_2, \ldots, x_k$ la función de probabilidad P asociada a X es:

$P(X = x_i) = p_i$, donde p_i es la probabilidad de que ocurra x_i. El conjunto de todas las probabilidades correspondientes al espacio muestral suma 1.

De la función de probabilidad derivan la función de distribución y densidad de probabilidad. Cabe mencionar que la función de probabilidad es usada para variables aleatorias que toman un conjunto discreto de valores, puesto que para variables aleatorias continuas se usa la función de densidad.

2.1.1. Función de Densidad de Probabilidad

Una función de densidad de probabilidad establece la probabilidad relativa según la cual una variable aleatoria puede tomar un determinado valor, por lo que la probabilidad cae en una región específica del espacio de posibilidades.

Una función de densidad de probabilidad caracteriza el comportamiento probable de una población en cuanto a la posibilidad de que una variable aleatoria continua X tome un valor cercano a x [6].

Una variable aleatoria X tiene densidad f de la forma:

$$P[a \leq X \leq b] = \int_a^b f(x)dx \tag{2.1}$$

Propiedades

- $f(x) \geq 0$ para toda x.
- El area bajo la curva es igual a 1:

$$\int_{-\infty}^{\infty} f(x)dx = 1 \tag{2.2}$$

Un ejemplo de esta función es la ley de potencia.

2.1.2. Ley de Potencia

La ley de potencia [7] es una relación matemática entre dos magnitudes que establece que dos escalares k e y se vinculan mediante la siguiente función:

$$y(k) = Ck^{-\tau} \tag{2.3}$$

Donde C es la constante de proporcionalidad y τ es el exponente de la ley de potencia. La ley de potencia de presenta en múltiples campos tanto en la naturaleza como en ámbitos artificiales. En términos generales se trata de una relación funcional entre dos cantidades, donde un cambio relativo en una cantidad se traduce en un cambio relativo en la otra cantidad que resulta en un cambio proporcional en dichas magnitudes, lo que es independiente del tamaño inicial de las cantidades.

En el área de grafos, la función de densidad de probabilidad de la ley de potencia puede establecer la probabilidad de que un nodo pueda tener k *grados* de aridad por medio de $Ck^{-\tau}$, donde el número de nodos que tienen k *grados* es la cantidad total de nodos por la probabilidad, es decir $nCk^{-\tau}$, lo que implica que estos nodos aportan una cantidad de $nCk^{1-\tau}$ aristas en el grafo [4].

2.1.3. Números Aleatorios con Distribución de la Ley de Potencia

Para generar una distribución de Ley de Potencia $P(k)$ a partir de una distribución uniforme $P(y)$, se tiene que $P(k) = Ck^{\tau}$ para $k \in [k_0, k_1]$. Entonces, al normalizar se tiene:

$$\int_{k_0}^{k_1} P(k)dk = C\frac{[k^{\tau+1}]_{k_0}^{k_1}}{\tau+1} = 1 \tag{2.4}$$

Se despeja C de la Ecuación 2.4 y se obtiene:

$$C = \frac{\tau + 1}{k_1^{\tau+1} - k_0^{\tau+1}} \tag{2.5}$$

Si se tiene que y es una variable uniformemente distribuida en $[0, 1]$. Entonces,

$$D(k) = \int_{k_0}^{k_1} P(k')dk' = C \int_{k_0}^{k} k'^{\tau} dk' \tag{2.6}$$

Al calcular la integral de la Ecuación 2.6 se tiene que:

$$y = \frac{C}{\tau + 1} \left(k^{\tau+1} - k_0^{\tau+1}\right) \tag{2.7}$$

Ahora, se despeja la variable k de la Ecuación 2.7. Se obtiene:

$$k = \left(\frac{\tau + 1}{C} y + k_0^{\tau+1}\right)^{\frac{1}{\tau+1}} \tag{2.8}$$

Se reemplaza C en la Ecuación 2.8 con la expresión de la Ecuación 2.5 y se tiene:

$$k = \left[\left(k_1^{\tau+1} - k_0^{\tau+1}\right) y + k_0^{\tau+1}\right]^{\frac{1}{\tau+1}} \tag{2.9}$$

Entonces los valores que puede tomar k, a partir de una serie de números y con distribución uniforme, se distribuyen a partir de la ley de potencia [8].

2.2. Grafos

Un grafo es un par $G = (N, E)$ donde N es un conjunto de $n = |N|$ nodos y E es un conjunto de $m = |E|$ aristas (dirigidas o no dirigidas). Una arista está dada por un par de nodos (u, v), donde u y v se denominan nodos adyacentes. Si se establece que $e = (u, v)$ es una arista dirigida, se tiene que, u es el nodo fuente, v es el nodo objetivo, por lo tanto e es una arista saliente del nodo u y es una arista entrante del nodo v. Si se establece que e es una arista no dirigida, entonces (u, v) es equivalente a (v, u). A partir de esto, se define que un grafo es dirigido cuando contiene sólo aristas dirigidas y es no dirigido cuando contiene sólo aristas no dirigidas [4].

Una arista de la forma (u, u) se denomina rulo. Si hay dos o más aristas que están asociadas al mismo par de nodos (u, v) se les denomina aristas paralelas. Un grado se dice simple, si es que no contiene aristas paralelas o rulos. En esta memoria sólo se consideran grafos simples.

2.2.1. Sobre la adyacencia de los nodos

Un nodo v es adyacente a u si y sólo si $(u, v) \in E$. Entonces, en un grafo no dirigido, si se tiene la arista (u, v), v es adyacente a u y simétricamente u es adyacente a v. Usando la misma arista como ejemplo, en un grafo dirigido, se tiene que sólo v es adyacente a u. En muchos casos, a las aristas se les asigna un peso o costo. A dichas estructuras se les llama grafos ponderados. En esta memoria no se considera el peso de las aristas [9].

La adyacencia de un nodo $u \in N$ es V, donde V es el conjunto de los nodos vecinos de u. Formalmente la $adyacencia(u) = \{v \in N, (u, v) \in E\}$. La adyacencia del nodo u no contiene al nodo u en sí.

2.2.2. Aridad de un Nodo

La aridad o grado de un nodo u en un grafo G simple, es la cantidad de aristas de E que están asociadas a u. Cuando G es un grafo no dirigido la aridad de u coincide con la cantidad de nodos vecinos [10]. En otras palabras, la aridad de un nodo u equivale a la cantidad de nodo adyacentes a éste.

Cuando se habla de grafos dirigidos, el término aridad se divide en 2 grupos, *in degree* (aridad o grados de entrada) y *out degree* (aridad o grados de salida). El indegree de u es la cantidad de aristas que tienen a este nodo como objetivo. El out degree de u equivale a la cantidad de aristas que tienen a este nodo como origen.

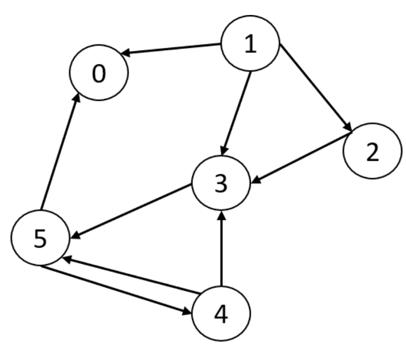


Figura 2.1: Grafo dirigido.

A modo de ejemplo, si se considera el grafo dirigido de la Figura 2.1 se puede determinar el in degree y el out degree de cada nodo como se aprecia en el Cuadro 2.1.

Cuadro 2.1: Aridad del grafo dirigido de la Figura 2.1.

Nodo	In-degree	Out-degree
0	2	0
1	0	3
2	1	1
3	3	1
4	1	2
5	2	2

2.3. Representación de Grafos

En esta sección se detallan diferentes estructuras de datos que permiten representar grafos y las características que presentan.

2.3.1. Lista de Aristas

Según la definición de grafo, éste es un par $G = (N, E)$ formado por un conjunto de nodos N y un conjunto de aristas E. Con esta definición, se puede representar un grafo fácilmente. Cuando la semántica de N es irrelevante, el conjunto de nodos se representa con números entre 0 y $n-1$, lo que sólo indica la cantidad de nodos que existen. Pero en el caso de E, éste se puede representar como una lista, por lo que a esta estructura se le llama lista de aristas [11], donde se tiene todos los pares de la forma (u, v).

Al tomar como ejemplo el grafo de la Figura 2.1, este se puede representar como una lista de aristas de la siguiente forma:

$E = \{(1,0),(1,2),(1,3),(2,3),(3,5),(4,3),(4,5),(5,0),(5,4)\}$

La ventaja de la lista de aristas es que ocupa $O(m)$ espacio para representar un grafo con n nodos y m aristas. La desventaja, es que la lista de arista no es muy útil para realizar operaciones sobre la estructura. Por ejemplo, hay que recorrer toda la estructura para conocer los todos nodos adyacentes de un nodo en específico.

2.3.2. Matriz de Adyacencia

Se puede hacer uso de una matriz M_G para representar cualquier grafo G usando la matriz que representa la relación que existe entre un nodo u y sus nodos adyacentes

($adyacencia(u)$). Si se habla de un grafo simple no dirigido, la relación que existe entre un nodo y sus adyacentes es simétrica, por lo que $M_G = (M_G)^T$, además de que la diagonal de M_G tiene solo 0's. A esta estructura se le llama matriz de adyacencia. Si G tiene n nodos, entonces M_G es una matriz de $n \times n$ [10]. Tomando como ejemplo el grafo no dirigido de la Figura 2.2, su representación como matriz de adyacencia se muestra en el Cuadro 2.2.

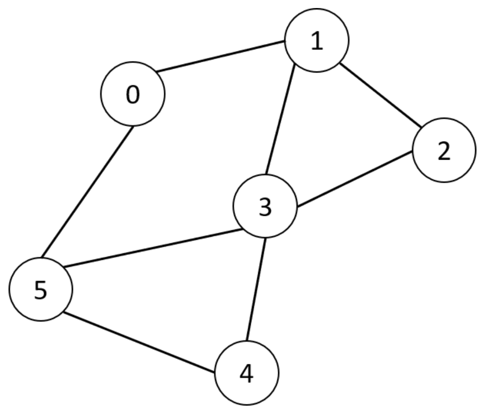


Figura 2.2: Grafo no dirigido.

Cuadro 2.2: Matriz de adyacencia.

	n0	n1	n2	n3	n4	n5
n0	0	1	0	0	0	1
n1	1	0	1	1	0	0
n2	0	1	0	1	0	0
n3	0	1	1	0	1	1
n4	0	0	0	1	0	1
n5	1	0	0	1	1	0

En la matriz del Cuadro 2.2, se puede apreciar que la conexión entre un nodo y otro se denota con el número 1, mientras que la inexistencia de esta se marca con el número 0. Naturalmente, también se puede usar una matriz de bits para ahorrar espacio de memoria.

2.3.3. Arreglo de Aridades

Un arreglo de aridades [4] es una estructura de datos que representa la distribución de aridades de un grafo. Para esto, cada índice del arreglo representa a un nodo específico y el valor contenido en dicha posición indica la aridad o cantidad de nodos vecinos que posee el nodo correspondiente al índice.

Si se considera la Figura 2.1 que muestra un grafo dirigido, su representación como arreglo de aridades es como se muestra en el Cuadro 2.3.

Cuadro 2.3: Arreglo de aridades para el grafo de la Figura 2.1.

0	3	1	1	2	2
0	1	2	3	4	5

En el Grafo Dirigido de la Figura 2.1 se aprecia que el nodo 1 posee 3 aristas que salen de él, por lo que se puede decir que $adyacencia(1) = \{0, 2, 3\}$. Por dicha razón, en el arreglo de aridades del Cuadro 2.3, en la posición 1, que corresponde al nodo 1, se almacena el valor 3, que indica la aridad del nodo. Este proceso se replica para todos los nodos para representar su aridad en el arreglo. Cabe destacar que el arreglo de aridades de un grafo dirigido equivale al out-degree de cada nodo almacenado en su respectiva posición (ver Cuadro 2.1, columna Out-degree).

Si se considera la Figura 2.2 que muestra un grafo no dirigido, su representación como arreglo de aridades es como se muestra en el Cuadro 2.4.

Cuadro 2.4: Arreglo de aridades para el grafo de la Figura 2.2.

2	3	2	4	2	3
0	1	2	3	4	5

Cabe mencionar que el arreglo de aridades del Cuadro 2.4 es equivalente a resumir la matriz de adyacencia del Cuadro 2.2. En efecto, el arreglo almacena la aridad correspondientes a cada nodo del grafo no dirigido de la Figura 2.2, lo que se puede comprobar sumando los valores de cada columna y asignándolos al nodo respectivo.

2.4. Algoritmo R-MAT

El Recursive Matrix o R-MAT [12] es un método para generar datos de grafos con el objetivo de satisfacer tres propiedades: parsimonia, realismo y velocidad. Parsimonia se refiere a que el método es ejecutado con pocos parámetros. Realismo hace referencia a generar grafos que se asemejan a estructuras que existen en la vida real y que pueden ser representadas por medio de estas. Velocidad en términos de generar grafos rápidamente.

La idea básica de R-MAT es generar las aristas del grafo calculando la probabilidad de cada una de estas. Como esto resulta difícil de realizar para cada arista, R-MAT realiza una simulación de tipo Montecarlo para poder repetir o duplicar características y comportamientos de un sistema real. Esto, pues el objetivo principal de las simulaciones de tipo Montecarlo es imitar el comportamiento de variables reales para analizar o predecir cómo estas cambian eventualmente.

Al observar un grafo del mundo real, se puede apreciar la existencia de nodos preferidos, es decir, nodos que poseen un mayor nivel de conexiones con respecto al resto. Un ejemplo es el grafo de seguidores de Twitter, ya que existen usuarios que tienen muchos seguidores y de forma inversa usuarios con pocos seguidores. Entonces, lo que hace R-MAT es concentrar la distribución de aristas en dichas zonas. Para esto, la idea es seleccionar una arista de forma aleatoria tratando de imitar este comportamiento y así realiza este proceso muchas veces.

Para simular la preferencia, se toma la matriz completa del grafo la que se divide en 4 cuadrantes (Figura 2.3) y se le asigna una probabilidad razonable a cada cuadrante. De esta forma, cada cuadrante tiene una preferencia distinta de acuerdo a la probabilidad asignada (a mayor probabilidad, mayor preferencia y viceversa). Después, dentro de cada cuadrante se vuelven a asignar la probabilidades siguiendo la misma distribución. Entonces, la idea es poder establecer nodos preferidos que se encuentran en un cuadrante preferencial de acuerdo a la probabilidad asignada. Así, se pueden generar aquellos nodos a los que le llegan más aristas [4]. Cabe mencionar que R-MAT selecciona una arista a la vez.

Para generar un grafo $G = (N, E)$, R-MAT define un procedimiento recursivo para agregar aristas a una matriz de adyacencia, la cual está vacía inicialmente, hasta que se haya agregado un número dado de aristas. El procedimiento recursivo subdivide la matriz de adyacencia del grafo en cuatro particiones llamadas a, b, c y d, donde la selección de x particiones sigue una probabilidad P_x que satisface $P_a + P_b + P_c + P_d = 1$, y agrega una arista a la matriz haciendo $(u, v) = 1$ al final de la recursión [13]. La selección de valores convenientes para P_a, P_b, P_c y P_d resulta ser crucial para un buen desempeño del algoritmo. En un estudio formal de R-MAT [14], los autores establecen que las probabilidades recomendadas son: P_a=0.55,

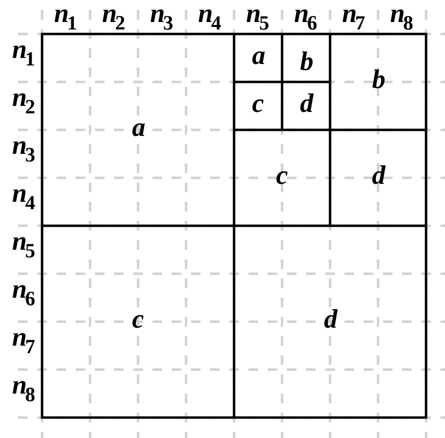


Figura 2.3: Ejemplo de matriz de adyacencia generada por R-MAT.

P_b=P_c=0.1, y P_d=0.25. Mientras que en otro estudio [13], después de varias pruebas y combinaciones, establecen como probabilidades para P_a=0.60, P_b=P_c=0.15, y P_d=0.10. Por ultimo, en [4] se utilizan los valores P_a=0.67, P_b=0.19, P_c=0.10, y P_d=0.04 para grafos dirigidos, mientras que para grafos no dirigidos se utilizan los valores P_a=0.74, P_b=0.17, P_c=0.08, y P_d=0.01.

2.4.1. Estadística de Kolmogorov-Smirnov (KS)

La estadística de Kolmogorov-Smirnov [15], es una prueba no paramétrica que se utiliza para determinar la bondad de ajuste de dos distribuciones de probabilidad entre sí. Cabe mencionar que la bondad de ajuste de un modelo estadístico describe lo bien que se ajusta a un conjunto de datos. Por lo general, esta estadística proporciona el medio para probar si un conjunto de observaciones son de una distribución continua especificada. Una de las ventajas que posee, es que se puede usar con tamaños de muestras pequeño.

En este sentido, la estadística de Kolmogorov-Smirnov permite comparar una distribución de muestra con una distribución de probabilidad de referencia (por ejemplo, la ley de potencia) cuantificando la distancia entre ellas. Las puntuaciones pequeñas para la estadística denotan un mejor ajuste entre las distribuciones en comparación.

2.5. Trabajo Relacionado

La propuesta de investigación está estrechamente ligada al estudio e implementación de R-MAT desarrollada y expuesta en *R^3MAT: A Rapid and Robust Graph Generator* [4], en donde se presentan tres implementaciones del algoritmo. La primera basada en matrices, en donde se expone el problema de memoria y dependencia de la potencia computacional. La segunda implementación basada en arreglos, demuestra ser escalable en términos de memoria logrando generar grafos grandes, pero se observó que tomaba mucho tiempo la generación de la aridad por el método recursivo que implementa. El tercer método basado en tablas de Hash demostró ser escalable en lo que respecta el uso de memoria, pero en términos de tiempo no supera al método basado en matriz. Pero es destacable que los métodos implementados basados en arreglos o en la tabla de Hash logran generar grafos que poseen la propiedad de la ley de potencia, por lo que también se demuestra la replicación de grafos sintéticos con las características esperadas. Adicionalmente esta metodología ofrece otras dos ventajas: el método de generación sigue un *set* específico de parámetros y no es necesario tener conocimiento riguroso para darle uso.

A continuación, se presentan investigaciones que han abordado temáticas similares al problema al que se está abordando con respecto a la generación de grafos.

En la investigación expuesta en el artículo *Towards a property graph generator for benchmarking* [16], los autores exponen el diseño conceptual de DataSynth, que es un framework para la generación de grafos con esquemas y características modificables. Si bien este marco de trabajo permite al usuario definir diferentes estructuras para los grafos y la correlación, el método se basa en un algoritmo de particiones de grafo llamado SBM-Part, que permite coincidir las propiedades con los nodos que se generan. Pero este algoritmo no garantiza una solución óptima, por lo que las restricciones estrictas que se establecen no son garantizadas.

En el artículo *Graph Mining: Laws, Generators, and Algorithms* [7], los autores hablan de patrones sobre grafos y las propiedades que debería producir un generador sobre un grafo, destacando que buscan una compensación entre parsimonia, realismo y eficiencia. Es por esto que deciden usar R-MAT que intenta abordar todas estas preocupaciones que se tienen a la hora de generar dichas estructuras. Si bien no realizan un estudio analítico exhaustivo del modelo, consideran que los tres parámetros que utilizan (tipo de grafo, número de aristas y número de nodos) no dan los suficientes grados de libertad para igualar todas las variedades de grafos que se pueden generar, pero el modelo garantiza y coincide con varios patrones que se esperan de un grafo sintético que simula uno real, como lo es la densidad de la ley de potencia y más patrones con los que coincide R-MAT.

En *A Mathematical Analysis of the R-MAT Random Graph Generator* [14], Chris Groër et al realizaron un análisis matemático de las distribuciones de la aridad de los grafos generados por R-MAT. Se obtuvieron las fórmulas exactas y asintóticas para la distribución de la aridad al modelar la creación de aristas como un problema de ocupación a partir de la teoría de probabilidad. También se proporcionan fórmulas para estimar la media y la varianza del número de aristas correspondientes al grafo final.

El artículo *SNAP: A general-purpose network analysis and graph-mining library* [17] habla sobre la plataforma de análisis de redes de Stanford (SNAP), la cual implementa diferentes métodos para la generación de grafos para máquinas individuales, pero con una gran memoria. La entrada de datos para el método de generación que implementa R-MAT (GenRMAT) es la cantidad de nodos, la cantidad de aristas, las probabilidades de partición y un generador de números aleatorios. La salida es un grafo dirigido. No se entrega más información acerca de la implementación.

En *Linear Work Generation of R-MAT Graphs* [18] los autores presentan un método que busca reducir el tiempo logarítmico de R-MAT a tiempo constante por arista. La idea básica es que pueden generar un número logarítmico de bits de dirección en cada iteración. Estos bits codifican los índices de columna y fila para una arista en cada iteración. Adicionalmente muestran que su algoritmo se puede paralelizar. Se da a entender que su algoritmo secuencial es 10 % más lento que otros métodos secuenciales, pero supone ser más simple que el resto.

2.6. Tecnologías Computacionales

En esta sección se describen las tecnologías y herramientas de software que se utilizan en el desarrollo de esta memoria.

- **Java**
  Es un lenguaje de programación y una plataforma informática diseñado por Sun Microsystems que en la actualidad es propiedad de Oracle Corporation. Al día de hoy, muchas aplicaciones y sitios funcionan gracias a Java, puesto que se considera rápido, seguro y fiable. Este lenguaje usa el paradigma orientado a objetos y permite la ejecución de un mismo programa en múltiples sistemas operativos [19].

- **R**
  Es un entorno y lenguaje de programación libre para computación estadística y gráficos. Es un proyecto GNU desarrollado en los Laboratorios Bell. Este proporciona una amplia variedad de técnicas estadísticas y gráficas, que a su vez es altamente extensible. Una de las características de R es la facilidad con la que se pueden producir gráficos bien diseñados con calidad de publicación [20].

- **Microsoft Excel**
  Es una hoja de cálculo desarrollada por Microsoft para diferentes sistemas operativos. Esta cuenta con herramientas gráficas, cálculos y el lenguaje de programación Visual Basic. Sus principales utilidades son que permite elaborar tablas y formatos de cálculos matemáticos por medio de diferentes fórmulas haciendo uso de múltiples operadores o funciones [21].

3. Metodología

En este capítulo se detallan las fases que se contemplan para el desarrollo de esta memoria y las actividades correspondientes. Se plantea seguir una metodología de investigación cuantitativa [22].

3.1. Fase Conceptual

En esta fase se fórmula y delimita el problema que se está abordando, a partir de la definición del contexto y basándose en la revisión literaria de tópicos relacionados al tema abordado. Las actividades que comprenden esta fase son:

- Estudio de los algoritmos R-MAT y R^{3}MAT.
- Investigación del estado del arte de los algoritmos para la generación de grafos sintéticos que funcionan como alternativas a R-MAT y R^{3}MAT.
- Estudio de funciones de distribución de probabilidades.
- Estudio de los principales componentes de un grafo y estructuras que se utilizan para su representación.
- Definición y delimitación del problema que se aborda.
- Construcción del marco teórico.
- Formulación de objetivos e hipótesis.

3.2. Fase de Planeación y Diseño de Investigación

En esta fase se definen los métodos y estrategias que se emplean para resolver el problema. Por esto, se establece un diseño experimental de investigación en donde se realiza intervenciones de forma activa al proponer, desarrollar y modificar los algoritmos que se consideran necesarios. Las actividades que comprenden esta fase son:

- Diseño, implementación y pruebas de un mecanismo que construya un arreglo de aridades usando la función de densidad de probabilidad de la ley de potencia.
- Diseño e implementación de un mecanismo que construya un arreglo de probabilidades acumuladas a partir de la función de densidad de probabilidad de la ley de potencia.
- Diseño e implementación de un mecanismo que construya un arreglo con la frecuencia de nodos por aridad en base a un arreglo de probabilidades acumuladas.
- Diseño, implementación y pruebas de un mecanismo que construya un arreglo de aridades en base a un arreglo de frecuencia de nodos por aridad.
- Estudio y diseño de un mecanismo que permite modificar la distribución uniforme que poseen los números aleatorios generados en Java a una distribución de ley de potencia.
- Estudio y modelado de una función que permita calcular la aridad máxima de un grafo, dado el tamaño y valor de τ de la estructura que se desea construir.
- Diseño, implementación y pruebas de un mecanismo que construya un arreglo de aridades con la generación de números aleatorios con distribución de ley de potencia y la función que calcula la aridad máxima.

3.3. Fase Empírica y Analítica

Esta fase corresponde a pasar a la ejecución del estudio, donde en primer lugar se lleva a cabo una recolección de datos y la preparación de estos para su análisis. En otras palabras, se lleva a cabo una serie de experimentos en los que se generan diversos grafos, en los que su posterior análisis permite sacar conclusiones generales que apuntan a esclarecer el problema formulado al inicio de esta memoria. En esta fase se realizan dos actividades:

- Experimentos preliminares para cada mecanismo que genera un arreglo de aridades.
- Evaluación experimental para la comparación de los diferentes mecanismos y la comparación con el estado del arte.

3.3.1. Método de Evaluación Experimentos Preliminares

En estos experimentos se consideran dos variables: tamaño del grafo y tipo de grafo. En términos de tamaño, se consideran grafos de 1 000, 10 000, 100 000 y 1 000 000 de nodos. Por su parte, el tipo de grafo indica la generación de aristas dirigidas o no dirigidas.

En base a los parámetros anteriores, se evalúa el método de generación en términos de eficiencia midiendo el tiempo transcurrido a medida que aumenta el número de nodos. Dada las características del grafo, se analiza si se alcanza la cantidad de aristas esperada con una fórmula que se propone más adelante y para todos estos experimentos se considera $\tau = 2$.

3.3.2. Configuración Experimental para Experimentos Preliminares

Estos experimentos se ejecutan en un ordenador con CPU Intel Core i3-7100U con 2.40 GHz, 4 GB de memoria (RAM) y 240 GB de memoria secundaria (SSD). El sistema operativo es Windows 10 Pro de 64 bits y se instaló Java versión 1.8.0_251. Se genera una aplicación Java para cada método que se plantea evaluar y se ejecuta por línea de comandos para la generación de los grafos.

3.4. Fase de Difusión

En esta fase se divulgan los resultados obtenidos, por lo que esta etapa comprende la formalización de todo el proceso realizado en el presente documento de memoria, por lo que se da cuenta de los objetivos propuestos, el diseño metodológico utilizado, resultados, dificultades y limitaciones. Se concluye el trabajo realizado junto con sugerencias o recomendaciones para nuevos estudios.

4. Generación de Grafos acorde a la Ley de Potencia

En este capítulo, se detallan los algoritmos propuestos a forma de solución, los resultados que han generado y su respectivo análisis.

4.1. Algoritmo 1

Este algoritmo hace uso de la función de probabilidad de la ley de potencia para generar la aridad de cada nodo directamente.

Se tiene que $P(x = k) = Ck^{-\tau}$ indica la probabilidad de que un nodo tenga k vecinos. Por lo tanto, el número total de nodos que tienen k vecinos en un grafo es $nCk^{-\tau}$, donde n es el número de nodos total del grafo. Esto último, implica que todos estos nodos agregan $knCk^{-\tau}$ aristas, lo que equivale a la expresión:

$$nCk^{1-\tau} \tag{4.1}$$

Como la suma de todas las probabilidades es 1, esto debe corresponder a sumar las probabilidades de que un nodo tenga k vecinos de 1 hasta $n-1$, vale decir:

$$C\sum_{k=1}^{n-1} k^{-\tau} = 1 \tag{4.2}$$

Al calcular el valor de la sumatoria de la Ecuación 4.2 y fijar un valor para τ, se puede obtener un valor para el factor de normalización (C).

El número de aristas m de un grafo que sigue la ley de potencia se obtiene al sumar los valores de la Ecuación 4.1, donde k toma valores desde 1 hasta $n-1$. Esto es:

$$m = nC \sum_{k=1}^{n-1} k^{1-\tau} \tag{4.3}$$

En el mundo real, el exponente de la ley de potencia τ para grafos naturales está en el rango $[1.92, 2.72]$ [7]. Usar un valor dentro de ese rango conlleva a determinar un valor para el factor de normalización haciendo uso de la Ecuación 4.2 cuando $\tau > 1$ [4]. Se calcula la sumatoria aproximándola por integrales. Como la función $k^{-\tau}$ es monótonamente decreciente, la sumatoria de la Ecuación 4.2 se aproxima de la forma:

$$\int_1^n k^{-\tau} dk < \sum_{k=1}^{n-1} k^{-\tau} < 1 + \int_1^{n-1} k^{-\tau} dk \tag{4.4}$$

La integral del lado izquierdo es $\frac{n^{1-\tau}-1}{1-\tau}$. Al lado derecho se suma 1 a la integral, debido a que es el valor exacto de $k^{-\tau}$ cuando $k = 1$. Por lo tanto, la integral es $\frac{(n-1)^{1-\tau}-1}{1-\tau}$. Se usan las aproximaciones de Taylor y la integral es $(n^{1-\tau} - \frac{1-\tau}{n^{\tau}} - 1)/(1-\tau)$. Como el $\lim_{n\to\infty} (\frac{1-\tau}{n^{\tau}}) = 0$, este término se omite. Entonces se tiene que la integral del lado derecho es $\frac{n^{1-\tau}-1}{1-\tau}$. Con esto, la Ecuación 4.4 se reescribe:

$$\frac{n^{1-\tau} - 1}{1 - \tau} < \sum_{k=1}^{n-1} k^{-\tau} < 1 + \frac{n^{1-\tau} - 1}{1 - \tau} \tag{4.5}$$

Al promediar los límites de la Ecuación 4.5, $\sum_{k=1}^{n-1} k^{-\tau} \approx \frac{n^{1-\tau}-1}{1-\tau} + \frac{1}{2} = \frac{2n^{1-\tau}-1-\tau}{2(1-\tau)}$. Entonces, se reemplaza en la Ecuación 4.2 y se obtiene el factor de normalización C al despejar:

$$C \approx \frac{2(1-\tau)}{2n^{1-\tau} - 1 - \tau} \tag{4.6}$$

Como el valor de $\tau > 1$, se aprecia que $\lim_{n\to\infty} (2n^{1-\tau}) = 0$, por lo que el factor de normalización es:

$$C \approx \frac{2(\tau - 1)}{1 + \tau} \tag{4.7}$$

Con respecto al número de aristas, la función $k^{1-\tau}$ es decreciente. Entonces la sumatoria de la Ecuación 4.3 se aproxima por integrales:

$$\int_1^n k^{1-\tau} dk < \sum_{k=1}^{n-1} k^{1-\tau} < 1 + \int_1^{n-1} k^{1-\tau} dk \tag{4.8}$$

La integral del lado izquierdo es $\frac{n^{2-\tau}-1}{2-\tau}$. Al lado derecho se suma 1 a la integral, debido a que es el valor exacto de $k^{-\tau}$ cuando $k=1$. Entonces, la integral es $\frac{(n-1)^{2-\tau}-1}{2-\tau}$, que aproximadamente es $\frac{n^{2-\tau}-n^{1-\tau}(2-\tau)-1}{2-\tau}$. Pero al ser $\tau>1$, el valor de $n^{1-\tau}$ es insignificante. Así, la integral de la derecha es $\frac{n^{2-\tau}-1}{2-\tau}$. Luego, la Ecuación 4.8 se reescribe:

$$\frac{n^{2-\tau}-1}{2-\tau} < \sum_{k=1}^{n-1} k^{1-\tau} < 1 + \frac{n^{2-\tau}-1}{2-\tau} \tag{4.9}$$

Al promediar los límites de la Ecuación 4.9, $\sum_{k=1}^{n-1} k^{1-\tau} \approx \frac{n^{2-\tau}-1}{2-\tau} + \frac{1}{2} = \frac{2n^{2-\tau}-\tau}{2(2-\tau)}$. Entonces, al reemplazar en 4.3, el número de aristas es:

$$m \approx n\frac{2(\tau-1)}{1+\tau}\frac{2n^{2-\tau}-\tau}{2(2-\tau)} = \frac{n(\tau-1)(2n^{2-\tau}-\tau)}{(1+\tau)(2-\tau)} \tag{4.10}$$

Es importante notar el caso cuando $\tau=2$, ya que la sumatoria en la Ecuación 4.3 se convierte en $\sum_{k=1}^{n-1} k^{-1} \approx ln(n) + 0.577\,22$. El factor de normalización es $C \approx \frac{2}{3}$, por lo que el número de aristas en este caso es:

$$m \approx \frac{2}{3} \cdot n \cdot ln(n) + 0.384\,81n \tag{4.11}$$

El cálculo de todas estas fórmulas y ecuaciones se encuentra con mayor detalle en el Artículo *R^3MAT: A Rapid and Robust Graph Generator* [4].

4.1.1. Generación de la Distribución de Aridades

El método de generación de grafos se basa principalmente en la creación de un arreglo de aridades que contiene la distribución de grados de los nodos. Este proceso se lleva a cabo mediante la función GenerateArrayPowerLaw, que se presenta en el Algoritmo 1.

El parámetro de entrada de la función GenerateArrayPowerLaw es el número de nodos N. La salida es un arreglo D de longitud N, tal que $D[i]$ define el número de aristas del nodo i, donde $0 \leq i < n$.

La función GenerateArrayPowerLaw comienza con la definición del valor de τ, la constante de proporcionalidad C y la inicialización el arreglo de aridades D. Se asigna el valor de $\tau = 2$ y se utiliza la fórmula de la Ecuación 4.7 para la constante C.

En el paso de generar la probabilidad (Línea 7) para una aridad k (valor posible dado el volumen N del grafo) se utiliza la fórmula de la Ecuación 2.3 de la definición de la ley de potencia. Se calcula el producto (Línea 8) de la probabilidad con el número de nodos N. Esto indica la cantidad de nodos que tienen aridad k en el grafo. Dado el valor del producto, se divide en dos casos. Caso uno, cuando la cantidad de nodos resultantes para la aridad k es mayor o igual a uno y el caso dos, cuando es menor a uno.

En el caso uno ($nodosAridad \geq 1$) se redondea el resultado de producto (Línea 10), ya que este valor es un *double* que posee cierta cantidad de cifras decimales y una parte entera, siendo esta última la que indica la cantidad de nodos con aridad k. En el paso de asignar la aridad k al nodo i (Línea 13) se está añadiendo la aridad k a un solo nodo en el arreglo de aridades. Por esto último, este paso se repite la cantidad de veces indicada por la cantidad de nodos con aridad k (valor de variable $nodosPorAridad$).

En el caso dos ($nodosAridad < 1$), inicialmente se obtiene un número aleatorio r entre 0 y 1 (Línea 19). El valor de r se compara con la cantidad de nodos para la aridad k (Línea 20), donde si se determina que r es menor a la cantidad, se crea un solo nodo i con la aridad k en el arreglo de aridades D (Línea 21).

Los pasos entre la línea 7 a la 25 se repiten $N-1$ veces desde la aridad $k = N-1$ hasta $k = 1$. El arreglo de aridades producido se retorna en la Línea 27, el que contiene el número de aristas de cada nodo i ordenada de mayor a menor. De este modo el nodo 0 tiene la mayor aridad, mientras que el nodo $N-1$ contiene la menor aridad. La representación gráfica del arreglo que retorna el Algoritmo 1 se aprecia en el Cuadro 4.1. Este arreglo tiene datos de un grafo de 1 000 nodos indexados del 0 en adelante generado con el método descrito.

Cuadro 4.1: Arreglo de aridades Algoritmo 1.

941	420	312	202	125	72	69	62	60	...	1
0	1	2	3	4	5	6	7	8	...	999

4.1.2. Experimentos Preliminares Algoritmo 1

En esta sección se evalúan los resultados que genera el Algoritmo 1, lo que permite analizar la eficiencia del método y características de los grafos producidos. El objetivo de estos experimentos es la búsqueda de fallas o falencias en el método de generación para su corrección o que den paso a crear nuevos algoritmos que resuelvan

Algorithm 1: GenerateArrayPowerLaw(int N)

Input: $N \leftarrow$ Número de nodos del grafo.
Output: Un arreglo D, tal que $D[i]$ contiene el grado del nodo i.

```
begin
    double τ ← 2
    double C ← (2 · (τ − 1))/(1 + τ)
    int i ← 0
    long D[] ← new long[N]
    for long k ← N − 1 to 1 do
        double pbb ← C · Math.pow(k, −τ)//Generar Probabilidad
        double nodosAridad ← N · pbb
        if nodosAridad ≥ 1 then
            long nodosPorAridad ← Math.round(nodosAridad)
            long j ← 0
            while j ≤ nodosPorAridad do
                D[i] ← k//Asignar aridad k al nodo i
                i ← i + 1
                j ← j + 1
                if i ≥ N then break;
            end
        else
            double r ← rand.nextDouble() //Caso 2
            if nodosAridad > r then
                D[i] ← k//Asignar aridad k al nodo i
                i ← i + 1
                if i ≥ N then break;
            end
        end
    end
    return D
end
```

los problemas de este primer método.

Resultados

Cada experimento corresponde a una combinación de tamaño del grafo y tipo de aristas. Se registra el tiempo de ejecución y el número de aristas generada. En general, se generan ocho grafos, con un tamaño máximo de un millón de nodos.

Cuadro 4.2: Grafos generados en la evaluación preliminar del Algoritmo 1.

GraphID	Tipo de Grafo	Número de Nodos	Número de Aristas	Tiempo (ms)
UDG1K	No Dirigido	1 000	2 105	12
DG1K	Dirigido	1 000	4 972	11
UDG10k	No Dirigido	10 000	33 606	74
DG10k	Dirigido	10 000	53 172	32
UDG100k	No Dirigido	100 000	402 687	391
DG100k	Dirigido	100 000	745 677	174
UDG1M	No Dirigido	1 000 000	4 890 285	9 441
DG1M	Dirigido	1 000 000	9 135 031	2 030

El Cuadro 4.2 presenta los resultados obtenidos. El tiempo de ejecución más bajo es 11 milisegundos que corresponde a la generación de un grafo dirigido de 1 000 nodos. Por otro lado, el tiempo de ejecución más alto es 9 441 milisegundos que corresponden a la generación de un grafo no dirigido de un millón de nodos. Se aprecia empíricamente que toma más tiempo la generación de grafos no dirigidos.

La fórmula de la Ecuación 4.11 indica la cantidad estimada de aristas para los tamaños de los grafos generados. Las estimaciones son 4 990 aristas para mil nodos, 65 250 aristas para 10 mil nodos, 806 009 aristas para 100 mil y 9 595 150 para un millón de nodos. Al observar la cantidad de aristas generada en cada grafo del Cuadro 4.2, se aprecia que en ningún caso se alcanzó la cantidad estimada.

El Cuadro 4.3 presenta las probabilidades generadas para las aridades posibles en un grafo de 1 000 nodos generado por el Algoritmo 1 (Línea 7). De acuerdo a la fórmula de la Ecuación 4.2 la sumatoria de todas las probabilidades es 1. Al realizar la suma de la probabilidades del Cuadro 4.3 el resultado es 1.095 955 711 lo cual resulta ser mayor a 1. En el Algoritmo 2 se aborda este inconveniente. Este problema se replica para grafos de mayor tamaño. También cabe mencionar que la fórmula que calcula la probabilidad para cada aridad k que es $P(x = k) = Ck^{-\tau}$ no depende de la cantidad de nodos, así como también es el caso de la fórmula de la Ecuación 4.7 para el factor de normalización C. Esto quiere decir, que al mantener el valor de $\tau = 2$, independiente del tamaño del grafo, las probabilidades para cada aridad mantienen el mismo valor. Esto último, implica que las aridades más bajas siempre abarcan una gran porción del total de nodos que se dispone, puesto que tienen mayor preferencia. A su vez, esto disminuye el número de aristas generadas, por lo que no se alcanza la

Cuadro 4.3: Cuadro resumen de las probabilidades generadas para cada aridad posible en un grafo de 1 000 nodos.

Aridad	**Probabilidad**
1	0.666 666 667
2	0.166 666 667
3	0.074 074 074
4	0.041 666 667
5	0.026 666 667
6	0.018 518 519
⋮	⋮
997	0.000 000 671
998	0.000 000 669
999	0.000 000 668

cantidad estimada. El Cuadro 4.3 completo se encuentra en el Anexo A.1.

4.2. Algoritmo 2

Este método de generación de grafos se basa en el cálculo de un arreglo de probabilidades acumuladas y un arreglo de frecuencias para cada aridad, de tal forma que a partir de estos se genera el arreglo de aridades que contiene la distribución de grados de los nodos. Este proceso se lleva a cabo mediante las funciones GeneratePbb, GenerateDegrees y GenerateArrayPowerLaw, que se presentan en los Algorithms 2, 3 y 4, respectivamente[1].

4.2.1. Generación de Probabilidades Acumuladas

A continuación, se describen los pasos y diseño del método para la generación del arreglo de probabilidades acumuladas que hace uso el Algoritmo 2. La idea de generar un arreglo de probabilidades acumuladas surge con el objetivo de resolver el problema del Algoritmo 1 (en donde la suma de las probabilidades en algunos casos puede ser mayor a 1.0).

El parámetro de entrada de la función GeneratePbb (ver pseudocódigo del Algorithm 2) es el número de nodos N. La salida es un arreglo PBB de longitud N,

[1]Por favor, note que Algorithm se refiere a un pseudocódigo en especifico, por ejemplo el que aparece en la página 43.

tal que $PBB[i]$ contiene la probabilidad acumulada que corresponde a la aridad i, donde $0 \leq i < N$.

La función GeneratePbb comienza con la definición del valor de $sumaPBB$ y $pbAcumulada$, como también la inicialización del arreglo PBB. Se define la asignación de $PBB[0] = 0$ para no generar nodos con aridad 0 y debido a que la probabilidad para esta aridad indefine la fórmula de la Ecuación 2.3 de la definición de la ley de potencia. Se asigna el valor de $sumaPBB = 0$ y $pbAcumulada = 0$.

En el paso de generar la probabilidad (Línea 7) para una aridad k (valor posible dado el volumen N del grafo) se utiliza la fórmula de la Ecuación 2.3 de la definición de la ley de potencia. En esta sección cabe mencionar que se utiliza la fórmula de la Ecuación 4.7 para la constante C y el valor de τ es por defecto 2, pero puede ser un parámetro de entrada ingresado al momento de ejecutar el programa. Tanto la constante C como τ son variables privadas de la clase en la que se han declarado las funciones. Una vez calculada la probabilidad para la aridad k, se asigna dicho valor en $PBB[k]$ (Línea 8) y este se suma a la variable $sumaPBB$ para obtener la suma de todas las probabilidades (Línea 9). Este proceso se realiza desde $k = 1$ hasta $N - 1$ (Línea 6). La representación gráfica de $PBB[]$ una vez finalizada la primera estructura for, se aprecia en el Cuadro 4.4 donde la probabilidad para $N - 1$ se aprecia como 0.00 debido a que es un valor decimal muy bajo. El valor de la variable $sumaPBB$ equivale a calcular $C \sum_{k=1}^{n-1} k^{-\tau}$.

Cuadro 4.4: Arreglo de probabilidades Algoritmo 2.

0	0.667	0.167	0.074	0.042	0.027	...	0.000
0	1	2	3	4	5	...	N-1

En el paso de ajustar la probabilidad (Línea 13) y calcular la probabilidad acumulada (Línea 14), primero se obtiene la probabilidad de la aridad i (Línea 12). Este valor se ajusta al dividirlo por la suma de probabilidades ($sumaPBB$), de tal forma que se obtiene la probabilidad ajustada para la aridad i. Con el valor ajustado, este se suma a la variable $pbAcumulada$, el cual a su vez se asigna al arreglo en $PBB[i]$ (Línea 15), lo que reemplaza la probabilidad de i por su probabilidad acumulada.

Los pasos entre la Línea 12 y 15 se repiten $N - 1$ veces desde la aridad $i = 1$ hasta $i = N - 1$. El arreglo de probabilidades acumuladas producido se retorna en la línea 17, el que contiene la probabilidad acumulada de cada aridad i. La representación gráfica del arreglo final se aprecia en el Cuadro 4.5. Se observa que la

Algorithm 2: GeneratePbb(int N)

Input: $N \leftarrow$ Número de nodos del grafo.
Output: Un arreglo PBB, tal que $PBB[i]$ contiene la probabilidad acumulada que corresponde a la aridad i.

```
begin
    double sumaPBB ← 0
    double pbAcumulada ← 0
    double PBB[] ← new double[N]
    PBB[0] ← 0
    for long k ← 1 to N − 1 do
        double pbb ← C · Math.pow(k, −τ)
        PBB[k] ← pbb
        sumaPBB ← sumaPBB + pbb
    end
    for long i ← 1 to N − 1 do
        double pb ← PBB[i]
        double pbAjustada ← pb/sumaPBB
        pbAcumulada ← pbAcumulada + pbAjustada
        PBB[i] ← pbAcumulada
    end
    return PBB
end
```

probabilidad para la aridad $N-1$ es uno, debido a que este último valor es la suma de todas las probabilidades ajustadas, lo que solventa el problema de que el total de las probabilidades sea mayor a 1 como ocurre en el primer algoritmo.

Cuadro 4.5: Arreglo de probabilidades acumuladas Algoritmo 2.

0	0.608	0.760	0.828	0.866	0.890	…	1
0	1	2	3	4	5	…	N-1

4.2.2. Generación de Arreglo de Frecuencias

A continuación, se describen los pasos y diseño del método para la generación del arreglo de frecuencia para cada aridad, el cual hace uso el Algoritmo 2.

Los parámetros de entrada de la función GenerateDegrees (que se ilustra en Algorithm 3) son el número de nodos N y un arreglo de probabilidades acumuladas $PBBA$, tal que $PBBA[i]$ contiene la probabilidad acumulada que corresponde a la aridad i. La salida es un arreglo $FREC$ de longitud N, tal que $FREC[i]$ contiene la cantidad de nodos que corresponde a la aridad i, donde $0 \leq i < N$.

La función GenerateDegrees comienza con la inicialización del arreglo $FREC$. A cada posición del arreglo $FREC$ se le asigna el valor 0 (Línea 4). Esto es debido a que la frecuencia asociada a cada aridad se produce de forma aleatoria. Entonces para diferentes aridades con probabilidades acumuladas muy altas podrían no generarse nodos.

Algorithm 3: GenerateDegrees(int N, double $PBBA[]$)

Input: $N \leftarrow$ Número de nodos del grafo. Un arreglo $PBBA$, tal que $PBBA[i]$ contiene la probabilidad acumulada que corresponde a la aridad i.

Output: Un arreglo $FREC$, tal que $FREC[i]$ contiene la cantidad de nodos que corresponde a la aridad i.

```
1  begin
2      long FREC[] ← new Long[N]
3      for long k ← 0 to N − 1 do
4          FREC[k] ← 0
5      end
6      for long i ← 1 to N do
7          double r ← Math.random()
8          long j ← 1
9          while j < N do
10             double pbaA ← PBBA[j − 1]
11             double pbaB ← PBBA[j]
12             if pbaB > r ∧ r > pbaA then
13                 FREC[j] ← FREC[j] + 1
14                 j = N
15             end
16             j ← j + 1
17         end
18     end
19     return FREC
20 end
```

En el paso de obtener un número aleatorio (Línea 7), se genera un número entre 0 y 1 que se almacena en la variable r. En otras palabras, se genera una probabilidad de forma aleatoria. Dado que se tiene el arreglo de probabilidades acumuladas $PBBA$, se busca el rango de valores en el cual se encuentra r. Para esto se obtiene la probabilidad $pbaA$ (Línea 10) y $pbaB$ (Línea 11), donde $pbaA < pbaB$. Con estas probabilidades acumuladas se verifica si $pbaB > r$ y $r > pbaA$ (Línea 12). En el caso verdadero, esto indica que r se encuentra en el rango de $PBBA[j]$ y $PBBA[j-1]$, donde $1 \leq j < N$. Entonces, en el arreglo $FREC$ se aumenta en 1 la cantidad de nodos asociada a la aridad j (Línea 13) y finaliza la búsqueda del rango al que pertenece r. En el caso falso, se sigue aumentando el valor del índice j, que a su vez representa la aridad, hasta que se encuentre el rango de probabilidades acumuladas al que pertenece r.

El paso de obtener un número aleatorio (Línea 7) se realiza N veces mientras que los pasos entre las Línea 10 y 16 se realizan a lo más N veces, puesto que el rango de valores al que pertenece r puede encontrarse antes de llegar al final del arreglo de probabilidades acumuladas. El arreglo de frecuencias se retorna en la Línea 19, el cual contiene la cantidad de nodos correspondientes a cada aridad i ($FREC[i]$).

4.2.3. Generación del Arreglo de Aridades

A continuación, se describen los pasos y diseño del método para la generación del arreglo de aridades, el cual hace uso del Algoritmo 2.

El parámetro de entrada de la función GenerateArrayPowerLaw (que se muestra en Algorithm 4) es el número de nodos N. La salida es un arreglo D de longitud N, tal que $D[i]$ contiene el grado del nodo i, donde $0 \leq i < N$.

La función GenerateArrayPowerLaw comienza con la inicialización del arreglo de aridades D. Se define el arreglo $PBBA$ el cual recibe el arreglo de probabilidades acumuladas que retorna la función GeneratePbb (Línea 3). Se define el arreglo $frec$ el cual recibe el arreglo de frecuencias que retorna la función GenerateDegrees (Línea 4) en base a N y $PBBA$.

Como el arreglo $frec$ contiene la cantidad de nodos para cada aridad, en la Línea 7 se obtiene la cantidad de nodos de la aridad k (el valor se almacena en la variable $nodosPorAridad$). Se comprueba si la cantidad de nodos es mayor a 0 (Línea 8). En caso de ser verdadero, se asigna la aridad k al nodo h (Línea 11). Se está añadiendo la aridad k a un solo nodo en el arreglo de aridades. Por esto último, este paso se repite la cantidad de veces indicada por la cantidad de nodos con aridad k. Cabe mencionar que el arreglo de frecuencias se recorre desde $N-1$ hasta 1. Esto les da prioridad a los nodos con mayor aridad para ser agregados al arreglo de aridades D.

Algorithm 4: GenerateArrayPowerLaw(int N)

```
   Input: N ← Número de nodos del grafo.
   Output: Un arreglo D, tal que D[i] contiene el grado del nodo i.
1  begin
2      long D[] ← new long[N]
3      double PBBA[] ← GeneratePbb(N)
4      long frec[] ← GenerateDegrees(N, PBBA)
5      long h ← 0
6      for long k ← N − 1 to 1 do
7          double nodosPorAridad ← frec[k]
8          if nodosPorAridad > 0 then
9              long cont ← 0
10             while cont < nodosPorAridad do
11                 D[h] ← k
12                 h ← h + 1
13                 if h ≥ N then break;
14                 cont ← cont + 1
15             end
16         end
17         if h ≥ N then break;
18     end
19     return D
20 end
```

Se debe tener en cuenta que si el índice h, que se utiliza para recorrer el arreglo de aridades D, llega a igualar a N (que es la cantidad de espacios disponibles en D) antes de que se termine de recorrer el arreglo de frecuencias $frec$, se da por terminado el proceso de producción del arreglo de aridades, por lo que se pierden frecuencias asociadas a aridades pequeñas.

El arreglo de aridades producido se retorna en la Línea 19, el que contiene el número de aristas de cada nodo i ordenadas de mayor a menor. De este modo, el nodo 0 tiene la mayor aridad, mientras que el nodo $N-1$ contiene la menor aridad. La representación gráfica del arreglo que se retorna tal como se aprecia en el Cuadro 4.20. Este arreglo tiene datos de un grafo no dirigido de 1 000 nodos indexado del cero en adelante generado con el método descrito.

Cuadro 4.6: Arreglo de aridades Algoritmo 2.

651	480	123	120	82	79	61	60	49	...	1
0	1	2	3	4	5	6	7	8	...	999

4.2.4. Experimentos Preliminares Algoritmo 2

En esta sección se evalúan los resultados que genera el Algoritmo 2, lo que permite analizar la eficiencia del método y características de los grafos producidos. Al igual que los experimentos preliminares del Algoritmo 1, el objetivo de estos es la búsqueda de fallas o falencias en el método de generación para su corrección o que den paso a crear nuevos algoritmos que resuelvan los problemas de este método.

Resultados

Cada experimento corresponde a una combinación de tamaño del grafo y tipo de aristas. Se registra el tiempo de ejecución y el número de aristas generada. En general, se generan ocho grafos, con un tamaño máximo de un millón de nodos.

Cuadro 4.7: Grafos generados en la evaluación preliminar del Algoritmo 2.

GraphID	**Tipo de Grafo**	**Número de Nodos**	**Número de Aristas**	**Tiempo (ms)**
UDG1K	No Dirigido	1 000	4 491	27
DG1K	Dirigido	1 000	6 473	13
UDG10k	No Dirigido	10 000	59 592	115
DG10k	Dirigido	10 000	54 514	38
UDG100k	No Dirigido	100 000	744 043	441
DG100k	Dirigido	100 000	750 447	177
UDG1M	No Dirigido	1 000 000	9 112 037	3 095
DG1M	Dirigido	1 000 000	9 186 680	2 062

El Cuadro 4.7 presenta los resultados obtenidos. El tiempo de ejecución más bajo fue de 13 milisegundos que corresponde a la generación de un grafo dirigido de 1 000 nodos. Por otro lado, el tiempo, de ejecución más alto es de 3 095 milisegundos que corresponden a la generación de un grafo no dirigido de un millón de nodos. Se aprecia empíricamente que toma más tiempo la generación de grafos no dirigidos.

La fórmula de la Ecuación 4.11 indica la cantidad estimada de aristas para los tamaños de los grafos generados. De acuerdo a las estimación de arista para cada caso, sólo el grafo dirigido de 1 000 nodos logra superar la cantidad de aristas estimada, mientras que en el resto de observaciones no alcanzan la cantidad esperada.

En el Cuadro 4.8 se muestra la cantidad de nodos con aridad 1 ($k = 1$) para cada grafo generado. Se aprecia que la cantidad de nodos con dicha aridad en cada grafo

Cuadro 4.8: Cantidad de nodos con aridad 1 para los diferentes grafos generados.

GraphID	**Tipo de Grafo**	**Número de Nodos**	**Nodos para k=1**
UDG1K	No Dirigido	1 000	605
DG1K	Dirigido	1 000	603
UDG10k	No Dirigido	10 000	6 053
DG10k	Dirigido	10 000	6 038
UDG100k	No Dirigido	100 000	61 131
DG100k	Dirigido	100 000	60 750
UDG1M	No Dirigido	1 000 000	608 395
DG1M	Dirigido	1 000 000	607 932

cubre el 60 % o más de los nodos. Es decir que en el arreglo de aridades, el 60 % de los espacios disponibles son ocupados por el valor 1. Esto se debe a que si bien, las probabilidades se ajustan, la aridad 1 sigue teniendo una alta preferencia. Para ejemplificar mejor esto, en la Figura 4.1 se muestra la distribución de aridades del grafo no dirigido de 100 000 nodos que se genera en este experimento, donde se tiene el número de nodos por aridad en el eje y y la aridad en el eje x, ambos en escala logarítmica. Se aprecia que la aridad 1 es la aridad con el mayor número de nodos (608 395), como se espera de la alta preferencia que posee, mientras que el número de nodos disminuye a medida que aumenta la aridad. Esto es un comportamiento esperado teniendo en cuenta el valor de las probabilidades. El problema es que la alta preferencia de la aridad 1 impide alcanzar la aridad estimada. ¿Por qué?, pues ¿Que pasaría si el espacio que usa un 1 en el arreglo de aridades se reemplaza por una aridad mayor? Pues, claramente se perdería un nodo con aridad 1, pero se obtiene un nodo con una aridad más alta, lo que aumenta el número de aristas totales del grafo. Esto permite alcanzar la cantidad de aristas estimadas. ¿Entonces hay que modificar las probabilidades de las aridades? No, porque la probabilidad de cada aridad se calcula con la función de probabilidad de la ley de potencia y luego se ajusta. La preferencia que posee cada aridad es la establecida en el proceso de generación de probabilidades acumuladas.

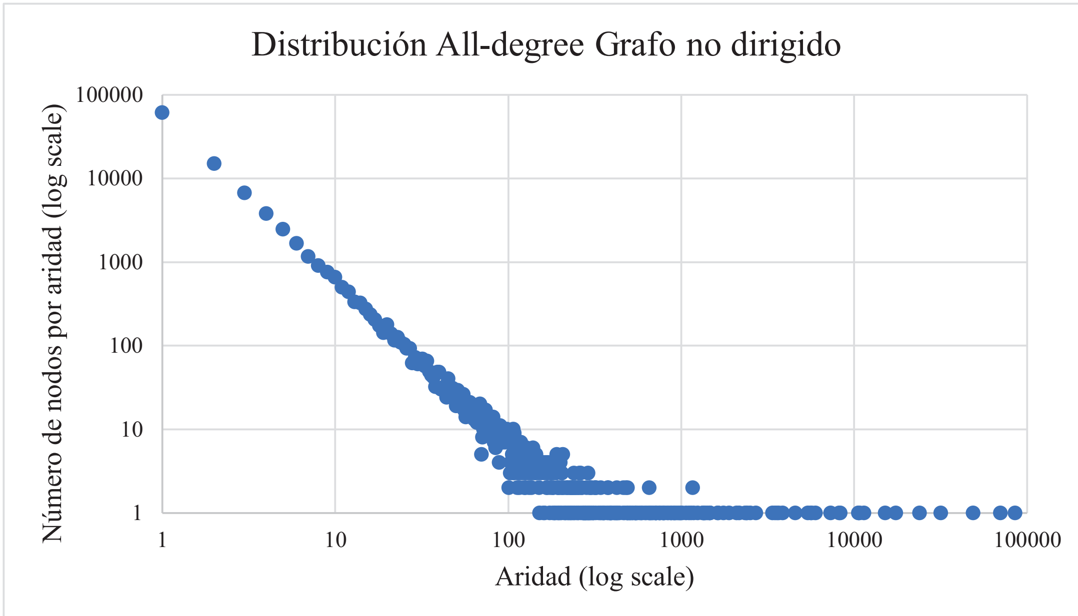


Figura 4.1: Distribución de la aridad en escala logarítmica de base 10 para un grafo de 100 000 nodos.

Entonces, si modificar las probabilidades no es factible en este caso, se analiza la distribución de los números aleatorios generados en Java con el fin de determinar su comportamiento y probar si se puede modificar para que estas probabilidades aleatorios prefieran aridades más altas.

4.3. Distribución de Números Aleatorios

En esta sección se analiza la distribución de los números aleatorios generados en Java haciendo uso del método $random()$ de la clase $Math$. El objetivo es determinar como la distribución afecta a la preferencia de las aridades cuando se genera el arreglo de frecuencias en el algoritmo.

Para llevar a cabo el análisis se generaron 1 000 números aleatorios con la instrucción r=$Math.random()$, la cual genera un *double* positivo, mayor o igual a 0.0 y menor que 1.0 que se almacena en la variable r. Se registra cada ejecución como se muestra en el Cuadro 4.9.

Con estos datos se gráfica cada uno de los números en el gráfico de la Figura 4.2 (a) Generación de los Números. Se aprecia que los números aleatorios se distribuyen equitativamente en el rango de valores de 0.0 a 1.0. Al ordenar los valores de menor

Cuadro 4.9: Registro de los mil números aleatorios obtenidos con Math.random().

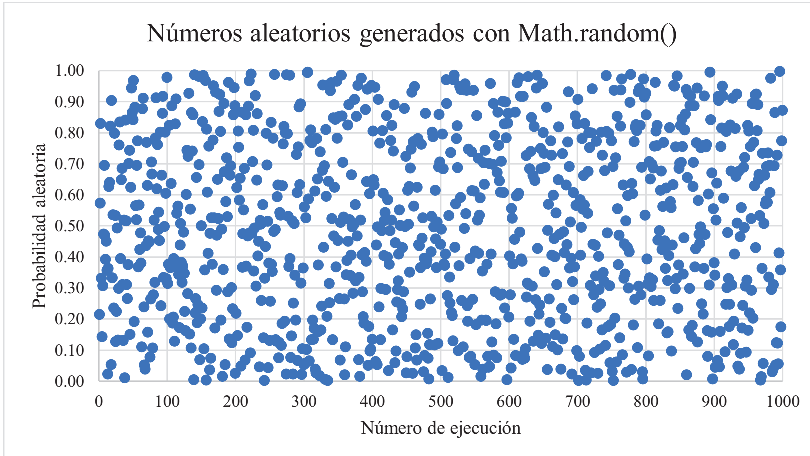
N° de Ejecución	**Número Aleatorio**
1	0.214 773 678
2	0.573 036 830
3	0.830 334 805
4	0.331 650 373
5	0.143 276 368
⋮	⋮
999	0.773 940 181
1 000	0.871 829 258

a mayor, se obtiene el comportamiento de los datos como se aprecia en la Figura 4.2 (b) Distribución. Se aprecia que los números generados por $Math.random()$ poseen una distribución uniforme.

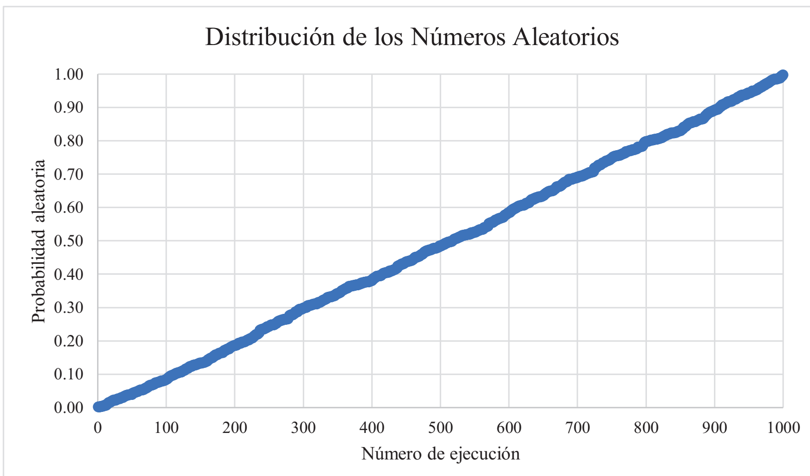
Ahora, de los mil números aleatorios, ¿cuántos pertenecen a cada aridad? En la Figura 4.3 los números generados están clasificados por la aridad a la que pertenecen. Como la probabilidad acumulada para la aridad uno ($k = 1$ en el gráfico) es 0.60 aproximadamente, todos los números aleatorios menores a este valor generan un nodo de aridad 1. Para $k = 2$, todos los números entre 0.60 y 0.76 generan nodos con aridad 2. En el caso de $k = 3$, los números entre 0.76 y 0.82 generan nodos con aridad 3. El resto de números representados con el color azul generan nodos para aridades mayores a 3 dependiendo del rango en el que se encuentre el valor.

En el gráfico (b) de la Figura 4.3 se aprecia que de las mil ejecuciones, más de la mitad corresponden a la aridad $k = 1$, las cuales son 621 para ser exactos. En el caso de las aridades $k = 2$ y $k = 3$, de la mil ejecuciones 143 y 79 les corresponden respectivamente, mientras que el resto se distribuye para aridades mayores a 3. Esto equivale a decir que se generan 621 nodos con aridad 1, 143 con aridad 2, 79 con aridad 3 y así sucesivamente. Es evidente que a medida que aumenta la aridad, la cantidad nodos generados disminuye con respecto a la aridad anterior. Esto se debe a que los rangos de las probabilidades se vuelven más estrechos con el aumento de la aridad y los números aleatorios se distribuyen uniformemente en el rango de 0.0 a 1.0. Esto explica la alta preferencia que posee la aridad 1, lo que a su vez dificulta alcanzar la cantidad de aristas esperada dada la cantidad de nodos y el valor de τ para un grafo.

Entonces, para modificar la preferencia de las aridades se modifica la distribución



(a) Generación de los números.



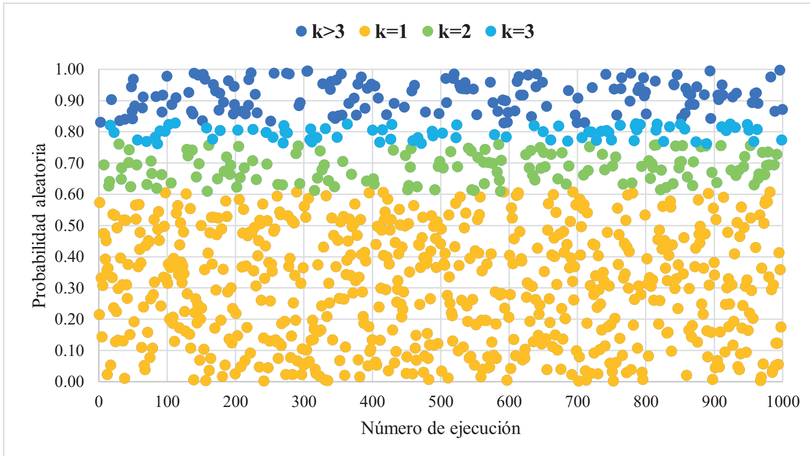
(b) Distribución.

Figura 4.2: Mil números aleatorios generados con Math.random().

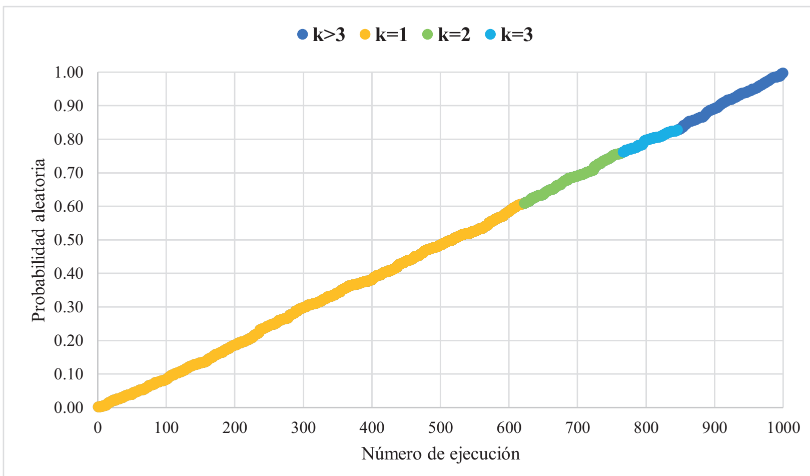
de los números aleatorios que se generan. Para esto, se utiliza la fórmula de la Ecuación 2.9, la cual genera números aleatorios con la distribución de la ley de potencia a partir de una distribución uniforme. Para utilizar la fórmula se definen los parámetros correspondientes a la ecuación.

$$k = \left[\left(k_1^{\tau+1} - k_0^{\tau+1}\right) y + k_0^{\tau+1}\right]^{\frac{1}{\tau+1}}$$

k_0 y k_1 actúan como límite inferior y superior respectivamente para generar números aleatorios con la distribución de la ley de potencia dentro un rango. τ es el exponente



(a) Generación de los números.



(b) Distribución.

Figura 4.3: Los mil números generados clasificados por aridad.

de la ley de potencia, mientras que y es una variable uniformemente distribuida entre 0.0 y 1.0. Entonces $k_0 = 0$ y $k_1 = 1$ para generar números dentro del rango de las probabilidades. $\tau = 2$ e $y = Math.random()$, ya que este método distribuye uniformemente dentro del rango de 0.0 a 1.0 como se ha comprobado. Entonces la Ecuación 2.9 que genera números aleatorios con la distribución de la ley de potencia se reescribe con los parámetros definidos como se muestra en la Ecuación 4.12.

$$k = \left[\left(1^{2+1} - 0^{2+1}\right) \cdot Math.random() + 0^{2+1}\right]^{\frac{1}{2+1}} \tag{4.12}$$

Al calcular las potencias de la Ecuación 4.12, la Ecuación se reescribe como se muestra en la Ecuación 4.13.

$$k = Math.random()^{\frac{1}{3}} \tag{4.13}$$

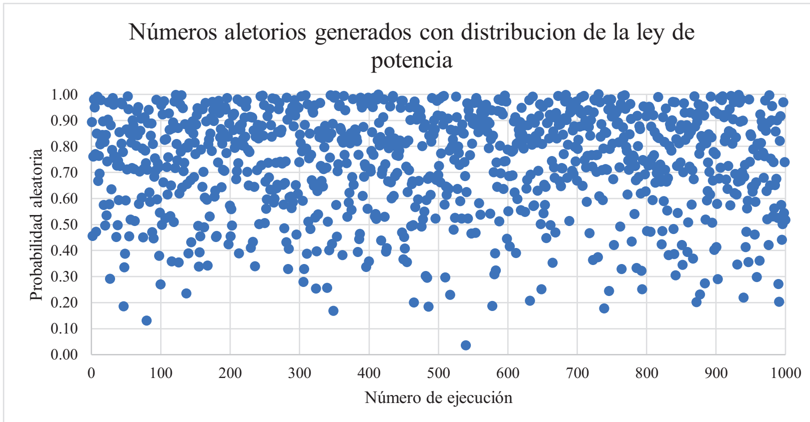
Para llevar a cabo el análisis de la distribución de los números aleatorios que genera la Ecuación 4.12, se generan 1 000 números aleatorios con la instrucción $r = Math.random()^{\frac{1}{3}}$. Se registra cada ejecución como se muestra en el Cuadro 4.10.

Cuadro 4.10: Registro de los mil números aleatorios obtenidos con distribución de la ley de potencia.

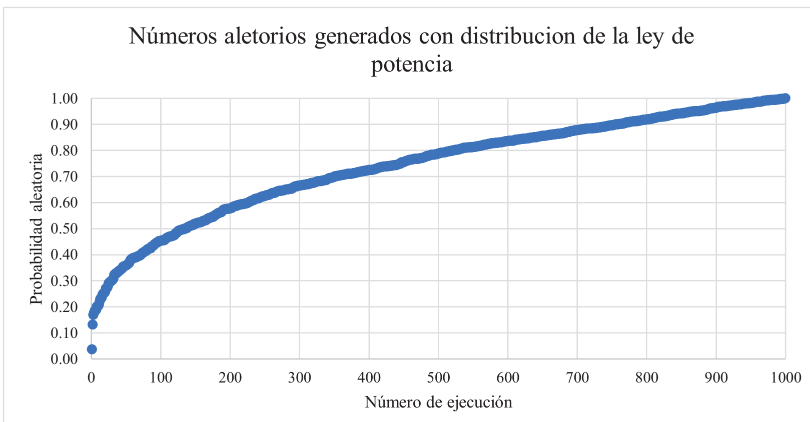
N° Ejecución	**Número Aleatorio**
1	0.892 377 99
2	0.455 242 02
3	0.759 804 11
4	0.979 544 33
⋮	⋮
999	0.739 325 46
1000	0.516 453 26

Con estos datos nuevamente se gráfica cada uno de los números en el gráfico de la Figura 4.4 (a) Generación de los Números. Se aprecia que los números aleatorios ya no se distribuyen uniformemente en el rango de 0.0 a 1.0, sino que se generan más números cercanos a 1.0, lo que aumenta la preferencia por generar nodos con aridades más altas. A su vez, se aprecia que disminuye la cantidad de números aleatorios generados en el rango de 0.0 a 0.60, por lo que disminuye la preferencia de generar nodos con aridad 1. Al ordenar los valores de menor a mayor, se obtiene el comportamiento de los datos como se aprecia en la Figura 4.4 (b) Distribución. Se observa que ahora los números generados poseen distribución de la ley de potencia.

En la Figura 4.5 los números generados están clasificados por la aridad a la que pertenecen. En el gráfico (b) de la Figura 4.5 se aprecia que de las mil ejecuciones, menos de la mitad generan números dentro del rango de la aridad $k = 1$. Para ser exactos, solo 233 ejecuciones pertenecen a la aridad 1. En el caso de las aridades $k = 2$ y $k = 3$, de la mil ejecuciones 221 y 125 les corresponden respectivamente, mientras que el resto se distribuye para aridades mayores a 3. Esto, al igual que en la distribución uniforme de los números aleatorios sin modificar, equivale a decir que se generan 233 nodos con aridad 1, 221 con aridad 2, 125 con aridad 3 y así



(a) Generación de los Números.



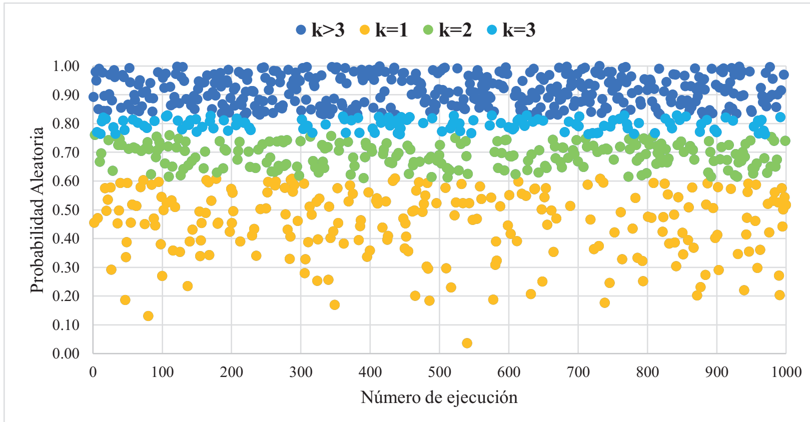
(b) Distribución.

Figura 4.4: Mil números aleatorios generados con distribución de ley de potencia.

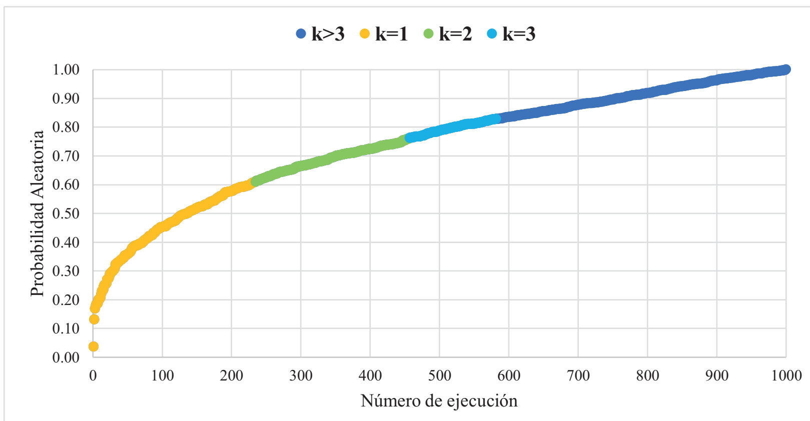
sucesivamente. Las cantidades han cambiado significativamente con la nueva distribución. De esta forma se corrige la alta preferencia que posee la aridad 1 y aumenta la preferencia por aridades más altas, lo que permite aumentar la cantidad de aristas que se generan en un grafo con el fin de alcanzar la cantidad estimada por la fórmula de la Ecuación 4.11.

4.4. Modelado de la Aridad Máxima

En esta sección se detalla el análisis y los experimentos que se llevan a cabo para determinar la aridad máxima asociada a un grafo G simple. Dada una cantidad de nodos N y un valor de τ específico, se busca obtener una función que determine la



(a) Generación de los Números.



(b) Distribución.

Figura 4.5: Mil números generados con la Ecuación 4.13 clasificados por aridad.

aridad máxima.

4.4.1. ¿Por qué Modelar una Función Para Obtener la Aridad máxima?

El modelado de una función que determine la aridad máxima de un grafo, surge de la necesidad de controlar la producción de aridades muy altas que se generan en los Algoritmos 1 y 2. Para analizar esto, se ha generado un grafo de 100 000 nodos con el Algoritmo 1, el Algoritmo 2 y R^3MAT [4]. Al analizar la distribución de aridades del Algoritmo 1 (ver Figura 4.6) y del Algoritmo 2 (ver Figura 4.7), en comparación con R^3MAT (ver Figura 4.8), se aprecia que se están generando nodos con aridades muy altas para un grafo de 100 000 nodos sintético, el cual se desea que posea una distribución de aridades que siga la ley de potencia. En el grafo generado

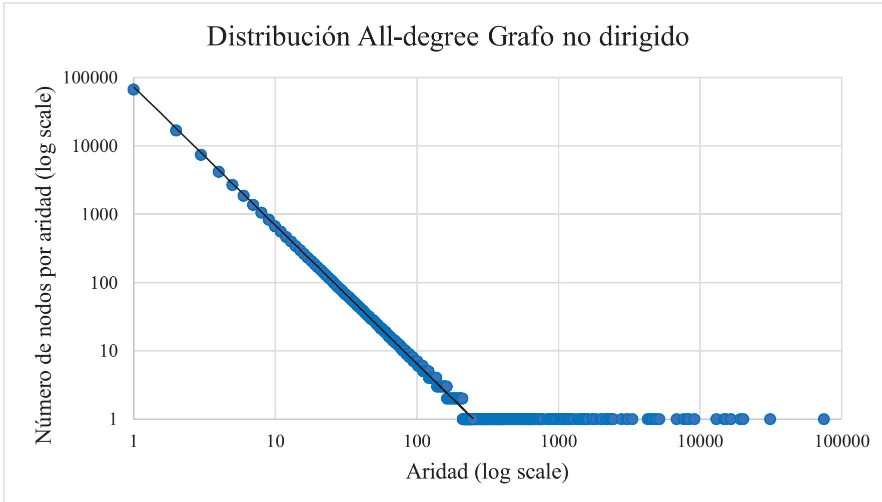


Figura 4.6: Distribución de la aridad de un grafo de 100 000 nodos Algoritmo 1.

por R^{3}MAT, la aridad máxima que se alcanza es 5 277, mientras que en los Algoritmos 1 y 2, las aridades máximas son 75 349 y 85 134, respectivamente. Dado el tamaño

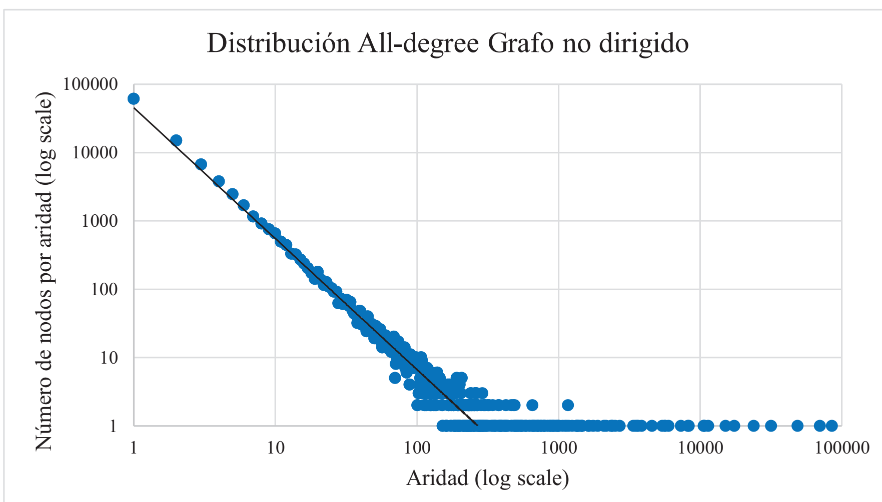


Figura 4.7: Distribución de la aridad de un grafo de 100 000 nodos Algoritmo 2.

del grafo, alcanzar esas aridades esta permitido, pero la aridad máxima que se debe generar debe ser similar a la alcanzada por R^{3}MAT, ya que está comprobado por la estadística de Kolmogorov-Smirnov que los grafos generados con este algoritmo poseen una buena distribución de la ley de potencia. Ahora, en los gráficos también se observa que se tiene una línea de tendencia que representa el comportamiento de

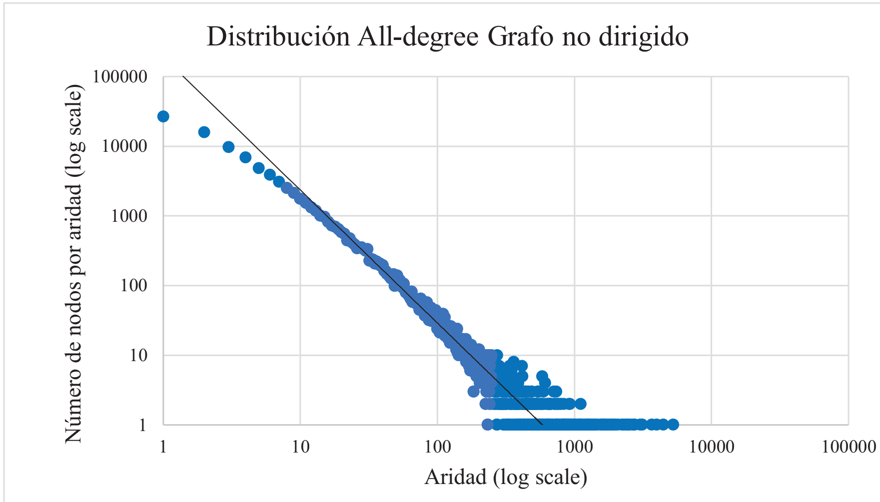


Figura 4.8: Distribución de la aridad de un grafo de 100 000 nodos R^{3}MAT.

la ley de potencia en un gráfico donde se tiene la distribución de las aridades con sus ejes en escala logarítmica. Entonces, el crear un grafo que siga estrictamente la ley de potencia en su distribución, la aridad máxima a la que debe aspirar un grafo de 100 000 nodos tiene que estar en el rango de 100 a 1 000. Este comportamiento se replica para grafos de menor o mayor tamaño.

Entonces, debido a la generación de nodos con aridades muy altas dado el tamaño de un grafo, surge la necesidad de corregir este problema. Para esto se plantea modelar un función que determine la aridad máxima, con el objetivo de limitar la generación de aridades dada una cantidad de nodos N y un valor de τ.

4.4.2. Experimento para Modelar la Función de Aridad Máxima

En esta sección se describe el experimento realizado para determinar la función de aridad máxima. Se especifican los pasos y diferentes procesos que se llevan a cabo.

Modelar la función de aridad máxima se basa en la generación de múltiples grafos. Para la generación de estos, se utiliza el algoritmo R^{3}MAT con el cual se consideran dos variables: tamaño del grafo y valor de τ. En términos de tamaño, se consideran grafos de 1 000, 10 000, 100 000, 1 000 000 y 10 000 000 de nodos. La variable τ considera valores desde 1.92 a 2.72 aumentando en 0.01, es decir que se generan grafos con τ igual a 1.92, 1.93, 1.94 y así hasta 2.72, lo que da un total de 81 valores diferentes para τ.

La generación de los grafos con R^{3}MAT corresponde a la combinación específica del tamaño del grafo y los diferentes valores de τ. Cada combinación se ejecuta 10 veces, se registra la aridad máxima de cada grafo y se calcula el promedio de la aridad máxima por τ.

Cuadro 4.11: Aridades máximas registradas para un grafo de 1 000 nodos.

τ	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**Promedio**
1.92	529	488	539	554	528	562	524	516	525	522	528.7
1.93	538	528	514	545	516	502	501	515	510	477	514.6
1.94	486	527	545	511	490	510	515	526	466	529	510.5
1.95	509	470	479	513	477	486	472	518	535	479	493.8
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
2.71	111	104	108	121	122	107	129	137	114	110	116.3
2.72	119	126	106	120	101	122	133	129	107	138	120.1

Las tablas de los Cuadros 4.11 y 4.12 muestran un resumen de las aridades máximas registradas para grafos de 1 000 y 10 000 000 de nodos respectivamente. Se aprecia que se han generado 10 grafos por cada valor de τ de los cuales se extrae su aridad máxima y se obtiene el promedio correspondiente. Como se tienen 81 valores de τ diferentes y se realizan 10 ejecuciones para cada valor, se generan 810 grafos para un tamaño específico de nodos. Entonces, como se tienen cinco tamaños diferentes

Cuadro 4.12: Aridades máximas registradas para un grafo de 10 000 000 nodos.

τ	**1**	**2**	**3**	**...**	**8**	**9**	**10**	**Promedio**
1.92	117 234	117 245	117 865	...	116 711	117 157	117 033	117 205.2
1.93	107 546	107 145	107 429	...	107 212	106 992	107 338	107 263.9
1.94	98 868	98 103	97 854	...	98 410	98 724	97 885	98 329.2
1.95	89 933	90 426	90 379	...	90 604	90 418	90 199	90 252.6
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
2.71	7 003	7 250	7 266	...	7 215	7 213	7 065	7 172.7
2.72	7 188	7 155	7 221	...	7 225	7 041	7 195	7 157.8

considerados para el experimento, todo el proceso conlleva a la generación 4 050 grafos totales de los cuales se extrae la aridad máxima. Las tablas con los valores de aridad máxima del resto de tamaños se encuentran en el Anexo A.2.1.

Con las columnas que contienen el promedio de la aridad máxima por valor de τ, se genera una tabla general que contiene todos estos valores reunidos y categorizados por el tamaño del grafo como se muestra en la tabla del Cuadro 4.13 como resumen de los registros. La tabla completa con los datos se encuentra en el Anexo A.11. Con

Cuadro 4.13: Aridad máxima promedio dada una cantidad de nodos y un valor de τ para un grafo.

τ	**1 000**	**10 000**	**100 000**	**1 000 000**	**10 000 000**
1.92	528.7	1 719.5	8 208.5	38 459.6	117 205.2
1.93	514.6	1 651.5	7 794.7	35 776.0	107 262.9
1.94	510.5	1 577.5	7 286.9	33 265.2	98 329.2
1.95	493.8	1 496.0	6 918.1	30 896.9	90 251.6
⋮	⋮	⋮	⋮	⋮	⋮
2.71	116.3	268.0	876.5	3 087.3	7 171.7
2.72	120.1	261.2	872.6	3 042.1	7 156.8

los datos tabulados en esta tabla, se modela el comportamiento de la aridad máxima para cada valor especifico de τ. Para esto, se genera un gráfico de dispersión como se ve en la Figura 4.9, donde se toma como variable independiente el tamaño del grafo y la aridad máxima correspondiente como variable dependiente. Los ejes del gráfico han sido modificados a una escala logarítmica para poder visualizar los datos como si tuvieran un comportamiento lineal, pero realmente no es así. En el gráfico se aprecia que se tiene una serie de datos para τ=1.92 (puntos azules) y τ=2.72 (puntos verdes) a los cuales se le ajusta una línea de tendencia potencial con un valor de R^2=0.9965 y R^2=0.9938 respectivamente. De esta forma se obtiene una función potencial y para cada valor de τ, que dada una cantidad de nodos x es capaz de determinar un valor de la aridad máxima para la estructura. El problema es que, si bien no se muestra en el gráfico de la Figura 4.9, se tienen 81 funciones, una por cada τ y cada una determina la aridad máxima sólo para dicho valor de τ.

Ahora, al analizar las funciones obtenidas, se aprecia que todas tienen dos componentes. La constante de proporcionalidad y el exponente de potencia. Tomando como ejemplo la función y de τ=1.92, la constante de proporcionalidad es 7.6706 y el exponente de potencia es 0.6041. Como estos valores se generan para un valor de τ, se plantea modelar una función que determine la constante de proporcionalidad y el exponente de potencia tomando como variable independiente a τ. Para esto, se lleva a cabo un registro de los valores de la constante de proporcionalidad y el exponente de potencia como se muestra en el Cuadro 4.14. Para visualizar la tabla completa el lector se debe dirigir al Anexo A.2.3.

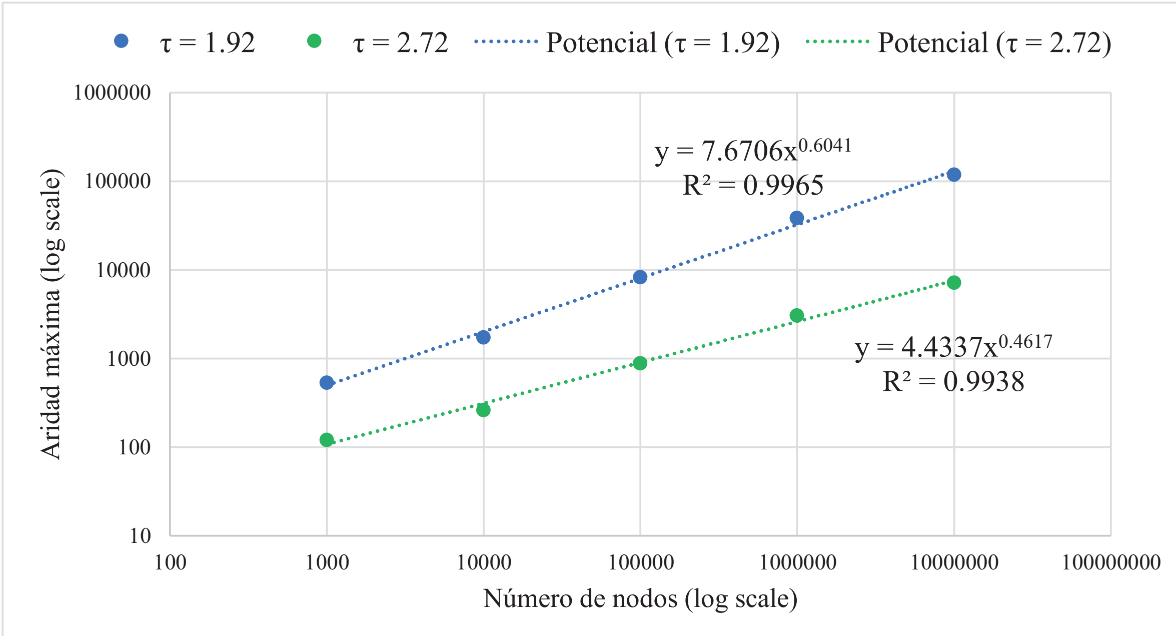


Figura 4.9: Gráfico de dispersión para la aridad máxima versus la cantidad de nodos.

Cuadro 4.14: Constante de proporcionalidad y exponente de potencia por cada valor de τ.

τ	**Constante de proporcionalidad**	**Exponente de la potencia**
1.92	7.6706	0.6041
1.93	7.8379	0.5974
1.94	8.1272	0.5893
1.95	8.154	0.5839
1.96	8.0931	0.5794
$\vdots$	$\vdots$	$\vdots$
2.70	4.3891	0.4639
2.71	4.3208	0.4642
2.72	4.4337	0.4617

Con los datos registrados, para analizar el comportamiento que posee cada componente se genera un gráfico de dispersión para la constante de proporcionalidad y para el exponente de la potencia como se observa en la Figura 4.10 y en la Figura 4.11, respectivamente. Ahora, de esta forma se visualiza la distribución de los datos con el fin de poder entender el comportamiento y ajustar una función que se adecue a los valores.

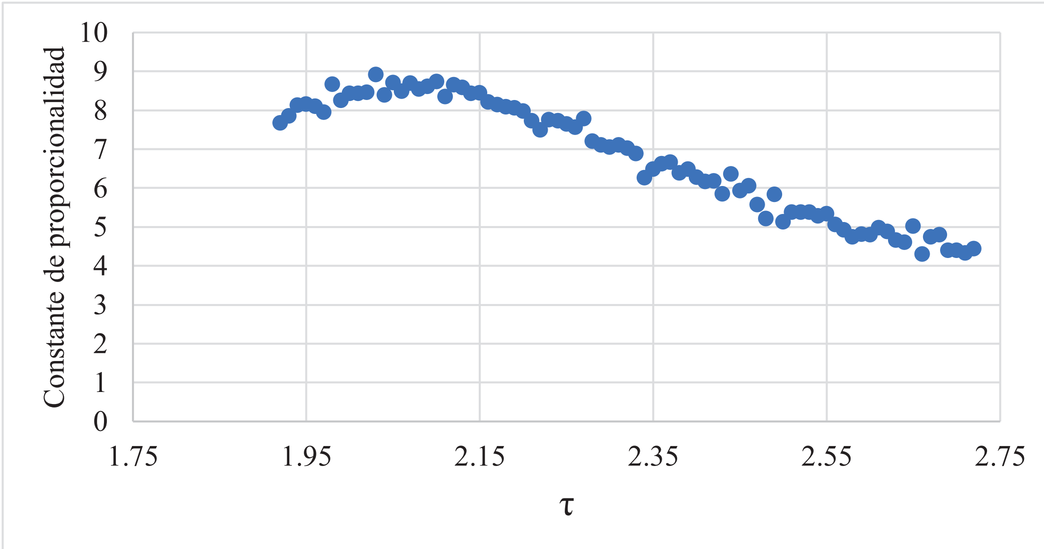


Figura 4.10: Gráfico de dispersión para la constante de proporcionalidad versus τ.

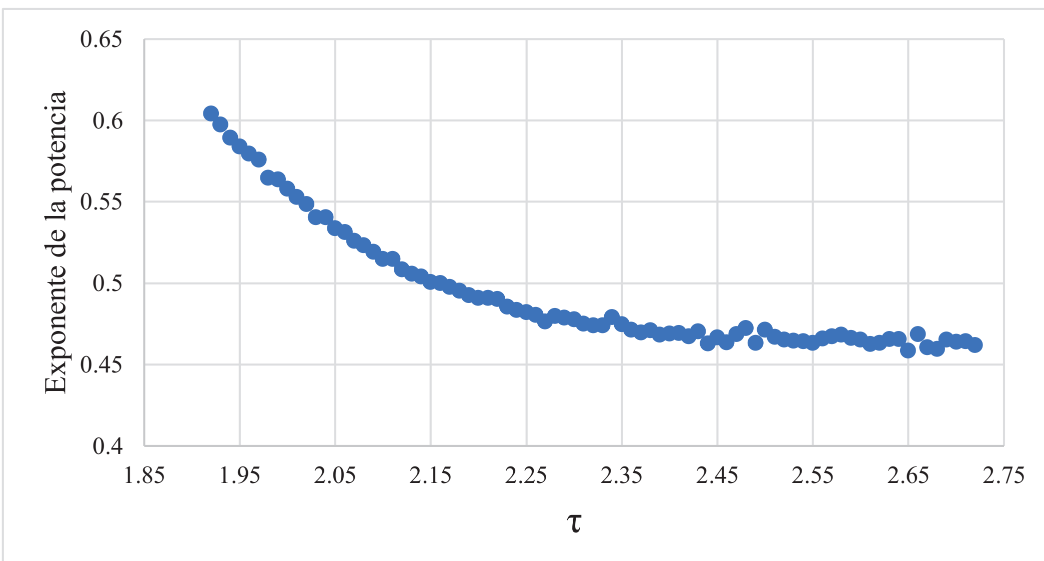


Figura 4.11: Gráfico de dispersión para el exponente de la potencia versus τ.

Se aprecia en el gráfico de la Figura 4.10 que la constante de proporcionalidad es creciente desde $\tau = 1.92$ hasta $\tau = 2.03$ aproximadamente, donde luego empieza a decrecer con el aumento de τ. En el gráfico de la Figura 4.11, se observa que el exponente de la potencia tiene un comportamiento exponencial decreciente y que converger a 0.45 aproximadamente. En el gráfico de la Figura 4.9 se ajusta la función potencial a los datos por medio de la herramienta de líneas de tendencia que proporciona Microsoft Excel. En el caso de la constante de proporcionalidad y el exponente de potencia, esta herramienta no es capaz de ajustar una línea de tendencia con un valor de R^2 aceptable.

4.4.3. Función que Determina la Constante de Proporcionalidad

Para ajustar una función que modela la constante de proporcionalidad a partir τ, se realizó un ajuste manualmente probando diferentes funciones con un comportamiento similar a los datos, hasta encontrar una función con un valor de R^2 aceptable. Al ver la distribución de los datos en el gráfico de la Figura 4.10, se aprecia que se tiene una curva sesgada a la izquierda. Ahora, ¿Qué función tiene un comportamiento similar? Si se tiene una función exponencial, cuyo exponente es negativo de la forma $ln^2(x)$ se obtiene la gráfica que se observa en la Figura 4.12. La función describe un comportamiento similar a la distribución de los datos. Entonces se toma como base para llegar a la verdadera función que se busca.

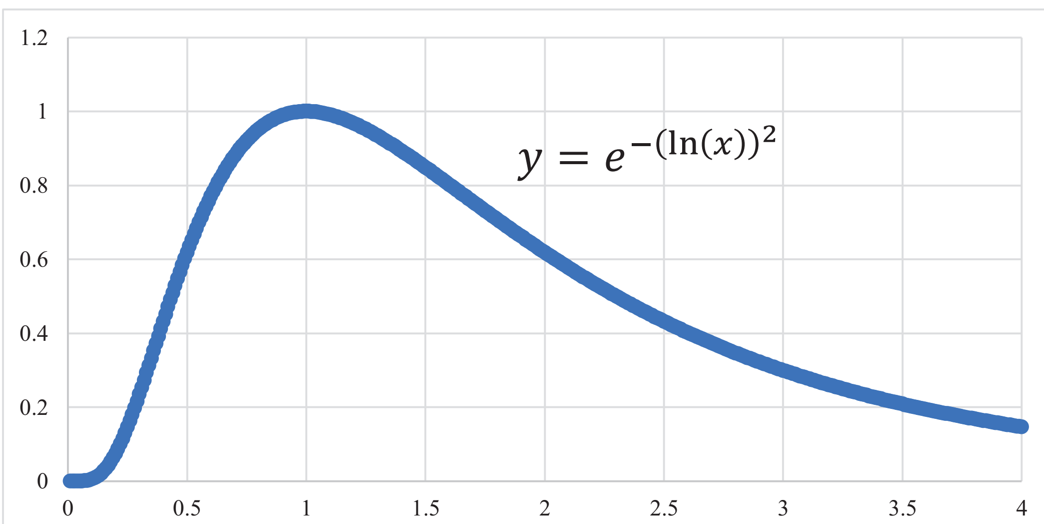


Figura 4.12: Gráfico de una función exponencial cuyo exponente negativo es un logaritmo natural al cuadrado.

De esta forma, a $y = e^{-ln^2(x)}$ se le agregan diferentes parámetros para que se ajuste a los valores que se tiene de la constante de proporcionalidad hasta que se llega a la función que se aprecia en el gráfico de la Figura 4.13. A la función y resultante, se le calcula el valor de R^2 y se obtiene que esta se ajusta a los datos con un valor de R^2=0.975 838 47, que resulta ser bastante bueno. Entonces se define a la fórmula

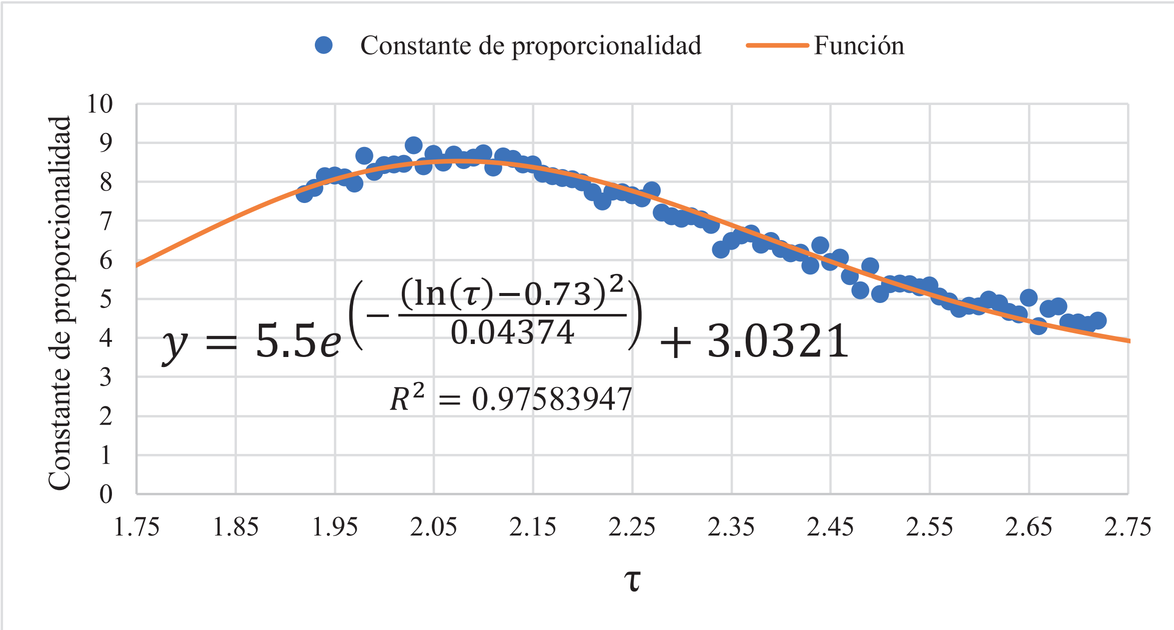


Figura 4.13: Gráfico de dispersión para la constante de proporcionalidad versus τ.

de la Ecuación 4.14 como la función que modela la constante de proporcionalidad teniendo a τ como variable independiente.

$$Constante(\tau) = 5.5e^{\left(-\frac{(ln(\tau)-0.73)^2}{0.043\,74}\right)} + 3.032\,1 \tag{4.14}$$

4.4.4. Función que Determina el Exponente de Potencia

Para ajustar una función que modela el exponente de potencia a partir de τ, se realiza un ajuste manualmente, tal como se realiza para la constante de proporcionalidad. De esta forma se ajusta una función hasta obtener un valor de R^2 aceptable. Al ver la distribución de los datos en el gráfico de la Figura 4.11, se aprecia que una función exponencial con exponente negativa posee el mismo comportamiento como se aprecie en la Figura 4.14.

Entonces, $y = e^{-x}$ se toma como base para modelar la función para el exponente de la potencia. De esta forma, la función se le agregan diferentes parámetros, los cuales son alterados constantemente con diferentes valores hasta que se logra un ajuste a los datos con un valor de R^2 aceptable. Así, se llega a la función que se

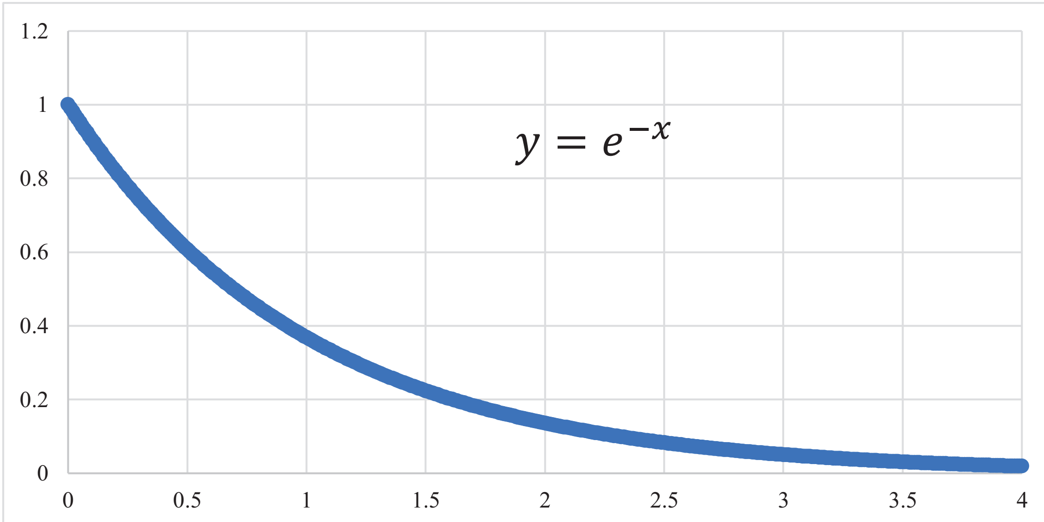


Figura 4.14: Gráfico de una función exponencial con exponente negativo.

aprecia en el gráfico de la Figura 4.15. A la función y resultante, se le calcula el valor de R^2 y se obtiene que el ajuste tiene un valor de R^2=0.994 565 37, que también resulta ser bastante bueno.

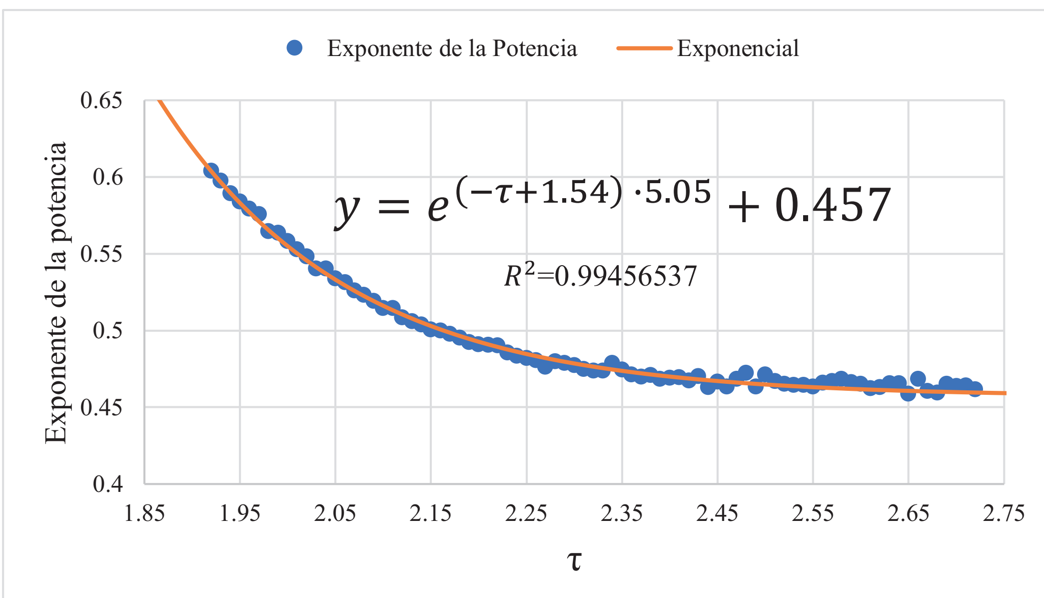


Figura 4.15: Gráfico de dispersión para del exponente de la potencia versus τ.

Entonces, se define la fórmula de la Ecuación 4.15 como la función que modela el exponente de la potencia teniendo τ como variable independiente.

$$Exponente(\tau) = e^{(-\tau+1.54)\cdot 5.05} + 0.457 \tag{4.15}$$

4.4.5. Función que Determina la Aridad Máxima

Dado que se ha modelado una función para la constante de proporcionalidad y otra para el exponente de la potencia, ahora se puede definir una función global que determine la aridad máxima de un grafo, dada la cantidad de nodos N que éste posee y un valor de τ. Entonces, usando las fórmulas de las Ecuaciones 4.14 y 4.15 se define la fórmula de la Ecuación 4.16 que modela la aridad máxima, donde la cantidad de nodos N y un valor de τ específico son las variables independientes.

$$AridadMaxima(N, \tau) = Constante(\tau) \cdot N^{Exponente(\tau)} \tag{4.16}$$

Ahora, para comprobar los resultados que genera la fórmula de la aridad máxima, se lleva a cabo una simulación en Microsoft Excel, en la cual se consideran diferentes tamaños de grafos y diferentes valores de τ. En términos de tamaños de grafos, se considera N={1 000, 10 000, 100 000, 1 000 000, 10 000 000}. Para la variable τ se consideran valores desde 1.92 a 2.72. Básicamente los mismos valores que se usan en el experimento para modelar la aridad máxima.

Entonces, al ejecutar la simulación con los parámetros establecidos, se genera una tabla con las aridades máximas que calcula la función. Un resumen de esta tabla se aprecia en el Cuadro 4.15. Para tener una visualización completa de la tabla, esta se encuentra en el Anexo A.2.4.

Cuadro 4.15: Aridades máximas obtenidas de la fórmula de la Ecuación 4.16.

τ	**1 000**	**10 000**	**100 000**	**1 000 000**	**10 000 000**
1.92	506.59	2 034.29	8 168.95	32 803.47	131 726.59
1.93	487.23	1 924.26	7 599.60	30 013.54	118 534.22
1.94	469.42	1 824.81	7 093.69	27 575.77	107 197.13
1.95	452.98	1 734.60	6 642.36	25 435.76	97 401.91
⋮	⋮	⋮	⋮	⋮	⋮
2.71	98.41	283.64	817.49	2 356.13	6 790.71
2.72	97.18	280.00	806.75	2 324.47	6 697.40

Como los datos de la tabla fueron generados con los mismos parámetros del experimento inicial para modelar la aridad máxima, la tabla del Cuadro 4.15 es equivalente a la tabla del Cuadro 4.13 que posee las aridades máximas promedio que se obtienen de los valores empíricos. Ahora, se comparan los valores de ambas tablas, ya que son correspondientes los valores de cada celda de una tabla con la otra. Así, se calcula la diferencia que hay entre los datos generados por la fórmula menos los datos obtenidos en el experimento. Esto se registra en la tabla del Cuadro 4.16. La tabla completa se encuentra en el Anexo A.2.5. En la tabla se aprecia

Cuadro 4.16: Diferencia entre la aridad registrada en el Cuadro 4.15 y el Cuadro 4.13.

τ	**1 000**	**10 000**	**100 000**	**1 000 000**	**10 000 000**
1.92	-22.11	314.79	-39.55	-5 656.13	14 521.39
1.93	-27.37	272.76	-195.10	-5 762.46	11 271.32
1.94	-41.08	247.31	-193.21	-5 689.43	8 867.93
1.95	-40.82	238.60	-275.74	-5 461.14	7 150.31
⋮	⋮	⋮	⋮	⋮	⋮
2.71	-17.89	15.64	-59.01	-731.17	-380.99
2.72	-22.92	18.80	-65.85	-717.63	-459.40

que hay diferencias positivas y negativas. Las diferencias positivas indican que la aridad máxima calculada por la fórmula de la Ecuación 4.16 supera a la aridad promedio registrada en el Cuadro 4.13. Entonces, las diferencias negativas indican que la aridad máxima calculada por la fórmula no alcanza la aridad promedio. Como la cantidades de cada columna varían bastante por los tamaños de los grafos, se transforman a porcentajes para analizar mejor los datos. En el Cuadro 4.17 se tiene

Cuadro 4.17: Diferencia entre la aridad registrada en el Cuadro 4.15 y el Cuadro 4.13 en porcentaje.

τ	**1 000**	**10 000**	**100 000**	**1 000 000**	**10 000 000**
1.92	-4.2 %	18.3 %	-0.5 %	-14.7 %	12.4 %
1.93	-5.3 %	16.5 %	-2.5 %	-16.1 %	10.5 %
1.94	-8.0 %	15.7 %	-2.7 %	-17.1 %	9.0 %
1.95	-8.3 %	15.9 %	-4.0 %	-17.7 %	7.9 %
⋮	⋮	⋮	⋮	⋮	⋮
2.69	-15.3 %	7.0 %	-6.4 %	-23.9 %	-5.3 %
2.70	-17.1 %	9.2 %	-8.0 %	-22.9 %	-5.8 %
2.71	-15.4 %	5.8 %	-6.7 %	-23.7 %	-5.3 %
2.72	-19.1 %	7.2 %	-7.5 %	-23.6 %	-6.4 %

a modo de resumen la diferencia entre las aridades en porcentaje. El Cuadro con todos los porcentajes se encuentra en el Anexo A.2.6. Cabe destacar que lo ideal es que los porcentajes sean 0 % o lo más cercano a este valor, ya que esto indica que la fórmula no genera aridades muy altas o bajas con respecto a las que se obtiene en el experimento para modelar la aridad máxima, que son los datos a los que se ajusta la función. El porcentaje más bajo que se registra es de aproximadamente -24 % en grafos de 1 000 000 de nodos, mientras que el porcentaje más alto es 26 % en grafos de 10 000 nodos. Con esto, se decide dar prioridad a los casos en que se genera una aridad por debajo de lo esperado. Se establece que cada valor que genera la fórmula de la Ecuación 4.16 se multiplica por 1.24 para cubrir la diferencia. Esto, con el fin de los grafos generados posean la cantidad de aristas que se esperan para cada caso.

4.5. Algoritmo 3

Este método de generación de grafos, al igual que el Algoritmo 2 se basa en el cálculo de un arreglo de probabilidades acumuladas y un arreglo de frecuencias para cada aridad, de tal forma que a partir de estas estructuras se genera el arreglo de aridades que contiene la distribución de grados de los nodos. La diferencia es que este método implementa la fórmula de la Ecuación 2.9 para la generación de números aleatorios con distribución de la ley de potencia y la fórmula de la Ecuación 4.16 que genera la aridad máxima para un grafo dada la cantidad de nodos N y un valor de τ. Este método lleva a cabo el proceso de generación mediante las funciones GeneratePbb, GenerateDegrees y GenerateArrayPowerLaw, que se presentan en los Algorithms 5, 6 y 7, respectivamente.

4.5.1. Generación de Probabilidades Acumuladas

A continuación, se describen los pasos y diseño del método para la generación del arreglo de probabilidades acumuladas que hace uso el Algoritmo 3.

El parámetro de entrada de la función GeneratePbb es el número de nodos N y la aridad máxima $LIMIT$ calculada con la formula de la Ecuación 4.16. La salida es un arreglo PBB de longitud $LIMIT$, tal que $PBB[i]$ contiene la probabilidad acumulada que corresponde a la aridad i, donde $0 \leq i < LIMIT$.

La función GeneratePbb, al igual que en el Algoritmo 2, comienza con la definición del valor de *sumaPBB* y *pbAcumulada*, como también la inicialización del arreglo

PBB. Se define la asignación de $PBB[0] = 0$ para no generar nodos con aridad 0 y debido a que la probabilidad para esta aridad indefine la fórmula de la Ecuación 2.3 de la definición de la ley de potencia. Se asigna el valor de $sumaPBB = 0$ y $pbAcumulada = 0$.

Algorithm 5: GeneratePbb(int N, int $LIMIT$)

Input: $N \leftarrow$ Número de nodos del grafo. $LIMIT \leftarrow$ Aridad máxima que establece el tamaño del arreglo de probabilidades acumuladas.

Output: Un arreglo PBB, tal que $PBB[i]$ contiene la probabilidad acumulada que corresponde a la aridad i.

```
1  begin
2      double sumaPBB ← 0
3      double pbAcumulada ← 0
4      double PBB[] ← new double[LIMIT]
5      PBB[0] ← 0
6      for long k ← 1 to N − 1 do
7          double pbb ← C · Math.pow(k, −τ)
8          if k < LIMIT then
9              PBB[k] ← pbb
10         end
11         sumaPBB ← sumaPBB + pbb
12     end
13     for long i ← 1 to LIMIT − 1 do
14         double pb ← PBB[i]
15         double pbAjustada ← pb/sumaPBB
16         pbAcumulada ← pbAcumulada + pbAjustada
17         PBB[i] ← pbAcumulada
18     end
19     return PBB
20 end
```

En el paso de generar la probabilidad (Línea 7) para una aridad k (valor posible dado el volumen N del grafo) se utiliza la fórmula de la Ecuación 2.3 de la definición de la ley de potencia. En esta sección cabe mencionar que se utiliza la fórmula de la Ecuación 4.7 para la constante C y el valor de τ es por defecto 2, pero puede ser un parámetro de entrada ingresado al momento de ejecutar el programa. Tanto la constante C como τ son variables privadas de la clase en la que se han declarado las funciones. Una vez calculada la probabilidad para la aridad k, se verifica que k sea menor a la aridad máxima $LIMIT$ (Línea 8), ya que no se generan nodos con aridad superior la indicada por este valor. Así se asigna el valor de la probabilidad en $PBB[k]$ (Línea 9) y este se suma a la variable $sumaPBB$ para obtener la suma de todas las probabilidades (Línea 11). El valor de $sumaPBB$ si considera todos los valores de k, ya que con este se ajustan las probabilidades (Línea 15). El proceso del calculo de probabilidades se realiza desde $k = 1$ hasta $N - 1$ (Línea 6 a la 12). La representación gráfica de $PBB[]$ una vez finalizada la primera estructura for, se aprecia en el Cuadro 4.18 donde la probabilidad para $LIMIT$ se aprecia como 0.000 debido a que es un valor decimal muy bajo. Estas probabilidades se obtienen de la generación de un grafo no dirigido de 1 000 nodos. El valor de la variable $sumaPBB$ equivale a calcular $C\sum_{k=1}^{n-1} k^{-\tau}$.

Cuadro 4.18: Arreglo de probabilidades Algoritmo 3.

0	0.667	0.167	0.074	0.042	0.027	...	0.000
0	1	2	3	4	5	...	LIMIT-1

En el paso de ajustar la probabilidad (Línea 15) y calcular la probabilidad acumulada (Línea 16), primero se obtiene la probabilidad de la aridad i (Línea 14). Este valor se ajusta al dividirlo por la suma de probabilidades ($sumaPBB$), de tal forma que se obtiene la probabilidad ajustadas para la aridad i. Con el valor ajustado, este se suma a la variable *pbAcumulada*, el cual a su vez se asigna al arreglo en $PBB[i]$ (Línea 17), lo que reemplaza la probabilidad de i por su probabilidad acumulada.

Los pasos entre la Línea 14 y 17 se repiten $LIMIT$-1 veces desde la aridad $i = 1$ hasta $i = LIMIT$-1. El arreglo de probabilidades acumuladas producido se retorna en la Línea 17, el que contiene la probabilidad acumulada de cada aridad i. La representación gráfica del arreglo final se aprecia en el Cuadro 4.19. Los valores del arreglo se obtienen de la generación de un grafo no dirigido de 1 000 nodos.

Cuadro 4.19: Arreglo de probabilidades acumuladas Algoritmo 3.

0	0.608	0.760	0.828	0.866	0.890	...	0.999
0	1	2	3	4	5	...	LIMIT-1

4.5.2. Generación del Arreglo de Frecuencias

A continuación, se describen los pasos y diseño del método para la generación del arreglo de frecuencia para cada aridad, del cual hace uso el Algoritmo 3.

Los parámetros de entrada de la función GenerateDegrees son el número de nodos N, la aridad máxima $LIMIT$ y un arreglo de probabilidades acumuladas $PBBA$, tal que $PBBA[i]$ contiene la probabilidad acumulada que corresponde a la aridad i. La salida es un arreglo $FREC$ de longitud $LIMIT$, tal que $FREC[i]$ contiene la cantidad de nodos que corresponde a la aridad i, donde $0 \leq i < LIMIT$.

La función GenerateDegrees comienza con la inicialización del arreglo $FREC$ y la definición del valor de x_0 al que se le asigna 0 y x_1 al que se le asigna la probabilidad acumulada de la aridad máxima. Estas dos variables son el limite inferior y superior de la fórmula de la Ecuación 2.9 para generar números aleatorios con distribución de la ley de potencia, por lo que se les asigna el rango de probabilidades contenida en $PBBA$. A cada posición del arreglo $FREC$ se le asigna el valor 0 (Línea 6). Esto es debido a que la frecuencia asociada a cada aridad se produce de forma aleatoria. Entonces para diferentes aridades con probabilidades acumuladas muy altas podrían no generarse nodos.

En el paso de obtener un número aleatorio a partir de la formula que genera números aleatorios con distribución de la ley de potencia (Línea 9), se genera un número entre 0 y $PBBA[LIMIT - 1]$ que se almacena en la variable r. Dado que se tiene el arreglo de probabilidades acumuladas $PBBA$, se busca el rango de valores en el cual se encuentra r. Para esto se obtiene la probabilidad $pbaA$ (Línea 12) y $pbaB$ (Línea 13), donde $pbaA < pbaB$. Con estas probabilidades acumuladas se verifica si $pbaB > r$ y $r > pbaA$ (Línea 14). En el caso verdadero, esto indica que r se encuentra en el rango de $PBBA[j]$ y $PBBA[j - 1]$, donde $1 \leq j < N$. Entonces, en el arreglo $FREC$ se aumenta en 1 la cantidad de nodos asociada a la aridad j (Línea 15) y finaliza la búsqueda del rango al que pertenece r. En el caso falso, se seguirá aumentando el valor del índice j, que a su vez representa la aridad, hasta que se encuentre el rango de probabilidades acumuladas al que pertenece r.

En la Línea 16 se tiene la variable global *edges_t* que lleva un conteo de la cantidad total de arista que se deben generar a medida que se va actualizando el arreglo de frecuencias (Línea 15). En la Línea 21, se tiene la variable global *edge_type* la que indica el tipo de grafo que se esta generando (0 si es no dirigido, 1 si es dirigido). Con esta variable se determina si seguir generando nodos para diferentes aridades de acuerdo al tipo de grafo. Si el grafo es no dirigido y se llega a las N generaciones de números aleatorios, se verifica si la cantidad de *edges_t* es menor a $edgesUndirect$ (variable global que contiene el doble de la cantidad de aristas estimada por la

Algorithm 6: GenerateDegrees(int N, int $LIMIT$, double $PBBA[]$)

Input: $N \leftarrow$ Número de nodos del grafo. $LIMIT \leftarrow$ Aridad máxima que establece un límite para la búsqueda de probabilidades. Un arreglo $PBBA$, tal que $PBBA[i]$ contiene la probabilidad acumulada que corresponde a la aridad i.

Output: Un arreglo $FREC$, tal que $FREC[i]$ contiene la cantidad de nodos que corresponde a la aridad i.

begin
  long $FREC[] \leftarrow$ new Long[$LIMIT$]
  double $x_0 \leftarrow 0$
  double $x_1 \leftarrow PBBA[LIMIT-1]$
  for long $k \leftarrow 0$ **to** $LIMIT-1$ **do**
    $FREC[k] \leftarrow 0$
  end
  for long $i \leftarrow 1$ **to** N **do**
    double $r \leftarrow \left[\left(x_1^{\tau+1} - x_0^{\tau+1}\right) \cdot Math.random() + x_0^{\tau+1}\right]^{\frac{1}{\tau+1}}$
    long $j \leftarrow 1$
    while $j < LIMIT$ **do**
      double $pbaA \leftarrow PBBA[j-1]$
      double $pbaB \leftarrow PBBA[j]$
      if $pbaB > r \wedge r > pbaA$ **then**
        $FREC[j] \leftarrow FREC[j] + 1$
        $edges\_t \leftarrow edges\_t + j$
        $j \leftarrow N$
      end
      $j \leftarrow j + 1$
    end
    if $edge\_type = 0$ **then**
      if $i = N \wedge edges\_t < edgesUndirect$ **then**
        $i \leftarrow i - 1$
      end
    else
      if $edges\_t \geq edges$ **then**
        $i \leftarrow N + 1$
      end
    end
  end
  return $FREC$
end

fórmula Ecuación 4.10). Si la condición de la Línea 22 es verdadera se seguirán generando números aleatorios para aumentar la cantidad de nodos por aridad y así alcanzar la cantidad de aristas estimada. En el caso de un grafo dirigido, si la cantidad de $edges\_t$ es mayor a $edges$ (variable global que contiene la cantidad de aristas estimada por la fórmula Ecuación 4.10) se termina la generación de números aleatorios para no generar un exceso de aristas con respecto a la cantidad estimada.

El paso de obtener un número aleatorio (Línea 9) se realiza a lo menos N veces si el grafo es no dirigido y a lo más N veces si es dirigido. Los pasos entre la Línea 12 y 19 se realiza a lo más $LIMIT - 1$ veces, puesto que el rango de valores al que pertenece r puede encontrarse antes de llegar al final del arreglo de probabilidades acumuladas. El arreglo de frecuencias se retorna en la Línea 31, el cual contiene la cantidad de nodos correspondientes a cada aridad i ($FREC[i]$).

4.5.3. Generación del Arreglo de Aridades

A continuación, se describen los pasos y diseño del método para la generación del arreglo de aridades, del cual hace uso el Algoritmo 3.

El parámetro de entrada de la función GenerateArrayPowerLaw es el número de nodos N. La salida es un arreglo D de longitud N, tal que $D[i]$ contiene el grado del nodo i, donde $0 \leq i < N$.

La función GenerateArrayPowerLaw comienza con la inicialización del arreglo de aridades D. Se define la variable $limit$ que almacena la aridad máxima calculada por la fórmula de la Ecuación 4.16 (Línea 3). Se define arreglo $PBBA$ el cual recibe el arreglo de probabilidades acumuladas que retorna la función GeneratePbb (Línea 4) en base a N y $limit + 1$. Se define el arreglo $frec$ el cual recibe el arreglo de frecuencias que retorna la función GenerateDegrees (Línea 5) en base a N, $limit + 1$ y $PBBA$.

Como el arreglo $frec$ contiene la cantidad de nodos para cada aridad, en la Línea 8 se obtiene la cantidad de nodos de la aridad k (el valor se almacena en la variable $nodosPorAridad$). Se comprueba si la cantidad de nodos es mayor a 0 (Línea 9). En caso de ser verdadero, se asigna la aridad k al nodo h (Línea 11). Se está añadiendo la aridad k a un solo nodo en el arreglo de aridades. Por esto último, este paso se repite la cantidad de veces indicada por la cantidad de nodos con aridad k. Cabe mencionar que el arreglo de frecuencias se recorre desde $limit$ hasta 1. Esto les da

Algorithm 7: GenerateArrayPowerLaw(int N)

```
   Input: N ← Número de nodos del grafo.
   Output: Un arreglo D, tal que D[i] contiene el grado del nodo i.
1  begin
2      long D[] ← new long[N]
3      int limit ← AridadMaxima(N, τ)
4      double PBBA[] ← GeneratePbb(N, limit + 1)
5      long frec[] ← GenerateDegrees(N, limit + 1, PBBA)
6      long h ← 0
7      for long k ← limit to 1 do
8          double nodosPorAridad ← frec[k]
9          if nodosPorAridad > 0 then
10             long cont ← 0
11             while cont < nodosPorAridad do
12                 D[h] ← k
13                 h ← h + 1
14                 if h ≥ N then break;
15                 cont ← cont + 1
16             end
17         end
18         if h ≥ N then break;
19     end
20     return D
21 end
```

prioridad a los nodos con mayor aridad para ser agregados al arreglo de aridades D.

Se debe tener en cuenta que si el índice h, que se utiliza para recorrer el arreglo de aridades D, llega a igualar a N (que es la cantidad de espacios disponibles en D) antes de que se termine de recorrer el arreglo de frecuencias $frec$, se da por terminado el proceso de producción del arreglo de aridades, por lo que se pierden frecuencias asociadas a aridades pequeñas.

El arreglo de aridades producido se retorna en la Línea 20, el que contiene el número de aristas de cada nodo i ordenada de mayor a menor. De este modo el nodo 0 tiene la mayor aridad, mientras que el nodo $N-1$ contiene la menor aridad. La representación gráfica del arreglo que se retorna se aprecia en el Cuadro 4.20. Este arreglo tiene datos de un grafo no dirigido de 1 000 nodos indexados del cero en adelante generado con el método descrito.

Cuadro 4.20: Arreglo de aridades Algoritmo 3.

472	385	273	209	186	178	167	153	147	...	1
0	1	2	3	4	5	6	7	8	...	999

4.5.4. Experimentos Preliminares Algoritmo 3

En esta sección se evalúan los resultados que genera el Algoritmo 3, lo que permite analizar la eficiencia del método y características de los grafos producidos. Al igual que los experimentos preliminares del Algoritmo 1 y 2, el objetivo de estos es la búsqueda de fallas o falencias en el método de generación para su corrección o que den paso crear nuevos algoritmos que resuelvan los problemas de este método.

Resultados

Cada experimento corresponde a una combinación de tamaño del grafo y tipo de aristas. Se registra el tiempo de ejecución y el número de aristas generada. En general, se generan ocho grafos, con un tamaño máximo de un millón de nodos.

Cuadro 4.21: Grafos generados en la evaluación preliminar del Algoritmo 3.

GraphID	**Tipo de Grafo**	**Número de Nodos**	**Número de Aristas**	**Tiempo (ms)**
UDG1K	No Dirigido	1 000	5 168	22
DG1K	Dirigido	1 000	5 529	20
UDG10k	No Dirigido	10 000	65 672	51
DG10k	Dirigido	10 000	70 068	48
UDG100k	No Dirigido	100 000	809 055	296
DG100k	Dirigido	100 000	853 518	255
UDG1M	No Dirigido	1 000 000	9 631 437	3 069
DG1M	Dirigido	1 000 000	10 010 100	2 935

El Cuadro 4.21 presenta los resultados obtenidos. El tiempo de ejecución más bajo fue de 20 milisegundos que corresponde a la generación de un grafo dirigido de mil nodos. Por otro lado, el tiempo de ejecución más alto es 3 069 milisegundos que corresponden a la generación de un grafo no dirigido de un millón de nodos. Se aprecia empíricamente que en la mayoría de casos toma más tiempo generar grafos no dirigidos.

La fórmula de la Ecuación 4.11 indica la cantidad estimada de aristas para los tamaños de grafos generados, ya que se crean con $\tau = 2$. De acuerdo a la estimación de aristas para cada caso, todos los grafos generados han superado la cantidad de aristas estimada.

Cuadro 4.22: Aridad máxima de los diferentes grafos generados.

GraphID	**Tipo de Grafo**	**Número de Nodos**	**Aridad Máxima**
UDG1K	No Dirigido	1 000	445
DG1K	Dirigido	1 000	439
UDG10k	No Dirigido	10 000	1 576
DG10k	Dirigido	10 000	1 650
UDG100k	No Dirigido	100 000	6 156
DG100k	Dirigido	100 000	6 126
UDG1M	No Dirigido	1 000 000	22 143
DG1M	Dirigido	1 000 000	21 894

En el Cuadro 4.22 se muestra la aridad máxima obtenida en cada grafo generado. Los valores de aridad máxima que establece la fórmula de la Ecuación 4.16 al ser multiplicada por 1.24 con $\tau = 2$ y $N = \{1\,000,\ 10\,000,\ 100\,000,\ 1\,000\,000\}$ son 479, 1 721, 6 176 y 22 167, respectivamente. Se observa que ningún grafo superó su respectiva aridad máxima y que se generaron valores cercanos a los indicados por la fórmula.

En la Figura 4.16 se muestra la distribución de aridades de un grafo no dirigido de 100 000 nodos que se generó en el experimento. En el gráfico se aprecia que ahora no se generan nodos con aridades muy altas como si ocurre en los Algoritmos 1 y 2. De hecho, la aridad máxima de este grafo es 6 156 como se registra en el Cuadro 4.22. También, gran parte de la distribución de las aridades es similar a la vista en el gráfico de de la Figura 4.8 del grafo de 100 000 nodos generado por R^{3}MAT.

En el gráfico se observa que la cantidad de nodos con aridad uno a decaído con respecto a los Algoritmos 1 y 2. El hecho de que disminuye la preferencia por la aridad uno es un resultado esperado, debido a la fórmula que altera la distribución de los números aleatorios. Pero la cantidad de nodos con aridad uno que se tienen es mucho menor, puesto que dicha cantidad en el arreglo de frecuencia es de 25 922 nodos, pero la cantidad que se observa en el grafo es de 6 143. ¿Por qué pasa esto? Pues esto se debe a que aumenta la preferencia de las aridades más altas con la fórmula para los números aleatorios, por lo que hay más nodos en la zona de aridades altas. A

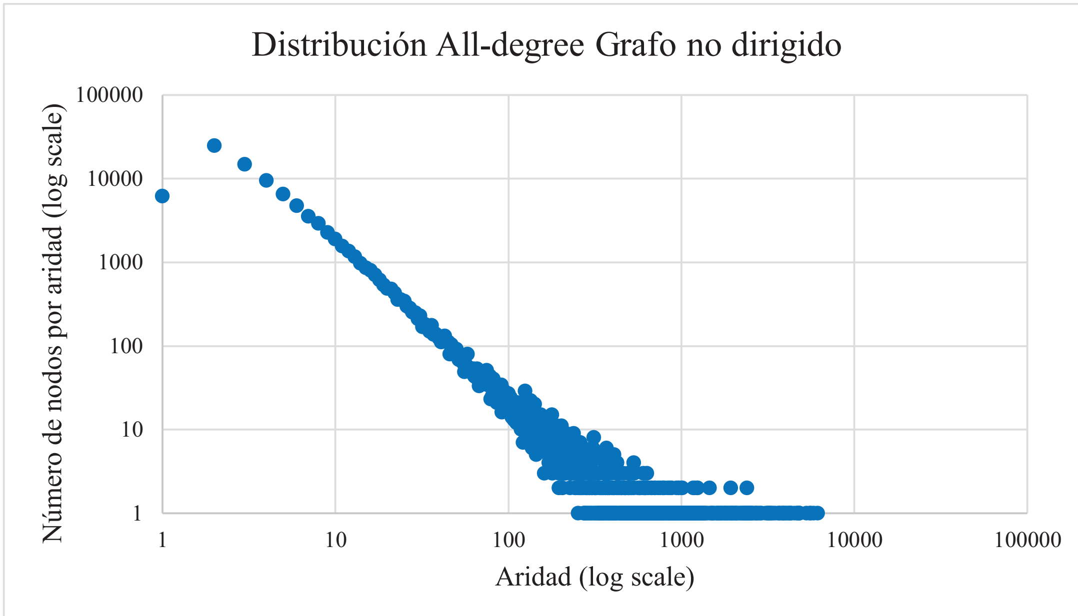


Figura 4.16: Distribución de la aridad de un grafo de 100 000 nodos Algoritmo 3.

su vez, el Algoritmo GenerateArrayPowerLaw (Algorithm 7) al momento de generar el arreglo de aridades, lo hace desde la mayor aridad a la menor. Entonces, como en el arreglo se le da la prioridad a las aridades altas y ahora estas consumen más espacios dentro de la estructura, debido a que se aumentó su preferencia, cuando hay que agregar los nodos con aridad uno, la estructura no da abasto para agregarlos a todos.

5. Evaluación Experimental

El objetivo de la evaluación experimental es probar y comparar la eficiencia y realismo de los algoritmos propuestos en el Capítulo 4. Además, también se realiza una comparación de rendimiento del Algoritmo 3 con R^3MAT y las alternativas secuenciales como lo son SNAP [17] y PaRMAT [23]. Por ende, este capítulo contiene el desarrollo de la metodología de evaluación, los resultados experimentales y la discusión correspondiente.

5.1. Fase Experimental 1

La Fase de Evaluación Experimental 1 se basa en la generación de grafos. Cada experimento combina tres variables: método de generación, tamaño del grafo y tipo de grafo. Se consideran los tres métodos presentados en el Capítulo 4 de este documento. En términos de tamaño, se consideran grafos de 1 000, 10 000, 100 000, 1 millón, 10 millones, 100 millones y 250 millones de nodos. La variable tipo de grafo indica la generación de aristas dirigidas o no dirigidas.

En base a estas variables, se evalúan los métodos de generación en términos de eficiencia y realismo. La eficiencia se mide como el tiempo transcurrido (o tiempo de ejecución) junto con el aumento creciente del número de nodos. El objetivo es determinar el método más rápido y más lento dado diferentes tamaños de grafos. Para hacerlo, se usa la función integrada en Java System.currentTimeMillis().

Como se menciona con anterioridad en el documento, R-MAT puede generar grafos con propiedades que se producen en redes reales, en particular en la distribución de la ley de potencia. Por lo tanto, el realismo de los métodos está dado por su capacidad de producir una distribución de ley de potencia para cualquier tamaño del grafo. Para verificar la noción de realismo, se calcula la estadística Kolmogorov-Smirnov.

En base a las métricas anteriores, se selecciona un solo algoritmo, con el cual se realiza la siguiente fase de experimentación para compararlo con R^{3}MAT y las alternativas secuenciales como lo son SNAP [17] y PaRMAT [23].

5.1.1. Configuración Experimental

Todos los experimentos de la Fase 1 se ejecutan en una máquina virtual creada en Google Cloud Platform. Se utiliza una máquina con una CPU IntelSkylake con 16 CPU virtuales, 60 Gb de memoria (RAM) y 65 GB de memoria secundaria (SSD). El sistema operativo es Debian GNU/Linux 10. Se instala Java Openjdk versión 11.0.7. Se genera una aplicación Java para cada uno de los algoritmos anteriormente descritos en el Capítulo 4 y se envían a la máquina virtual. Los tres algoritmos poseen una interfaz basada en jar que permite generar grafos desde la línea de comandos, ejecutando la instrucción:

```
java -jar GraphGenerador<Número>.jar
               <N> <T> <F> <TAU>
```

donde $< Número >$ define el algoritmo que se va utilizar, $< N >$ es la cantidad de nodos del grafo (Obligatorio), $< T >$ indica el tipo de grafo (0 para no dirigido que es el valor por defecto, 1 para dirigido), $< F >$ indica el formato de los datos de salida (0 para la lista de aristas, 1 para graphml) y $< TAU >$ es el valor para el exponente de la potencia (sólo acepta valores positivos y se recomiendan dentro del rango [1.92, 2.72]). Los grafos para esta fase de la evaluación experimental se generan configurando los parámetros $< N >$ y $< T >$. Como formato de salida, se utiliza la representación de lista de arista que es el formato por defecto y para el exponente se utiliza $\tau = 2$ (valor por defecto) que se encuentra dentro del rango recomendado.

5.1.2. Resultados Experimentales Fase 1

Cada experimento corresponde a una combinación específica de método de generación, tamaño del grafo y tipo de arista. Para cada experimento, cada aplicación se ejecuta 3 veces, se registra el tiempo promedio de ejecución, el número de aristas más alto y el tamaño del archivo de salida más pesado. Las tablas de los Cuadros 5.2, 5.3 y 5.4 muestran un resumen de los grafos generados. En términos generales, se generan catorce tipos de grafos, con un tamaño máximo de 250 millones de nodos.

Cuadro 5.1: Número de aristas estimadas para $\tau = 2$.

Número de Nodos	**Número de Aristas Estimadas**
1 000	4 990
10 000	65 250
100 000	806 009
1 000 000	9 595 150
10 000 000	111 302 071
100 000 000	1 266 526 383
250 000 000	3 319 031 079

Al utilizar $\tau = 2$, el número de aristas está dado por la fórmula de la Ecuación 4.11. Dado los parámetros establecidos, se determina la cantidad de aristas esperadas para cada caso, lo que se muestra en el Cuadro 5.1.

Cuadro 5.2: Grafos generados en la evaluación experimental con el Algoritmo 1.

GraphID	**Tipo de Grafo**	**Número de Nodos**	**Número de Aristas**	**Tamaño del archivo (bytes)**
UDG1K	No Dirigido	1 000	2 967	16 852
DG1K	Dirigido	1 000	6 157	38 843
UDG10k	No Dirigido	10 000	32 721	219 747
DG10k	Dirigido	10 000	73 274	553 750
UDG100k	No Dirigido	100 000	422 785	3 208 840
DG100k	Dirigido	100 000	805 431	7 169 750
UDG1M	No Dirigido	1 000 000	5 022 802	43 636 784
DG1M	Dirigido	1 000 000	10 969 895	109 511 240
UDG10M	No Dirigido	10 000 000	54 364 490	537 996 603
DG10M	Dirigido	10 000 000	113 826 751	1 291 831 799
UDG100M	No Dirigido	100 000 000	658 648 148	6 932 165 680
DG100M	Dirigido	100 000 000	1 212 399 045	15 406 239 541
UDG250M	No Dirigido	250 000 000	1 643 894 149	18 527 339 255
DG250M	Dirigido	250 000 000	3 443 251 139	45 060 733 249

Eficiencia

La tabla del Cuadro 5.5 presenta los tiempos de ejecución de los tres algoritmos de generación. El tiempo de ejecución más bajo es $t = 7$ ms, que corresponden al algoritmo 1, cuando genera un grafo no dirigido de 1 000 nodos. El tiempo de

Cuadro 5.3: Grafos generados en la evaluación experimental con el Algoritmo 2.

GraphID	**Tipo de Grafo**	**Número de Nodos**	**Número de Aristas**	**Tamaño del archivo (bytes)**
UDG1K	No Dirigido	1 000	4 523	26 982
DG1K	Dirigido	1 000	6 020	37 934
UDG10k	No Dirigido	10 000	60 085	427 617
DG10k	Dirigido	10 000	74 648	559 387
UDG100k	No Dirigido	100 000	763 560	6 153 435
DG100k	Dirigido	100 000	977 059	8 458 346
UDG1M	No Dirigido	1 000 000	9 058 915	85 481 507
DG1M	Dirigido	1 000 000	9 440 164	94 854 971
UDG10M	No Dirigido	10 000 000	107 368 690	1 089 942 064
DG10M	Dirigido	10 000 000	108 235 290	1 221 784 843
UDG100M	No Dirigido	100 000 000	1 211 190 297	14 037 819 654
DG100M	Dirigido	100 000 000	1 135 697 577	14 387 209 565
UDG250M	No Dirigido	250 000 000	3 177 728 620	39 424 790 065
DG250M	Dirigido	250 000 000	2 769 839 061	36 984 352 695

Cuadro 5.4: Grafos generados en la evaluación experimental con el Algoritmo 3.

GraphID	**Tipo de Grafo**	**Número de Nodos**	**Número de Aristas**	**Tamaño del archivo (bytes)**
UDG1K	No Dirigido	1 000	5 345	32 702
DG1K	Dirigido	1 000	5 573	35 158
UDG10k	No Dirigido	10 000	65 461	508 264
DG10k	Dirigido	10 000	70 056	544 092
UDG100k	No Dirigido	100 000	809 338	7 477 082
DG100k	Dirigido	100 000	850 361	7 847 960
UDG1M	No Dirigido	1 000 000	9 637 627	103 600 519
DG1M	Dirigido	1 000 000	10 019 287	107 178 852
UDG10M	No Dirigido	10 000 000	111 932 512	1 359 912 208
DG10M	Dirigido	10 000 000	115 464 055	1 398 244 703
UDG100M	No Dirigido	100 000 000	1 273 325 547	17 370 644 366
DG100M	Dirigido	100 000 000	1 306 527 631	17 720 796 918
UDG250M	No Dirigido	250 000 000	3 337 637 291	47 538 272 197
DG250M	Dirigido	250 000 000	3 418 869 542	48 451 309 944

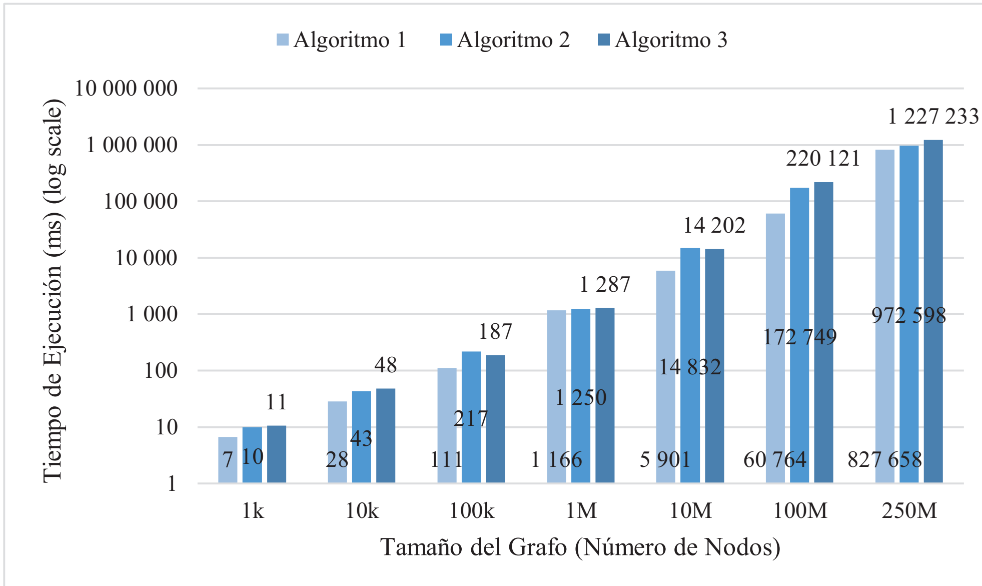
ejecución más alto es t =1 227 233 ms, lo que se obtiene mediante el Algoritmo 3 para generar un grafo no dirigido de 250 millones de nodos. Para comparar la eficiencia de los métodos, se incluye la Figura 5.1. Esta figura muestra los tiempos

Cuadro 5.5: Resultados de los tiempos de ejecución de cada algoritmo.

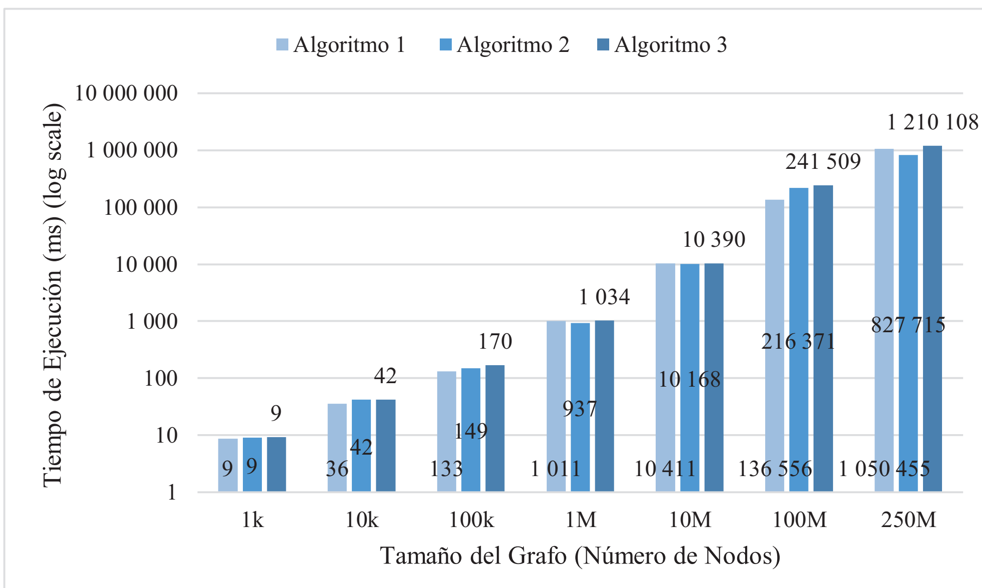
GraphID	**Tipo de Grafo**	**Algortimo 1 (ms)**	**Algoritmo 2 (ms)**	**Algoritmo 3 (ms)**
UDG1K	No Dirigido	7	10	11
DG1K	Dirigido	9	9	9
UDG10k	No Dirigido	28	43	48
DG10k	Dirigido	36	42	42
UDG100k	No Dirigido	111	217	187
DG100k	Dirigido	133	149	170
UDG1M	No Dirigido	1 166	1 250	1 287
DG1M	Dirigido	1 011	937	1 034
UDG10M	No Dirigido	5 901	14 843	14 202
DG10M	Dirigido	10 411	10 168	10 390
UDG100M	No Dirigido	60 764	172 749	220 121
DG100M	Dirigido	136 556	216 371	241 509
UDG250M	No Dirigido	827 658	972 598	1 227 233
DG250M	Dirigido	1 050 455	827 715	1 210 108

de ejecución de los tres algoritmos para generar grafos de diferentes tamaños (de 1 000 a 250 millones de nodos), ya sean no dirigidos (ver Figura 5.1 (a)) como no dirigidos (ver Figura 5.1 (b)). Si bien, los tres métodos generan todos los grafos, se puede ver que el Algoritmo 1 tiene un mejor rendimiento en función del tiempo de ejecución cuando se trata grafos no dirigidos. Por otro lado, cuando se generan grafos dirigidos se puede ver que en función del tamaño del grafo, los resultados indican que para grafos de 10 mil, 100 mil y 100 millones el Algoritmo 1 también posee mejor rendimiento. Pero para grafos de 1 millón, 10 millones y 250 millones es el Algoritmo 2 que posee mejor rendimiento. Se demuestra empíricamente que el Algoritmo 3 es más lento que los Algoritmos 1 y 2 en la mayoría de resultados experimentales.

Si bien, los Algoritmos 1 y 2 son los que tiene un mejor rendimiento en función del tiempo de ejecución, cabe analizar la cantidad de aristas generadas por cada método en comparación con la cantidad estimada señalada en la tabla del Cuadro 5.1. Para esta comparación, se tienen las tablas de los Cuadros 5.6 y 5.7 para grafo dirigido y no dirigido respectivamente. La cantidad de aristas más baja es 4 351, que corresponden al Algoritmo 2, cuando genera un grafo no dirigido de 1 000 nodos. La cantidad de aristas más alta es 3 443 251 139, que corresponden al Algoritmo 1 cuando genera un grafo dirigido de 250 millones.



(a) Grafos no dirigidos



(b) Grafos dirigidos

Figura 5.1: Tiempo de ejecución de los métodos para generar grafos de diferentes tamaños.

En el caso de grafos no dirigidos, se aprecia en el Cuadro 5.6 que de todos los grafos generados, solamente el Algoritmo 3 supera la cantidad de aristas estimadas, mientras que los Algoritmos 1 y 2 no alcanzan la cantidad esperada. El Algoritmo

Cuadro 5.6: Número de aristas generada por los métodos para generar grafos no dirigidos de diferentes tamaños.

Tamaño	Algoritmo 1	Algoritmo 2	Algoritmo 3	Estimado
1k	2 967	4 523	5 345	4 990
10k	32 721	60 085	65 461	65 250
100k	422 785	763 560	809 338	806 009
1M	5 022 802	9 058 915	9 637 627	9 595 150
10M	54 364 490	107 368 690	111 932 512	111 302 071
100M	658 648 148	1 211 190 297	1 273 325 547	1 266 526 383
250M	1 643 894 149	3 177 728 620	3 337 637 291	3 319 031 079

Cuadro 5.7: Número de aristas generada por los métodos para generar grafos dirigidos de diferentes tamaños.

Tamaño	Algoritmo 1	Algoritmo 2	Algoritmo 3	Estimado
1k	6 157	4 351	5 573	4 990
10k	73 274	74 648	70 056	65 250
100k	805 431	977 059	850 361	806 009
1M	10 969 895	9 440 164	10 019 287	9 595 150
10M	113 826 751	108 235 290	115 464 055	111 302 071
100M	1 212 399 045	1 135 697 577	1 306 527 631	1 266 526 383
250M	3 443 251 139	2 769 839 061	3 418 869 542	3 319 031 079

1 se queda muy atrás con respecto al número de arista que debería alcanzar. El Algoritmo 2 se queda cerca de la cantidad esperada, pero no la alcanza. Esto explica el porqué los Algoritmos 1 y 2 en la mayoría de casos tienen un mejor rendimiento en términos de tiempo con respecto al Algoritmo 3 como se ve en la Figura 5.1 (a), puesto que un menor número de aristas equivale a un menor tiempo de generación del archivo con la lista de aristas.

En el caso de grafos dirigidos, se aprecia en el Cuadro 5.6 que el Algoritmo 1 supera la cantidad de aristas estimada en la mayoría de casos, pero en grafos de 100 mil y 100 millones de nodos se queda por detrás de lo esperado. Por otro lado, el Algoritmo 2 no alcanza la cantidad de aristas estimadas en la mayoría de casos, excepto cuando se generan grafos de 10 mil y 100 mil nodos donde supera la cantidad estimada. En el caso del Algoritmo 3, este supera la cantidad de aristas estimada en todos los grafos generados. Esto explica que los tiempos de los algoritmos resulten similares para los diferentes tamaños de los grafos generados como se ve en la Figura 5.1 (b). Si bien, en el caso de los grafos dirigidos, los Algoritmos 1 y 3 superan la cantidad de aristas estimadas en la mayoría de experimentos, es el Algoritmo 3 quien la supera en todos estos. Además genera una cantidad de aristas que supera en menor medida la cantidad esperada en comparación con el Algoritmo 1, que genera una mayor cantidad de aristas con respecto al Algoritmo 3, excepto en el caso de un grafo de 10 millones de nodos. Superar por mucho la cantidad de aristas estimada, implica que los métodos tienen que usar más tiempo en generar el arreglo de aristas, por lo que no se considera positivo el superar la cantidad estimada por una cantidad muy alta. En este aspecto, es el Algoritmo 3 el que posee un mejor rendimiento con respecto al resto de los métodos. Una representación visual de esta comparación se encuentra en el Anexo B.2.

Realismo

El objetivo de este experimento, es verificar que los algoritmos generen grafos que sigan la distribución de la ley de potencia. Para hacer esto, se presentan una verificación estadística y visual. La verificación estadística se refiere al uso de la estadística de Kolmogorov-Smirnov (KS-stat) [24]. En este caso, la métrica KS-stat se usa para probar qué tan bien una distribución de muestra se ajusta a la distribución de ley de potencia.

En este experimento, se consideran grafos dirigidos y no dirigidos con los siguientes tamaños: n_1=1k, n_2=10k, n_3=100k, n_4=1M y n_5=10M. Cabe destacar que una puntuación pequeña para $KS - stat$ denota un mejor ajuste entre la distribución observada y la distribución de la ley de potencia.

Cuadro 5.8: Puntuaciones KS-stat para grafos no dirigidos.

Tamaño	**Algoritmo 1**	**Algoritmo 2**	**Algoritmo 3**
1k	0.072 144 20	0.124 262 90	0.111 965 30
10k	0.035 416 06	0.055 945 63	0.065 162 19
100k	0.018 208 59	0.041 933 74	0.033 694 74
1M	0.009 236 51	0.026 818 82	0.021 474 73
10M	0.004 737 66	0.018 445 77	0.017 333 50

En el caso de los grafos no dirigidos (ver Cuadro 5.8), todos los puntajes tienen un valor inferior a 0.13, lo que se considera muy bueno. Por lo tanto, todos los grafos no dirigidos siguen una distribución de la ley de potencia.

Cuadro 5.9: Puntuaciones KS-stat para grafos dirigidos all-degree.

Tamaño	Algoritmo 1	Algoritmo 2	Algoritmo 3
1k	0.112 676 10	0.101 105 60	0.096 464 30
10k	0.089 067 66	0.064 630 18	0.100 051 60
100k	0.065 466 66	0.057 971 01	0.048 765 48
1M	0.074 415 77	0.054 859 62	0.043 670 06
10M	0.035 558 30	0.042 539 42	0.028 063 07

Cuadro 5.10: Puntuaciones KS-stat para grafos dirigidos in-degree.

Tamaño	Algoritmo 1	Algoritmo 2	Algoritmo 3
1k	0.111 111 10	0.103 448 30	0.119 500 20
10k	0.070 630 88	0.100 902 10	0.104 896 10
100k	0.054 852 75	0.046 420 93	0.051 020 63
1M	0.050 054 22	0.032 765 4	0.0371 348 5
10M	0.019 433 26	0.025 653 58	0.042 218 19

Cuadro 5.11: Puntuaciones KS-stat para grafos dirigidos out-degree.

Tamaño	Algoritmo 1	Algoritmo 2	Algoritmo 3
1k	0.072 084 20	0.053 775 74	0.109 090 90
10k	0.035 400 29	0.075 469 22	0.084 323 95
100k	0.018 207 91	0.044 929 03	0.031 210 13
1M	0.009 236 53	0.024 315 89	0.023 979 08
10M	0.004 737 66	0.017 519 07	0.017 490 36

En el caso de grafos no dirigidos, se ha calculado el puntajes de KS-stat para el all degree, in-degree y out-degree (ver Cuadros 5.9, 5.10 y 5.11). Para el all degree, in degree y out degree todos los puntajes son inferiores a 0.12, lo que se considera un valor aceptable. Por lo tanto, todos los grafos dirigidos siguen una distribución de la ley de potencia.

Se debe tener en cuenta que los puntajes disminuyen con el aumento del tamaño del grafo, ya sea dirigido o no dirigido. Por lo tanto, se conjetura que los grafos mayores a 10 millones seguirán una distribución de la ley de potencia.

La verificación visual se realiza con los diagramas de dispersión que muestran la distribución de la aridad de los grafos. Estos gráficos se encuentran en el Anexo B.1. Se han generado gráficos para 1 000 y 1 000 000 de nodos para cada algoritmo. Con esto se demuestra que los grafos generados siguen una distribución de ley de potencia.

5.2. Fase Experimental 2

Las Fase de Evaluación Experimental 2 se basa en experimentos para verificar que el Algoritmo 3 es competitivo con el estado del arte en términos de tiempo ejecución y realismo. ¿Por qué el Algoritmo 3? Porque este algoritmo es el que presenta los mejores resultados en cuanto a la distribución de la aridad al alcanzar el número de aristas estimada y no genera nodos con grados muy altos gracias a la fórmula de la aridad máxima. Por este motivo, se compara con R^3MAT [4] (la versión basada en arreglos), SNAP [17] y PaRMAT [23] para generar grafos de 1 000 a 100 millones de nodos tanto para grafos dirigido como no dirigidos.

El experimento se realiza en la misma máquina virtual creada en Google Cloud Plataform que se utiliza en la Fase Experimental 1. Para cada combinación de parámetros experimentales, se realizan 3 ejecuciones para registrar el tiempo promedio en milisegundos. Según la documentación de SNAP, sólo puede generar grafos dirigidos. Como SNAP y PaRMAT aceptan valores de probabilidades como R-MAT para los cuadrantes, se les asignan las probabilidades alpha=0.75, beta=0.05, gamma=0.18 y delta=0.02 para grafos no dirigidos y alpha=0.75, beta=0.05, gamma=0.19 y delta=0.01 para grafos dirigidos. A R^3MAT se le configuraron estas probabilidades, ya que en [4] establecen que es la mejor combinación de estos parámetros.

5.2.1. Resultados Experimentales Fase 2

En la tabla del Cuadro 5.12 se muestran los tiempos de ejecución en milisegundos. Se usa negrita para indicar el método más rápido. Como se puede ver, el Algoritmo 3 es el más rápido en la generación de todos los grafos excepto en el caso del grafo no dirigido de 10 000 nodos, donde gana PaRMAT. Pero cabe mencionar que todos los tiempo marcados con un asterisco (*) en la columna de PaRMAT, son grafos generados con aristas repetidas o que pueden tener como nodo fuente y objetivo el mismo, lo que sucede mayoritariamente en grafos no dirigidos, puesto que el método no es capaz de generar este tipo de grafos con un tamaño superior a 10 000 nodos sin repetir aristas. Por otro lado, no se puede completar la serie con SNAP, ya que en un grafo dirigido de 10 millones estuvo ejecutándose por más de 5 horas, por lo que se decide detener su ejecución y no continuar con su evaluación. En el caso de R^3MAT, este método es capaz de completar la serie sin problema alguno y a pesar de que en algunos casos pereciera ganarle PaRMAT, este último logra tiempos más bajos que R^{3}MAT sólo en los grafos que se genera con aristas duplicadas, mientras que R^{3}MAT garantiza que no genera aristas duplicadas.

Cuadro 5.12: Comparación del Algoritmo 3 con PaRMAT, SNAP y R^{3}MAT.

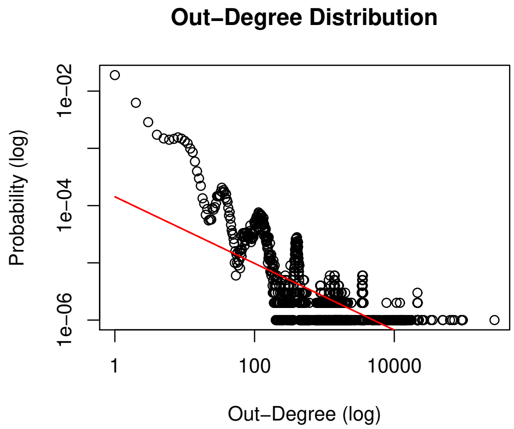
GraphID	**PaRMAT**	**SNAP**	**R^{3}MAT$_a$**	**Algoritmo 3**
UDG1K	57	-	16	**11**
DG1K	39	27	15	**9**
UDG10k	**23***	-	97	48
DG10k	1 963	329	89	**42**
UDG100k	209*	-	1 117	**187**
DG100k	15 339	9 233	867	**170**
UDG1M	2 848*	-	10 825	**1 287**
DG1M	159 162	471 268	10 418	**1 034**
UDG10M	38 525*	-	146 646	**14 202**
DG10M	1 457 991	>18 000 000	141 815	**10 390**
UDG100M	629 009*	-	2 095 889	**220 121**
DG100M	455 657*	-	2 020 027	**241 509**

Como se aprecia, el Algoritmo 3 es el más rápido de los métodos, mientras que R^{3}MAT es la segunda mejor alternativa si no se desea la generación de aristas repetidas en grafos no dirigidos de más de 10 000 nodos, puesto que en ese caso PaRMAT se presenta como una buena opción en términos de tiempo.

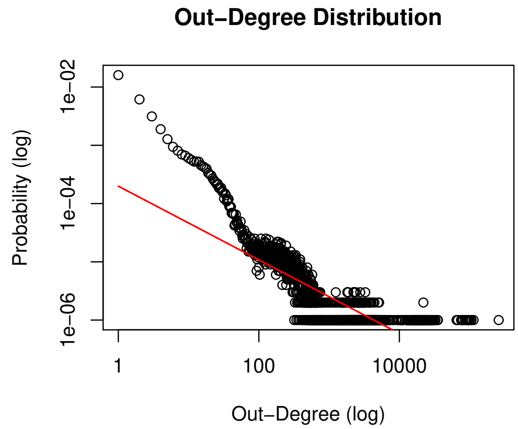
Finalmente, la Figura 5.2 muestra la distribución del Out-degree de grafos dirigidos de un millón de nodos generados con los cuatro métodos. SNAP y PaRMAT muestran distribuciones erráticas de la ley de potencia, mientras que R^{3}MAT y el Algoritmo 3 presentan las mejores distribuciones de grados. Se calcula la estadística de Kolmogorov–Smirnov y se obtiene los siguientes puntajes: ks=0.070 282 15 para PaRMAT, ks=0.069 568 54 para SNAP, ks=0.023 864 37 para R^{3}MAT y ks=0.023 979 08 para el Algoritmo 3. En este caso, R^{3}MAT es el método que presenta un mejor ajuste para de la distribución de la ley de potencia, quedando por detrás el Algoritmo 3 con un resultado bastante cercano.

5.2.2. Comparación de R^{3}MAT y Algoritmo 3

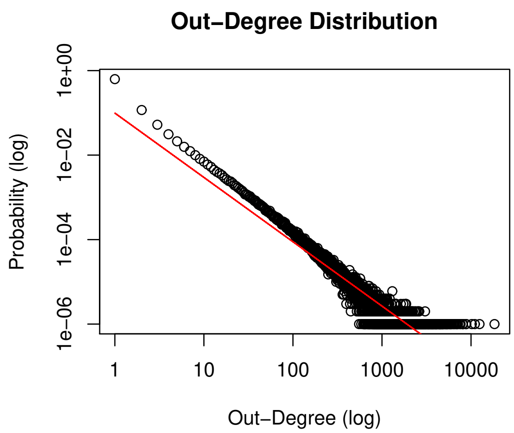
Para finalizar la Evaluación Experimental, se lleva a cabo la comparación en términos de tiempo entre R^{3}MAT y el Algoritmo 3 en la producción del arreglo de aridades. Para esto se generan grafos de 1 000 a 100 millones de nodos tanto para



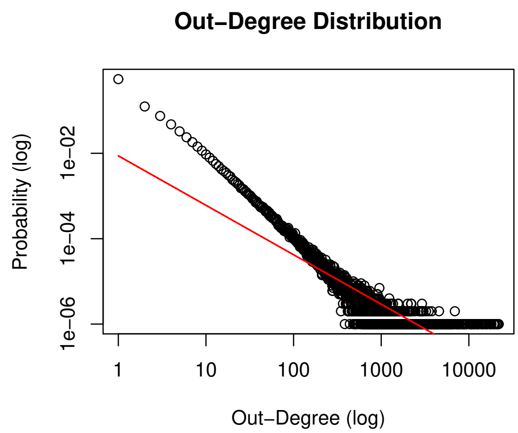
(a) PaRMAT



(b) SNAP



(c) R^3MAT_a



(d) Algoritmo 3

Figura 5.2: Distribución de Out Degree de grafos de un millon de nodos con PaRMAT (a), SNAP (b), R^3MAT_a (c) y Algoritmo 3 (d).

grafos dirigido como no dirigidos. Se registra el tiempo con la función integrada en Java System.currentTimeMillis().

En la tabla del Cuadro 5.13, se han registrado el tiempo en milisegundos que tardan R^3MAT y el Algoritmo 3 en la generar el arreglo de aridades. Como se puede ver, el Algoritmo 3 disminuye considerablemente el tiempo de generación de la estructura con el aumento del tamaño de nodos, llegando ha ser hasta 54 veces más rápido en la generación del arreglo de aridades en un grafo no dirigido de 100 millones de nodos.

Esto se debe a que el Algoritmo 3 hace uso de la función de probabilidad de la

Cuadro 5.13: Registro del tiempo de generación del arreglo de aridades en R^{3}MAT y el Algoritmo 3.

GraphID	**Número de Nodos**	**R^{3}MAT Tiempo Arreglo (ms)**	**Algoritmo 3 Tiempo Arreglo (ms)**
UDG1K	1 000	9	2
DG1K	1 000	9	2
UDG10k	10 000	64	8
DG10k	10 000	59	6
UDG100k	100 000	759	44
DG100k	100 000	705	30
UDG1M	1 000 000	9 962	315
DG1M	1 000 000	9 306	188
UDG10M	10 000 000	135 924	3 058
DG10M	10 000 000	128 580	1 765
UDG100M	100 000 000	1 777 123	32 090
DG100M	100 000 000	1 729 162	18 030

ley de potencia para generar un arreglo de probabilidades acumuladas y se generan números aleatorios que se buscan en este arreglo para generar la frecuencia de nodos asociada a cada aridad. Una vez calculadas todas las frecuencias, estas se pasan al arreglo de aridades. Todo es un proceso secuencial, mientras que R^{3}MAT simula la partición de los cuadrante como R-MAT para generar el arreglo de aridades, por lo que realiza un proceso recursivo. De esta forma se demuestra empíricamente que el Algoritmo 3 es más rápido que R^{3}MAT en términos de tiempo para la generación del arreglo de aridades.

6. Conclusiones

El método de generación de grafos R-MAT permite producir grafos sintéticos con propiedades semejantes a las observadas en los grafos del mundo real, en particular, que la distribución de aridades sigue la ley de potencia. Dada una cantidad de nodos N y un valor m de la cantidad de aristas esperadas, el método básico de R-MAT mantiene en la memoria principal la matriz de adyacencia completa que representa el conjunto de aristas E. A pesar de su simplicidad, su requisito de memoria $O(n^2)$ hace que la versión básica de R-MAT no sea adecuada para generar grafos de escalas grandes. Debido a esto, R^3MAT se desarrolló para resolver los problemas de escalabilidad de R-MAT utilizando un arreglo de aridades a cambio de la matriz, pero asegurando de mantener la propiedad de la ley de potencia en la distribución de aridades.

El problema de R^3MAT es que simula el proceso de subdividir recursivamente la matriz para generar arista, por lo que le cuesta un tiempo considerable el generar grafos grandes, pero la idea de utilizar un arreglo de aridades como una estructura que resume la matriz de adyacencia no solamente es algo positivo por reducir el requisito de memoria $O(n^2)$ a $O(n)$, sino también porque revela y motiva nuevas formas para generar grafos que sigan o no la ley de potencia.

Es esencialmente por lo anterior es que en este trabajo se aborda el problema de disminuir el tiempo de producción del arreglo de aridades, desde la perspectiva de utilizar la función de probabilidad de la ley de potencia para inducir su creación.

A continuación se resumen los principales logros de esta memoria y se analiza su impacto.

Producción del arreglo de aridades

Se propusieron tres algoritmos de producción del arreglo de aridades, donde en algunos se hace uso de un arreglo de probabilidades acumuladas y un arreglo de frecuencias.

El Algoritmo 1, que se considera como el método básico, es el que hace uso directamente de la función de probabilidad de la ley de potencia para determinar la fracción de nodos que le corresponde a cada aridad dado el tamaño del grafo. En la Evaluación Experimental ha resultado ser el que tiene la mejor evaluación de la estadística de Kolmogorov-Smirnov en grafos no dirigidos y en el out-degree de grafos dirigidos, lo que indica que posee una buena distribución de la ley de potencia. Uno de los problema de este algoritmo es que en grafos no dirigidos no alcanza la cantidad de aristas esperadas quedando muy por debajo de dicho valor. También está el hecho de que genera nodos con aridades muy altas como se demuestra al observar la distribución de aridades en un gráfico de dispersión. Pero su principal problema, es que la suma de todas las probabilidades que se utilizan al momento de calcular la cantidad de nodos por aridad es superior a 1, lo que no es aceptable.

El Algoritmo 2, surge principalmente con el objetivo de resolver el problema de la suma de probabilidades del Algoritmo 1. Como resultado, esto conlleva a la generación de un arreglo de probabilidades acumuladas donde dichas probabilidades son ajustadas previamente. Esto resuelve el problema del Algoritmo 1. El problema de este método, es que en la mayoría de casos no alcanza la cantidad de aristas estimadas y a su vez mantiene el problema de generar nodos con aridades muy altas. La falta de aristas para alcanzar la cantidad dada por la fórmula sumado a la alta preferencia por la aridad uno, fue lo que motivo el realizar un análisis de la generación de los números aleatorios de Math.random(). Esto conlleva a la búsqueda de la fórmula que genera números aleatorios con distribución de la ley de potencia a partir de una distribución uniforme. Por otro lado, el hecho de generar nodos con aridades muy altas en comparación a los resultados que genera R^{3}MAT conduce a la realización de un experimento en búsqueda de modelar la función que determina la aridad máxima de un grafo en función del número de nodos N y el valor de τ. Básicamente, con esto quiero decir que este algoritmo no destaca particularmente en algún apartado. En los resultados de la Evaluación Experimental tuvo resultados medianamente buenos al ser rápido en ciertos casos y teniendo algunos valores de la estadística de Kolmogorov-Smirnov mejores que los Algoritmos 1 y 3 en ciertos casos. Pero se considera esencial en el desarrollo de esta memoria, pues fue la estructura que más se estudio y la que motivo la búsqueda de diferentes mecanismos para resolver los principales problemas que surgieron durante el desarrollo e investigación.

El Algoritmo 3, es el mecanismo que implementa la solución a los nodos con aridades muy elevadas por medio de la función que modela la aridad máxima, la cual usa como límite en la producción de aridades permitidas en el arreglo de aridades. Ahora bien, el uso de esta función también permite reducir el tamaño del arreglo de probabilidades acumuladas y del arreglo de frecuencias, puesto que en estas estructuras sólo se utiliza el tamaño indicado por la aridad máxima. De esta forma se reduce la cantidad de memoria que utiliza el Algoritmo 2, ya que todas las estructuras en este último son del tamaño del grafo (número de nodos). Este uso de la aridad máxima es muy útil, pues se demuestra empíricamente que este valor no se acerca a la aridad media ($N/2$) dado el tamaño de un grafo (N). Esto implica que la estructuras para el arreglo de probabilidades acumuladas y el arreglo de frecuencia son altamente escalable con el aumento del número de nodos, sobre todo la primera estructura que es un arreglo de *double*. Por otro lado, el Algoritmo 3 es el método que implementa la generación de número aleatorios con distribución de la ley de potencia, lo que modifica la alta preferencia de la aridad uno y se generan más nodos con aridades más elevadas. De esta forma, el Algoritmo 3 alcanza con mayor facilidad la cantidad de aristas estimada. Por último, en la Evaluación Experimental, se demuestra que este método es muy rápido en comparación a otros algoritmos, que son una alternativa para generar grafos con distribución de la ley de potencia. Además, demuestra tener buenos valores para la estadística de Kolmogorov-Smirnov, quedando sólo por detrás de R^{3}MAT en la comparación con los métodos previamente existentes. En la comparación final de los tiempos de generación del arreglo de aridades entre R^{3}MAT y el Algoritmo 3, se comprueba la reducción del tiempo de producción, donde este último demuestra ser hasta 54 veces más rápido en generar esta estructura en una grafo de gran escala como los son 100 millones de nodos.

Luego de este análisis, se considera al Algoritmo 3 como el mejor método obtenido durante el desarrollo de esta memoria para producir un grafo con la distribución de la ley de potencia. Ha reducido el tiempo de generación, produce la cantidad de aristas esperadas y los grafos resultantes mantienen la propiedad la ley de potencia en la distribución de aridades.

Trabajo Futuro

Este trabajo de memoria se puede expandir tanto en el tema de la distribución de aridades como en la generación de las aristas.

Como se ha demostrado que la producción del arreglo de aridades es muy rápida. Esto permite diseñar e implementar un método que mejore la distribución de las aridades en el arreglo que use más tiempo de ejecución que se ha ganado con respecto a la competencia, de tal forma que todos los valores generados en el arreglo de frecuencias puedan generarse y no se pierdan nodos con aridad uno que deberían generarse.

Mejorar la fórmula que determina la aridad máxima podría ser posible por el estudio de las lineas de tendencia que se genera a partir de la distribución de aridades en los gráficos de dispersión.

Realizar el estudio sobre los grafos generados en esta memoria del punto o aridad en la que inicia y termina el comportamiento de la ley de potencia. Esto hace referencia a que en los gráficos de dispersión se observan secciones lineales que corresponden a la distribución de la ley de potencia, pero estas secciones no esta presentes en toda la distribución. Por lo que se plantea que la ley de potencia se encuentra presente en un rango de aridades más allá del promedio de nodos vecinos, pero que no llega a las aridades más altas que se generan.

Se podría implementar un método distribuido para la generación de las aristas una vez que se ha generado el arreglo de aridades, lo que podría mejorar el tiempo global del algoritmo al reducir el tiempo de producción de las aristas.

Realizar pruebas con grafos de mayor escala y con diferentes valores de τ para determinar el comportamiento del algoritmo, el funcionamiento de la fórmula de la aridad máxima y el número de aristas generadas. Cabe mencionar en este apartado que se han generado grafos de 1 000 millones de nodos con $\tau = 2.29$ y grafos de 1 500 millones de nodos con $\tau = 2.72$.

Bibliografía

[1] Arijit Khan and Sameh Elnikety. Systems for big-graphs. *Proceedings of the VLDB Endowment*, 7(13):1709–1710, 2014.

[2] Shuai Zhao, Le Yu, and Bo Cheng. Probabilistic community using link and content for social networks. *IEEE Access*, 5:27189–27202, 2017.

[3] Zhibo Zhu, Qinke Peng, Zhi Li, Xinyu Guan, and Owais Muhammad. Fast pagerank computation based on network decomposition and dag structure. *IEEE Access*, 6:41760–41770, 2018.

[4] R. Angles, R. Paredes, and R. García. R3MAT: A rapid and robust graph generator. *IEEE Access*, 8:130048–130065, 2020.

[5] Francisco Javier Martín Pliego and Luis Ruiz-Maya Pérez. *Fundamentos de probabilidad*. Editorial Paraninfo, 2006.

[6] Luis Rodríguez Ojeda. Construcción de kernel y funciones de densidad de probabilidad. *ESPOL, Matemáticas.*, page 18, 2012.

[7] Deepayan Chakrabarti and Christos Faloutsos. Graph mining: Laws, generators, and algorithms. *ACM computing surveys (CSUR)*, 38(1):2, 2006.

[8] MathWorld Team. Random number. https://mathworld.wolfram.com/RandomNumber.html, 2020. Accedido 28-03-2020.

[9] R. Paredes. *Graphs for Metric Space Searching*. PhD thesis, University of Chile, Chile, July 2008. Advisor: G. Navarro. Dept. of Computer Science Tech Report TR/DCC-2008-10. Available at http://www.dcc.uchile.cl/˜raparede/publ/08PhDthesis.pdf.

[10] Jorge P. Apuntes de Matem aticas Discretas Fundamentos de Matem aticas Discretas. page 117, 2009.

[11] Francisco Soulignac. Notas de la clase 3–introducción a grafos.

[12] Oded Green, Robert Mccoll, and David A Bader. A Fast Algorithm for Streaming Betweenness Centrality. *2012 International Conference on Privacy, Security, Risk and Trust and 2012 International Confernece on Social Computing*, page 20, 2012.

[13] Deepayan Chakrabarti, Yiping Zhan, and Christos Faloutsos. R-mat: A recursive model for graph mining. In *Proceedings of the 2004 SIAM International Conference on Data Mining*, page 446. SIAM, 2004.

[14] Chris Groër, Blair D Sullivan, and Steve Poole. A mathematical analysis of the r-mat random graph generator. *Networks*, 58(3):159–170, 2011.

[15] Hubert W Lilliefors. On the kolmogorov-smirnov test for normality with mean and variance unknown. *Journal of the American statistical Association*, 62(318):399–402, 1967.

[16] Arnau Prat-Pérez, Joan Guisado-Gámez, Xavier Fernández Salas, Petr Koupy, Siegfried Depner, and Davide Basilio Bartolini. Towards a property graph generator for benchmarking. page 6, 2017.

[17] Jure Leskovec and Rok Sosič. Snap: A general-purpose network analysis and graph-mining library. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 8(1):1–20, 2016.

[18] Lorenz Hübschle-Schneider and Peter Sanders. Linear work generation of r-mat graphs. *arXiv preprint arXiv:1905.03525*, 2019.

[19] ORACLE. ¿qué es java?, 2020. https://www.java.com/es/about/.

[20] The R Foundation. What is R?, 2020. https://www.r-project.org/about.html.

[21] Microsoft. Ayuda y aprendizaje de excel, 2020. https://support.office.com/es-es/excel.

[22] Monje Álvarez Carlos Arturo. Metodología de la investigación cuantitativa y cualitativa. *Universidad Surcolombiana, Facultad de Ciencias Sociales y Humanas, Programa de Comunicación Social y Periodismo, Neiva*, page 217, 2011.

[23] Farzad Khorasani, Rajiv Gupta, and Laxmi N Bhuyan. Scalable simd-efficient graph processing on gpus. pages 39–50, 2015.

[24] Aaron Clauset, Cosma Rohilla Shalizi, and Mark EJ Newman. Power-law distributions in empirical data. *SIAM review*, 51(4):661–703, 2009.

ANEXOS

A. Experimentos Preliminares

En este anexo se encuentran los Cuadros con la información completa de los experimentos preliminares de los algoritmos propuestos en el Capítulo 4.

A.1. Experimentos Preliminares Algoritmo 1

Aquí se encuentra la tabla de probabilidades asociada al problema del Algoritmo 1.

Cuadro A.1: Probabilidades para las aridades de un grafo de 1 000 nodos.

Aridad	**Probabilidad**	**Aridad**	**Probabilidad**	**Aridad**	**Probabilidad**
1	0.666666667	334	0.000005976	667	0.000001499
2	0.166666667	335	0.000005940	668	0.000001494
3	0.074074074	336	0.000005905	669	0.000001490
4	0.041666667	337	0.000005870	670	0.000001485
5	0.026666667	338	0.000005835	671	0.000001481
6	0.018518519	339	0.000005801	672	0.000001476
7	0.013605442	340	0.000005767	673	0.000001472
8	0.010416667	341	0.000005733	674	0.000001468
9	0.008230453	342	0.000005700	675	0.000001463
10	0.006666667	343	0.000005667	676	0.000001459
11	0.005509642	344	0.000005634	677	0.000001455
12	0.004629630	345	0.000005601	678	0.000001450
13	0.003944773	346	0.000005569	679	0.000001446

Aridad	Probabilidad	Aridad	Probabilidad	Aridad	Probabilidad
14	0.003401361	347	0.000005537	680	0.000001442
15	0.002962963	348	0.000005505	681	0.000001438
16	0.002604167	349	0.000005473	682	0.000001433
17	0.002306805	350	0.000005442	683	0.000001429
18	0.002057613	351	0.000005411	684	0.000001425
19	0.001846722	352	0.000005381	685	0.000001421
20	0.001666667	353	0.000005350	686	0.000001417
21	0.001511716	354	0.000005320	687	0.000001413
22	0.001377410	355	0.000005290	688	0.000001408
23	0.001260239	356	0.000005260	689	0.000001404
24	0.001157407	357	0.000005231	690	0.000001400
25	0.001066667	358	0.000005202	691	0.000001396
26	0.000986193	359	0.000005173	692	0.000001392
27	0.000914495	360	0.000005144	693	0.000001388
28	0.000850340	361	0.000005116	694	0.000001384
29	0.000792707	362	0.000005087	695	0.000001380
30	0.000740741	363	0.000005059	696	0.000001376
31	0.000693722	364	0.000005032	697	0.000001372
32	0.000651042	365	0.000005004	698	0.000001368
33	0.000612182	366	0.000004977	699	0.000001364
34	0.000576701	367	0.000004950	700	0.000001361
35	0.000544218	368	0.000004923	701	0.000001357
36	0.000514403	369	0.000004896	702	0.000001353
37	0.000486973	370	0.000004870	703	0.000001349
38	0.000461681	371	0.000004844	704	0.000001345
39	0.000438308	372	0.000004818	705	0.000001341
40	0.000416667	373	0.000004792	706	0.000001338
41	0.000396589	374	0.000004766	707	0.000001334
42	0.000377929	375	0.000004741	708	0.000001330
43	0.000360555	376	0.000004716	709	0.000001326
44	0.000344353	377	0.000004691	710	0.000001322
45	0.000329218	378	0.000004666	711	0.000001319

Aridad	Probabilidad	Aridad	Probabilidad	Aridad	Probabilidad
46	0.000315060	379	0.000004641	712	0.000001315
47	0.000301796	380	0.000004617	713	0.000001311
48	0.000289352	381	0.000004593	714	0.000001308
49	0.000277662	382	0.000004569	715	0.000001304
50	0.000266667	383	0.000004545	716	0.000001300
51	0.000256312	384	0.000004521	717	0.000001297
52	0.000246548	385	0.000004498	718	0.000001293
53	0.000237332	386	0.000004474	719	0.000001290
54	0.000228624	387	0.000004451	720	0.000001286
55	0.000220386	388	0.000004428	721	0.000001282
56	0.000212585	389	0.000004406	722	0.000001279
57	0.000205191	390	0.000004383	723	0.000001275
58	0.000198177	391	0.000004361	724	0.000001272
59	0.000191516	392	0.000004338	725	0.000001268
60	0.000185185	393	0.000004316	726	0.000001265
61	0.000179163	394	0.000004295	727	0.000001261
62	0.000173430	395	0.000004273	728	0.000001258
63	0.000167968	396	0.000004251	729	0.000001254
64	0.000162760	397	0.000004230	730	0.000001251
65	0.000157791	398	0.000004209	731	0.000001248
66	0.000153046	399	0.000004188	732	0.000001244
67	0.000148511	400	0.000004167	733	0.000001241
68	0.000144175	401	0.000004146	734	0.000001237
69	0.000140027	402	0.000004125	735	0.000001234
70	0.000136054	403	0.000004105	736	0.000001231
71	0.000132249	404	0.000004085	737	0.000001227
72	0.000128601	405	0.000004064	738	0.000001224
73	0.000125102	406	0.000004044	739	0.000001221
74	0.000121743	407	0.000004025	740	0.000001217
75	0.000118519	408	0.000004005	741	0.000001214
76	0.000115420	409	0.000003985	742	0.000001211
77	0.000112442	410	0.000003966	743	0.000001208

Aridad	Probabilidad	Aridad	Probabilidad	Aridad	Probabilidad
78	0.000109577	411	0.000003947	744	0.000001204
79	0.000106820	412	0.000003927	745	0.000001201
80	0.000104167	413	0.000003908	746	0.000001198
81	0.000101611	414	0.000003890	747	0.000001195
82	0.000099147	415	0.000003871	748	0.000001192
83	0.000096773	416	0.000003852	749	0.000001188
84	0.000094482	417	0.000003834	750	0.000001185
85	0.000092272	418	0.000003816	751	0.000001182
86	0.000090139	419	0.000003797	752	0.000001179
87	0.000088079	420	0.000003779	753	0.000001176
88	0.000086088	421	0.000003761	754	0.000001173
89	0.000084164	422	0.000003744	755	0.000001170
90	0.000082305	423	0.000003726	756	0.000001166
91	0.000080506	424	0.000003708	757	0.000001163
92	0.000078765	425	0.000003691	758	0.000001160
93	0.000077080	426	0.000003674	759	0.000001157
94	0.000075449	427	0.000003656	760	0.000001154
95	0.000073869	428	0.000003639	761	0.000001151
96	0.000072338	429	0.000003622	762	0.000001148
97	0.000070854	430	0.000003606	763	0.000001145
98	0.000069416	431	0.000003589	764	0.000001142
99	0.000068020	432	0.000003572	765	0.000001139
100	0.000066667	433	0.000003556	766	0.000001136
101	0.000065353	434	0.000003539	767	0.000001133
102	0.000064078	435	0.000003523	768	0.000001130
103	0.000062840	436	0.000003507	769	0.000001127
104	0.000061637	437	0.000003491	770	0.000001124
105	0.000060469	438	0.000003475	771	0.000001122
106	0.000059333	439	0.000003459	772	0.000001119
107	0.000058229	440	0.000003444	773	0.000001116
108	0.000057156	441	0.000003428	774	0.000001113
109	0.000056112	442	0.000003412	775	0.000001110

Aridad	Probabilidad	Aridad	Probabilidad	Aridad	Probabilidad
110	0.000055096	443	0.000003397	776	0.000001107
111	0.000054108	444	0.000003382	777	0.000001104
112	0.000053146	445	0.000003367	778	0.000001101
113	0.000052210	446	0.000003351	779	0.000001099
114	0.000051298	447	0.000003337	780	0.000001096
115	0.000050410	448	0.000003322	781	0.000001093
116	0.000049544	449	0.000003307	782	0.000001090
117	0.000048701	450	0.000003292	783	0.000001087
118	0.000047879	451	0.000003278	784	0.000001085
119	0.000047078	452	0.000003263	785	0.000001082
120	0.000046296	453	0.000003249	786	0.000001079
121	0.000045534	454	0.000003234	787	0.000001076
122	0.000044791	455	0.000003220	788	0.000001074
123	0.000044065	456	0.000003206	789	0.000001071
124	0.000043358	457	0.000003192	790	0.000001068
125	0.000042667	458	0.000003178	791	0.000001066
126	0.000041992	459	0.000003164	792	0.000001063
127	0.000041333	460	0.000003151	793	0.000001060
128	0.000040690	461	0.000003137	794	0.000001057
129	0.000040062	462	0.000003123	795	0.000001055
130	0.000039448	463	0.000003110	796	0.000001052
131	0.000038848	464	0.000003097	797	0.000001050
132	0.000038261	465	0.000003083	798	0.000001047
133	0.000037688	466	0.000003070	799	0.000001044
134	0.000037128	467	0.000003057	800	0.000001042
135	0.000036580	468	0.000003044	801	0.000001039
136	0.000036044	469	0.000003031	802	0.000001036
137	0.000035520	470	0.000003018	803	0.000001034
138	0.000035007	471	0.000003005	804	0.000001031
139	0.000034505	472	0.000002992	805	0.000001029
140	0.000034014	473	0.000002980	806	0.000001026
141	0.000033533	474	0.000002967	807	0.000001024

Aridad	Probabilidad	Aridad	Probabilidad	Aridad	Probabilidad
142	0.000033062	475	0.000002955	808	0.000001021
143	0.000032601	476	0.000002942	809	0.000001019
144	0.000032150	477	0.000002930	810	0.000001016
145	0.000031708	478	0.000002918	811	0.000001014
146	0.000031275	479	0.000002906	812	0.000001011
147	0.000030851	480	0.000002894	813	0.000001009
148	0.000030436	481	0.000002881	814	0.000001006
149	0.000030029	482	0.000002870	815	0.000001004
150	0.000029630	483	0.000002858	816	0.000001001
151	0.000029238	484	0.000002846	817	0.000000999
152	0.000028855	485	0.000002834	818	0.000000996
153	0.000028479	486	0.000002823	819	0.000000994
154	0.000028110	487	0.000002811	820	0.000000991
155	0.000027749	488	0.000002799	821	0.000000989
156	0.000027394	489	0.000002788	822	0.000000987
157	0.000027046	490	0.000002777	823	0.000000984
158	0.000026705	491	0.000002765	824	0.000000982
159	0.000026370	492	0.000002754	825	0.000000979
160	0.000026042	493	0.000002743	826	0.000000977
161	0.000025719	494	0.000002732	827	0.000000975
162	0.000025403	495	0.000002721	828	0.000000972
163	0.000025092	496	0.000002710	829	0.000000970
164	0.000024787	497	0.000002699	830	0.000000968
165	0.000024487	498	0.000002688	831	0.000000965
166	0.000024193	499	0.000002677	832	0.000000963
167	0.000023904	500	0.000002667	833	0.000000961
168	0.000023621	501	0.000002656	834	0.000000958
169	0.000023342	502	0.000002645	835	0.000000956
170	0.000023068	503	0.000002635	836	0.000000954
171	0.000022799	504	0.000002625	837	0.000000952
172	0.000022535	505	0.000002614	838	0.000000949
173	0.000022275	506	0.000002604	839	0.000000947

Aridad	Probabilidad	Aridad	Probabilidad	Aridad	Probabilidad
174	0.000022020	507	0.000002594	840	0.000000945
175	0.000021769	508	0.000002583	841	0.000000943
176	0.000021522	509	0.000002573	842	0.000000940
177	0.000021280	510	0.000002563	843	0.000000938
178	0.000021041	511	0.000002553	844	0.000000936
179	0.000020807	512	0.000002543	845	0.000000934
180	0.000020576	513	0.000002533	846	0.000000931
181	0.000020349	514	0.000002523	847	0.000000929
182	0.000020126	515	0.000002514	848	0.000000927
183	0.000019907	516	0.000002504	849	0.000000925
184	0.000019691	517	0.000002494	850	0.000000923
185	0.000019479	518	0.000002485	851	0.000000921
186	0.000019270	519	0.000002475	852	0.000000918
187	0.000019065	520	0.000002465	853	0.000000916
188	0.000018862	521	0.000002456	854	0.000000914
189	0.000018663	522	0.000002447	855	0.000000912
190	0.000018467	523	0.000002437	856	0.000000910
191	0.000018274	524	0.000002428	857	0.000000908
192	0.000018084	525	0.000002419	858	0.000000906
193	0.000017898	526	0.000002410	859	0.000000903
194	0.000017714	527	0.000002400	860	0.000000901
195	0.000017532	528	0.000002391	861	0.000000899
196	0.000017354	529	0.000002382	862	0.000000897
197	0.000017178	530	0.000002373	863	0.000000895
198	0.000017005	531	0.000002364	864	0.000000893
199	0.000016835	532	0.000002356	865	0.000000891
200	0.000016667	533	0.000002347	866	0.000000889
201	0.000016501	534	0.000002338	867	0.000000887
202	0.000016338	535	0.000002329	868	0.000000885
203	0.000016178	536	0.000002320	869	0.000000883
204	0.000016019	537	0.000002312	870	0.000000881
205	0.000015864	538	0.000002303	871	0.000000879

Aridad	**Probabilidad**	**Aridad**	**Probabilidad**	**Aridad**	**Probabilidad**
206	0.000015710	539	0.000002295	872	0.000000877
207	0.000015559	540	0.000002286	873	0.000000875
208	0.000015409	541	0.000002278	874	0.000000873
209	0.000015262	542	0.000002269	875	0.000000871
210	0.000015117	543	0.000002261	876	0.000000869
211	0.000014974	544	0.000002253	877	0.000000867
212	0.000014833	545	0.000002244	878	0.000000865
213	0.000014694	546	0.000002236	879	0.000000863
214	0.000014557	547	0.000002228	880	0.000000861
215	0.000014422	548	0.000002220	881	0.000000859
216	0.000014289	549	0.000002212	882	0.000000857
217	0.000014158	550	0.000002204	883	0.000000855
218	0.000014028	551	0.000002196	884	0.000000853
219	0.000013900	552	0.000002188	885	0.000000851
220	0.000013774	553	0.000002180	886	0.000000849
221	0.000013650	554	0.000002172	887	0.000000847
222	0.000013527	555	0.000002164	888	0.000000845
223	0.000013406	556	0.000002157	889	0.000000844
224	0.000013287	557	0.000002149	890	0.000000842
225	0.000013169	558	0.000002141	891	0.000000840
226	0.000013052	559	0.000002133	892	0.000000838
227	0.000012938	560	0.000002126	893	0.000000836
228	0.000012824	561	0.000002118	894	0.000000834
229	0.000012713	562	0.000002111	895	0.000000832
230	0.000012602	563	0.000002103	896	0.000000830
231	0.000012494	564	0.000002096	897	0.000000829
232	0.000012386	565	0.000002088	898	0.000000827
233	0.000012280	566	0.000002081	899	0.000000825
234	0.000012175	567	0.000002074	900	0.000000823
235	0.000012072	568	0.000002066	901	0.000000821
236	0.000011970	569	0.000002059	902	0.000000819
237	0.000011869	570	0.000002052	903	0.000000818

Aridad	Probabilidad	Aridad	Probabilidad	Aridad	Probabilidad
238	0.000011769	571	0.000002045	904	0.000000816
239	0.000011671	572	0.000002038	905	0.000000814
240	0.000011574	573	0.000002030	906	0.000000812
241	0.000011478	574	0.000002023	907	0.000000810
242	0.000011384	575	0.000002016	908	0.000000809
243	0.000011290	576	0.000002009	909	0.000000807
244	0.000011198	577	0.000002002	910	0.000000805
245	0.000011106	578	0.000001996	911	0.000000803
246	0.000011016	579	0.000001989	912	0.000000802
247	0.000010927	580	0.000001982	913	0.000000800
248	0.000010839	581	0.000001975	914	0.000000798
249	0.000010753	582	0.000001968	915	0.000000796
250	0.000010667	583	0.000001961	916	0.000000795
251	0.000010582	584	0.000001955	917	0.000000793
252	0.000010498	585	0.000001948	918	0.000000791
253	0.000010415	586	0.000001941	919	0.000000789
254	0.000010333	587	0.000001935	920	0.000000788
255	0.000010252	588	0.000001928	921	0.000000786
256	0.000010173	589	0.000001922	922	0.000000784
257	0.000010094	590	0.000001915	923	0.000000783
258	0.000010015	591	0.000001909	924	0.000000781
259	0.000009938	592	0.000001902	925	0.000000779
260	0.000009862	593	0.000001896	926	0.000000777
261	0.000009787	594	0.000001889	927	0.000000776
262	0.000009712	595	0.000001883	928	0.000000774
263	0.000009638	596	0.000001877	929	0.000000772
264	0.000009565	597	0.000001871	930	0.000000771
265	0.000009493	598	0.000001864	931	0.000000769
266	0.000009422	599	0.000001858	932	0.000000767
267	0.000009352	600	0.000001852	933	0.000000766
268	0.000009282	601	0.000001846	934	0.000000764
269	0.000009213	602	0.000001840	935	0.000000763

Aridad	Probabilidad	Aridad	Probabilidad	Aridad	Probabilidad
270	0.000009145	603	0.000001833	936	0.000000761
271	0.000009078	604	0.000001827	937	0.000000759
272	0.000009011	605	0.000001821	938	0.000000758
273	0.000008945	606	0.000001815	939	0.000000756
274	0.000008880	607	0.000001809	940	0.000000754
275	0.000008815	608	0.000001803	941	0.000000753
276	0.000008752	609	0.000001798	942	0.000000751
277	0.000008689	610	0.000001792	943	0.000000750
278	0.000008626	611	0.000001786	944	0.000000748
279	0.000008564	612	0.000001780	945	0.000000747
280	0.000008503	613	0.000001774	946	0.000000745
281	0.000008443	614	0.000001768	947	0.000000743
282	0.000008383	615	0.000001763	948	0.000000742
283	0.000008324	616	0.000001757	949	0.000000740
284	0.000008266	617	0.000001751	950	0.000000739
285	0.000008208	618	0.000001746	951	0.000000737
286	0.000008150	619	0.000001740	952	0.000000736
287	0.000008094	620	0.000001734	953	0.000000734
288	0.000008038	621	0.000001729	954	0.000000733
289	0.000007982	622	0.000001723	955	0.000000731
290	0.000007927	623	0.000001718	956	0.000000729
291	0.000007873	624	0.000001712	957	0.000000728
292	0.000007819	625	0.000001707	958	0.000000726
293	0.000007766	626	0.000001701	959	0.000000725
294	0.000007713	627	0.000001696	960	0.000000723
295	0.000007661	628	0.000001690	961	0.000000722
296	0.000007609	629	0.000001685	962	0.000000720
297	0.000007558	630	0.000001680	963	0.000000719
298	0.000007507	631	0.000001674	964	0.000000717
299	0.000007457	632	0.000001669	965	0.000000716
300	0.000007407	633	0.000001664	966	0.000000714
301	0.000007358	634	0.000001659	967	0.000000713

Aridad	Probabilidad	Aridad	Probabilidad	Aridad	Probabilidad
302	0.000007310	635	0.000001653	968	0.000000711
303	0.000007261	636	0.000001648	969	0.000000710
304	0.000007214	637	0.000001643	970	0.000000709
305	0.000007167	638	0.000001638	971	0.000000707
306	0.000007120	639	0.000001633	972	0.000000706
307	0.000007073	640	0.000001628	973	0.000000704
308	0.000007028	641	0.000001623	974	0.000000703
309	0.000006982	642	0.000001617	975	0.000000701
310	0.000006937	643	0.000001612	976	0.000000700
311	0.000006893	644	0.000001607	977	0.000000698
312	0.000006849	645	0.000001602	978	0.000000697
313	0.000006805	646	0.000001598	979	0.000000696
314	0.000006762	647	0.000001593	980	0.000000694
315	0.000006719	648	0.000001588	981	0.000000693
316	0.000006676	649	0.000001583	982	0.000000691
317	0.000006634	650	0.000001578	983	0.000000690
318	0.000006593	651	0.000001573	984	0.000000689
319	0.000006551	652	0.000001568	985	0.000000687
320	0.000006510	653	0.000001563	986	0.000000686
321	0.000006470	654	0.000001559	987	0.000000684
322	0.000006430	655	0.000001554	988	0.000000683
323	0.000006390	656	0.000001549	989	0.000000682
324	0.000006351	657	0.000001544	990	0.000000680
325	0.000006312	658	0.000001540	991	0.000000679
326	0.000006273	659	0.000001535	992	0.000000677
327	0.000006235	660	0.000001530	993	0.000000676
328	0.000006197	661	0.000001526	994	0.000000675
329	0.000006159	662	0.000001521	995	0.000000673
330	0.000006122	663	0.000001517	996	0.000000672
331	0.000006085	664	0.000001512	997	0.000000671
332	0.000006048	665	0.000001508	998	0.000000669
333	0.000006012	666	0.000001503	999	0.000000668

A.2. Experimento para la Aridad Máxima

En esta sección de Anexos se presentan las tablas con los datos asociadas al experimento que determina la aridad máxima.

A.2.1. Aridades Máximas Clasificadas por Tamaño del Grafo.

Cuadro A.2: Aridades máximas registradas para un grafo de 1 000 nodos.

τ	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**Promedio**
1.92	529	488	539	554	528	562	524	516	525	522	528.7
1.93	538	528	514	545	516	502	501	515	510	477	514.6
1.94	486	527	545	511	490	510	515	526	466	529	510.5
1.95	509	470	479	513	477	486	472	518	535	479	493.8
1.96	422	483	453	461	477	472	465	487	497	491	470.8
1.97	419	455	450	444	430	442	474	435	461	453	446.3
1.98	436	472	479	439	452	467	490	474	452	477	463.8
1.99	454	431	429	403	408	419	422	425	426	446	426.3
2.00	411	407	416	421	447	432	408	416	414	423	419.5
2.01	410	405	418	430	421	393	394	401	410	400	408.2
2.02	404	403	396	399	400	408	361	375	391	410	394.7
2.03	391	392	403	405	362	391	405	397	410	419	397.5
2.04	362	387	362	359	371	370	357	374	388	353	368.3
2.05	387	375	367	376	328	345	386	362	382	371	367.9
2.06	354	339	360	373	356	329	353	348	340	364	351.6
2.07	337	356	321	345	373	359	311	372	371	351	349.6
2.08	316	346	337	357	327	337	353	353	327	342	339.5
2.09	342	351	322	328	316	338	330	303	306	341	327.7
2.10	309	350	327	341	307	328	341	327	313	332	327.5
2.11	295	344	256	296	289	338	302	308	316	305	304.9
2.12	283	353	302	318	319	294	319	294	300	340	312.2
2.13	307	325	293	312	292	313	282	321	282	295	302.2
2.14	310	281	279	263	309	297	317	298	304	283	294.1

2.15	277	284	287	289	269	289	291	287	264	316	285.3
2.16	300	236	273	262	269	280	309	277	270	264	274
2.17	280	291	273	279	265	265	250	262	272	269	270.6
2.18	273	282	260	260	266	254	281	280	237	256	264.9
2.19	256	257	259	272	265	258	252	289	264	264	263.6
2.20	283	216	266	256	246	255	270	252	268	256	256.8
2.21	236	246	236	250	240	248	251	273	236	254	247
2.22	259	245	206	243	214	226	248	235	238	247	236.1
2.23	250	247	245	235	254	277	223	242	226	211	241
2.24	252	231	210	222	218	213	248	251	245	249	233.9
2.25	218	241	234	213	238	246	248	229	221	235	232.3
2.26	212	201	234	250	225	246	228	228	217	215	225.6
2.27	229	226	249	243	217	229	229	217	211	247	229.7
2.28	214	217	230	241	202	215	209	221	225	184	215.8
2.29	223	186	229	181	207	213	201	200	230	228	209.8
2.30	224	206	220	190	212	233	198	204	199	198	208.4
2.31	211	203	192	221	207	178	209	212	203	235	207.1
2.32	206	208	198	215	214	193	215	207	187	212	205.5
2.33	210	180	210	193	211	201	177	191	201	197	197.1
2.34	189	173	199	186	171	195	181	180	175	178	182.7
2.35	165	177	194	192	218	195	178	180	164	200	186.3
2.36	199	173	191	214	183	186	181	183	190	216	191.6
2.37	200	180	198	201	186	181	196	168	178	179	186.7
2.38	195	171	187	162	183	165	194	179	187	185	180.8
2.39	189	192	192	188	177	156	165	198	188	169	181.4
2.40	171	166	171	186	164	199	176	200	172	175	178
2.41	182	156	177	177	162	180	170	170	153	180	170.7
2.42	156	180	174	168	170	173	192	171	178	165	172.7
2.43	144	189	146	147	173	170	168	154	167	177	163.5
2.44	175	162	190	158	166	182	172	141	198	172	171.6
2.45	185	165	149	169	172	149	166	174	144	178	165.1
2.46	167	156	180	155	155	169	182	168	167	157	165.6

2.47	146	141	152	157	182	156	148	163	141	158	154.4
2.48	148	145	146	159	145	151	123	138	163	160	147.8
2.49	161	158	179	141	162	141	151	164	150	158	156.5
2.50	134	153	148	142	151	146	159	131	137	139	144
2.51	147	133	172	149	149	139	148	136	147	136	145.6
2.52	146	132	172	143	153	145	147	139	141	147	146.5
2.53	131	136	142	150	140	147	172	135	159	155	146.7
2.54	158	150	138	158	141	161	139	145	118	143	145.1
2.55	137	148	127	158	145	152	153	149	134	149	145.2
2.56	145	138	142	137	137	136	135	139	149	134	139.2
2.57	144	125	139	130	139	141	156	131	116	123	134.4
2.58	126	129	130	128	129	125	137	139	141	138	132.2
2.59	125	154	140	136	119	117	146	124	131	125	131.7
2.60	117	118	126	158	139	138	135	140	116	130	131.7
2.61	131	117	119	131	126	151	157	139	139	139	134.9
2.62	131	123	151	116	139	125	139	138	145	121	132.8
2.63	123	123	130	138	121	149	136	123	124	120	128.7
2.64	121	117	135	107	104	140	105	142	145	130	124.6
2.65	129	130	109	133	118	157	142	146	134	127	132.5
2.66	123	111	128	108	116	122	125	131	103	121	118.8
2.67	128	104	112	123	143	126	126	127	139	139	126.7
2.68	125	127	138	137	107	138	145	130	113	128	128.8
2.69	117	118	130	107	108	113	111	118	146	125	119.3
2.70	107	126	115	115	138	113	120	131	115	122	120.2
2.71	111	104	108	121	122	107	129	137	114	110	116.3
2.72	119	126	106	120	101	122	133	129	107	138	120.1

Cuadro A.3: Aridades máximas registradas para un grafo de 10 000 nodos Parte 1.

τ	**1**	**2**	**3**	**4**	**5**	**6**
1.92	1686	1692	1742	1739	1692	1716

1.93	1674	1676	1654	1721	1662	1616
1.94	1543	1579	1553	1602	1479	1596
1.95	1461	1469	1504	1477	1535	1503
1.96	1491	1438	1459	1477	1386	1494
1.97	1410	1388	1359	1365	1374	1379
1.98	1295	1270	1350	1281	1353	1328
1.99	1237	1334	1297	1299	1233	1237
2.00	1226	1218	1284	1217	1257	1202
2.01	1174	1146	1181	1184	1177	1132
2.02	1137	1119	1150	1080	1200	1110
2.03	1106	1144	1073	1062	1101	1142
2.04	1048	1030	1060	1088	1046	1056
2.05	1053	982	1018	1088	1066	1000
2.06	937	969	954	994	1001	1001
2.07	944	964	928	941	943	957
2.08	900	906	890	895	844	911
2.09	912	910	917	879	909	843
2.10	870	854	880	835	810	874
2.11	831	809	812	867	857	836
2.12	763	814	788	755	784	807
2.13	768	756	757	832	752	785
2.14	742	701	716	782	705	751
2.15	725	741	742	696	735	721
2.16	695	737	685	704	673	710
2.17	692	683	709	681	643	691
2.18	658	646	646	685	656	702
2.19	620	601	642	651	636	660
2.20	628	636	640	616	661	601
2.21	603	632	618	608	581	578
2.22	612	564	571	605	600	611
2.23	560	548	570	579	578	559
2.24	549	547	554	596	565	573

2.25	551	555	564	539	532	566
2.26	515	560	555	564	528	533
2.27	533	534	511	494	574	532
2.28	478	498	526	488	515	499
2.29	485	524	497	523	485	492
2.30	482	452	483	466	483	489
2.31	457	465	483	478	474	479
2.32	458	444	456	449	456	438
2.33	492	472	474	454	473	438
2.34	453	457	421	447	414	461
2.35	467	413	459	462	420	439
2.36	417	436	449	400	393	423
2.37	421	430	426	417	418	425
2.38	403	429	419	403	426	420
2.39	383	414	444	384	403	386
2.40	403	378	407	406	378	405
2.41	422	403	376	407	400	405
2.42	391	375	374	365	377	384
2.43	351	381	408	385	422	376
2.44	386	368	401	387	391	378
2.45	365	389	374	340	365	355
2.46	381	341	354	394	353	372
2.47	390	348	346	355	336	378
2.48	363	343	336	325	375	326
2.49	371	346	341	363	366	333
2.50	349	346	319	347	352	338
2.51	322	326	327	344	338	326
2.52	351	338	347	331	332	325
2.53	318	306	315	322	316	330
2.54	325	319	298	298	299	317
2.55	318	322	305	306	319	313
2.56	308	319	329	312	298	311

2.57	307	311	311	322	327	330
2.58	308	286	291	292	333	295
2.59	300	295	307	276	295	304
2.60	296	301	283	294	304	300
2.61	310	264	293	291	299	303
2.62	331	276	301	305	285	288
2.63	272	324	270	288	257	286
2.64	288	300	294	281	302	268
2.65	290	283	271	277	296	293
2.66	268	302	271	273	280	248
2.67	270	287	255	272	278	278
2.68	268	294	282	273	277	249
2.69	277	259	265	281	259	282
2.70	238	256	280	267	271	272
2.71	262	258	270	275	275	287
2.72	266	268	246	287	268	254

Cuadro A.4: Aridades máximas registradas para un grafo de 10 000 nodos Parte 2.

τ	**7**	**8**	**9**	**10**
1.92	1 690	1 771	1 732	1 735
1.93	1 591	1 625	1 677	1 619
1.94	1 576	1 627	1 640	1 580
1.95	1 559	1 485	1 451	1 516
1.96	1 430	1 390	1 445	1 392
1.97	1 363	1 379	1 378	1 410
1.98	1 356	1 284	1 290	1 324
1.99	1 323	1 297	1 269	1 292
2.00	1 245	1 218	1 250	1 257
2.01	1 152	1 193	1 139	1 214
2.02	1 112	1 128	1 141	1 160

2.03	1 123	1 107	1 135	1 078
2.04	1 036	1 054	1 067	1 041
2.05	1 010	992	1 039	1 000
2.06	937	1 005	1 000	980
2.07	1 000	881	944	887
2.08	899	898	913	943
2.09	851	883	890	884
2.10	850	832	846	860
2.11	837	843	844	825
2.12	791	779	792	842
2.13	761	799	795	730
2.14	719	757	764	754
2.15	736	712	734	754
2.16	667	778	782	727
2.17	661	664	689	679
2.18	656	670	651	676
2.19	645	636	633	610
2.20	616	596	606	593
2.21	582	585	619	608
2.22	566	623	580	538
2.23	570	577	586	601
2.24	578	576	583	570
2.25	576	529	525	526
2.26	529	547	539	529
2.27	509	535	489	532
2.28	529	477	497	507
2.29	502	506	484	522
2.30	456	504	482	499
2.31	466	457	497	488
2.32	465	463	459	490
2.33	472	478	443	453
2.34	402	432	465	461

2.35	436	426	413	453
2.36	422	429	400	434
2.37	432	413	424	454
2.38	418	396	428	399
2.39	418	416	409	420
2.40	402	379	371	418
2.41	384	400	375	399
2.42	362	365	414	391
2.43	370	359	359	381
2.44	390	371	395	362
2.45	356	373	347	375
2.46	334	386	352	357
2.47	353	342	346	365
2.48	345	321	321	350
2.49	351	374	350	338
2.50	316	352	322	313
2.51	337	363	332	352
2.52	334	333	315	334
2.53	326	338	340	356
2.54	316	353	328	313
2.55	325	345	319	318
2.56	301	316	314	313
2.57	299	304	301	308
2.58	290	295	329	283
2.59	316	307	310	285
2.60	294	279	296	301
2.61	285	260	310	307
2.62	309	270	280	264
2.63	287	277	272	276
2.64	262	293	271	300
2.65	297	286	302	293
2.66	277	255	283	284

2.67	288	266	251	309
2.68	263	283	268	267
2.69	265	288	259	288
2.70	290	253	252	252
2.71	278	261	259	255
2.72	272	258	231	262

Cuadro A.5: Aridades máximas registradas para un grafo de 100 000 nodos Parte 1.

τ	**1**	**2**	**3**	**4**	**5**	**6**
1.92	8 172	8 231	8 035	8 306	8 158	8 184
1.93	7 787	7 723	7 856	7 815	7 738	7 824
1.94	7 361	7 324	7 155	7 319	7 194	7 358
1.95	6 912	7 017	6 894	6 901	6 879	6 739
1.96	6 560	6 626	6 663	6 673	6 292	6 498
1.97	6 257	6 170	6 166	6 149	6 177	6 174
1.98	5 824	5 788	5 960	5 864	5 906	5 779
1.99	5 589	5 567	5 556	5 507	5 584	5 540
2.00	5 454	5 237	5 383	5 343	5 218	5 409
2.01	5 060	5 020	5 136	5 031	5 001	5 079
2.02	4 879	4 716	4 803	4 698	4 739	4 802
2.03	4 526	4 453	4 644	4 510	4 540	4 674
2.04	4 376	4 244	4 206	4 422	4 272	4 444
2.05	4 157	4 087	4 186	4 183	4 220	4 046
2.06	4 005	3 876	3 981	3 845	3 882	3 932
2.07	3 839	3 847	3 721	3 837	3 814	3 831
2.08	3 572	3 661	3 603	3 656	3 518	3 500
2.09	3 522	3 398	3 484	3 476	3 554	3 447
2.10	3 237	3 218	3 326	3 389	3 282	3 330
2.11	3 175	3 243	3 158	3 215	3 251	3 221
2.12	2 980	3 079	2 986	3 123	3 027	3 099

2.13	2 923	2 912	2 870	2 965	2 918	2 973
2.14	2 919	2 889	2 871	2 935	2 891	2 854
2.15	2 659	2 787	2 711	2 729	2 768	2 737
2.16	2 509	2 540	2 570	2 508	2 662	2 546
2.17	2 522	2 486	2 640	2 558	2 485	2 552
2.18	2 476	2 435	2 354	2 439	2 427	2 378
2.19	2 258	2 383	2 379	2 249	2 318	2 287
2.20	2 326	2 334	2 272	2 237	2 276	2 288
2.21	2 201	2 264	2 200	2 216	2 230	2 249
2.22	2 072	2 143	2 123	2 190	2 140	2 179
2.23	2 071	2 088	2 103	2 018	2 015	2 005
2.24	2 081	2 016	2 089	2 019	2 039	1 977
2.25	2 023	1 989	2 009	1 968	1 913	1 908
2.26	1 847	1 872	1 900	1 868	1 858	1 851
2.27	1 893	1 866	1 894	1 816	1 783	1 801
2.28	1 849	1 779	1 780	1 802	1 803	1 861
2.29	1 734	1 761	1 752	1 693	1 795	1 708
2.30	1 701	1 712	1 773	1 635	1 734	1 730
2.31	1 685	1 649	1 667	1 650	1 705	1 590
2.32	1 663	1 630	1 615	1 679	1 680	1 636
2.33	1 629	1 574	1 526	1 586	1 521	1 575
2.34	1 613	1 561	1 585	1 600	1 476	1 573
2.35	1 558	1 485	1 495	1 506	1 395	1 540
2.36	1 465	1 415	1 459	1 409	1 446	1 444
2.37	1 511	1 507	1 456	1 486	1 462	1 409
2.38	1 414	1 383	1 431	1 457	1 460	1 450
2.39	1 443	1 382	1 346	1 375	1 469	1 345
2.40	1 377	1 300	1 403	1 355	1 330	1 315
2.41	1 350	1 343	1 329	1 391	1 355	1 359
2.42	1 265	1 248	1 302	1 337	1 313	1 355
2.43	1 285	1 348	1 278	1 253	1 246	1 319
2.44	1 259	1 206	1 328	1 304	1 263	1 264

2.45	1 210	1 199	1 284	1 237	1 206	1 281
2.46	1 289	1 187	1 185	1 250	1 188	1 184
2.47	1 235	1 179	1 209	1 211	1 177	1 180
2.48	1 159	1 197	1 191	1 160	1 188	1 236
2.49	1 184	1 184	1 157	1 240	1 215	1 153
2.50	1 163	1 137	1 116	1 147	1 095	1 207
2.51	1 124	1 205	1 169	1 222	1 178	1 144
2.52	1 113	1 116	1 106	1 128	1 107	1 101
2.53	1 093	1 051	1 067	1 074	1 038	1 133
2.54	1 116	1 087	1 108	1 053	1 097	1 101
2.55	1 061	1 123	1 076	1 060	1 113	1 088
2.56	1 124	1 010	1 005	1 043	1 067	1 112
2.57	1 026	1 048	1 042	1 110	1 019	1 045
2.58	1 062	988	988	995	1 010	1 069
2.59	1 025	1 038	956	1 004	1 015	989
2.60	946	982	976	968	1 003	995
2.61	994	918	1 054	996	1 016	1 025
2.62	975	999	979	918	1 017	1 061
2.63	997	983	963	972	964	995
2.64	951	937	1 007	973	1 003	926
2.65	958	936	904	938	935	995
2.66	905	939	1 000	903	892	909
2.67	920	912	925	925	937	913
2.68	936	887	897	876	923	960
2.69	909	927	894	953	874	899
2.70	830	874	936	958	917	896
2.71	874	919	868	835	858	859
2.72	875	867	897	911	862	849

Cuadro A.6: Aridades máximas registradas para un grafo de 100 000 nodos Parte 2.

τ	**7**	**8**	**9**	**10**	**Promedio**
1.92	8 311	8 179	8 236	8 273	8 209
1.93	7 929	7 752	7 813	7 710	7 795
1.94	7 352	7 404	7 199	7 203	7 287
1.95	6 916	6 993	6 970	6 960	6 918
1.96	6 632	6 465	6 409	6 516	6 533
1.97	6 243	6 226	6 111	6 120	6 179
1.98	5 765	5 863	5 827	5 812	5 839
1.99	5 617	5 467	5 583	5 455	5 547
2.00	5 292	5 375	5 364	5 341	5 342
2.01	5 057	4 998	4 885	4 923	5 019
2.02	4 751	4 883	4 767	4 820	4 786
2.03	4 502	4 538	4 444	4 473	4 530
2.04	4 436	4 230	4 367	4 342	4 334
2.05	4 124	4 033	4 147	4 221	4 140
2.06	3 852	3 960	3 884	3 924	3 914
2.07	3 752	3 864	3 791	3 800	3 810
2.08	3 535	3 565	3 529	3 492	3 563
2.09	3 398	3 422	3 412	3 512	3 463
2.10	3 366	3 308	3 274	3 264	3 299
2.11	3 120	3 234	3 248	3 198	3 206
2.12	3 093	3 037	3 032	3 005	3 046
2.13	2 902	2 925	2 972	2 961	2 932
2.14	2 812	2 819	2 740	2 805	2 854
2.15	2 754	2 698	2 757	2 773	2 737
2.16	2 569	2 667	2 682	2 688	2 594
2.17	2 582	2 625	2 485	2 555	2 549
2.18	2 329	2 506	2 510	2 390	2 424
2.19	2 369	2 319	2 389	2 283	2 323
2.20	2 283	2 264	2 212	2 266	2 276
2.21	2 226	2 294	2 154	2 175	2 221
2.22	2 145	2 117	2 166	2 181	2 146

2.23	2 100	2 129	2 074	2 103	2 071
2.24	1 964	2 038	2 028	2 015	2 027
2.25	1 954	2 003	1 984	1 936	1 969
2.26	2 023	1 885	1 963	1 975	1 904
2.27	1 873	1 890	1 804	1 828	1 845
2.28	1 799	1 835	1 794	1 845	1 815
2.29	1 780	1 713	1 735	1 747	1 742
2.30	1 691	1 703	1 724	1 764	1 717
2.31	1 651	1 719	1 684	1 601	1 660
2.32	1 622	1 594	1 625	1 579	1 632
2.33	1 588	1 581	1 643	1 597	1 582
2.34	1 641	1 594	1 567	1 500	1 571
2.35	1 565	1 490	1 478	1 490	1 500
2.36	1 533	1 532	1 476	1 468	1 465
2.37	1 522	1 469	1 427	1 492	1 474
2.38	1 503	1 356	1 368	1 411	1 423
2.39	1 386	1 338	1 410	1 422	1 392
2.40	1 364	1 348	1 413	1 294	1 350
2.41	1 370	1 379	1 384	1 320	1 358
2.42	1 326	1 359	1 335	1 308	1 315
2.43	1 228	1 282	1 263	1 346	1 285
2.44	1 307	1 288	1 274	1 247	1 274
2.45	1 244	1 286	1 296	1 244	1 249
2.46	1 193	1 205	1 267	1 218	1 217
2.47	1 238	1 230	1 204	1 152	1 202
2.48	1 190	1 235	1 158	1 243	1 196
2.49	1 156	1 184	1 201	1 194	1 187
2.50	1 166	1 199	1 102	1 115	1 145
2.51	1 151	1 125	1 199	1 172	1 169
2.52	1 132	1 054	1 150	1 073	1 108
2.53	1 158	1 076	1 115	1 118	1 092
2.54	1 094	1 098	1 081	1 074	1 091

2.55	1 089	1 075	1 062	1 037	1 078
2.56	1 068	1 034	1 043	1 075	1 058
2.57	1 067	1 075	989	1 067	1 049
2.58	982	996	1 012	999	1 010
2.59	1 008	1 056	1 043	1 034	1 017
2.60	954	1 042	1 008	977	985
2.61	999	1 016	985	1 005	1 001
2.62	963	937	1 011	971	983
2.63	931	1 018	970	980	977
2.64	950	993	944	893	958
2.65	936	998	933	955	949
2.66	914	962	866	918	921
2.67	925	940	976	927	930
2.68	938	910	957	911	920
2.69	931	814	856	922	898
2.70	887	894	909	903	900
2.71	907	889	872	884	877
2.72	843	895	861	866	873

Cuadro A.7: Aridades máximas registradas para un grafo de 1 000 000 nodos Parte 1.

τ	**1**	**2**	**3**	**4**	**5**	**6**
1.92	38 500	38 533	38 300	38 375	38 380	38 506
1.93	35 875	35 502	35 873	35 810	35 922	35 968
1.94	33 475	33 379	33 494	33 561	33 339	33 128
1.95	30 824	30 670	30 923	30 858	30 851	30 976
1.96	28 852	29 013	28 930	28 964	28 866	28 814
1.97	27 253	27 108	26 746	27 033	26 641	27 114
1.98	25 175	25 499	25 186	25 292	24 956	25 170
1.99	23 791	23 639	23 914	23 462	23 654	23 444
2.00	22 200	22 381	22 142	22 396	22 389	22 410
2.01	20 910	21 062	21 050	20 755	20 752	20 938
2.02	19 692	19 824	19 485	19 772	19 643	19 758
2.03	18 402	18 590	18 528	18 242	18 333	18 566
2.04	17 348	17 448	17 187	17 390	17 504	17 482
2.05	16 380	16 385	16 705	16 624	16 458	16 454
2.06	15 586	15 654	15 556	15 389	15 499	15 493
2.07	14 582	14 817	14 821	14 729	14 729	14 912
2.08	13 779	14 147	13 891	13 913	14 023	14 159
2.09	13 424	13 170	13 108	13 319	13 411	13 228
2.10	12 773	12 594	12 729	12 849	12 564	12 796
2.11	12 060	12 084	11 992	12 250	12 133	12 063
2.12	11 695	11 349	11 548	11 468	11 585	11 327
2.13	10 913	11 047	10 934	10 900	11 128	10 976
2.14	10 601	10 511	10 595	10 429	10 352	10 621
2.15	10 053	9 971	10 074	10 049	10 020	10 092
2.16	9 647	9 603	9 691	9 511	9 644	9 724
2.17	9 298	9 280	9 333	9 226	9 271	9 237
2.18	8 812	8 922	9 089	9 006	8 981	8 960
2.19	8 547	8 527	8 603	8 783	8 631	8 602
2.20	8 288	8 609	8 395	8 280	8 183	8 124

2.21	8 071	7 979	7 976	8 145	8 149	8 117
2.22	7 651	7 533	7 670	7 778	7 754	7 558
2.23	7 338	7 541	7 554	7 527	7 360	7 504
2.24	7 304	7 231	7 189	7 356	7 168	7 286
2.25	7 077	7 042	7 053	6 972	7 114	7 112
2.26	6 783	6 799	6 661	6 862	6 664	6 954
2.27	6 567	6 650	6 638	6 576	6 593	6 585
2.28	6 357	6 486	6 394	6 316	6 456	6 334
2.29	6 283	6 189	6 312	6 312	6 232	6 177
2.30	5 997	6 293	6 181	6 115	6 030	6 097
2.31	5 951	5 887	5 926	5 942	5 778	5 941
2.32	5 638	5 804	5 800	5 745	5 739	5 871
2.33	5 480	5 581	5 624	5 488	5 614	5 700
2.34	5 492	5 438	5 402	5 432	5 677	5 412
2.35	5 366	5 340	5 398	5 403	5 443	5 270
2.36	5 227	5 145	5 378	5 227	5 372	5 293
2.37	5 124	5 154	5 186	5 166	4 972	5 112
2.38	5 063	4 898	5 109	4 976	5 004	5 025
2.39	4 779	4 996	4 952	4 875	4 887	4 914
2.40	4 742	4 955	4 753	4 789	4 794	4 750
2.41	4 627	4 822	4 701	4 710	4 672	4 726
2.42	4 661	4 526	4 674	4 624	4 542	4 641
2.43	4 483	4 581	4 391	4 530	4 474	4 424
2.44	4 492	4 460	4 489	4 553	4 422	4 481
2.45	4 392	4 380	4 309	4 381	4 259	4 412
2.46	4 278	4 327	4 464	4 248	4 291	4 251
2.47	4 250	4 166	4 192	4 131	4 230	4 204
2.48	4 130	4 148	4 240	4 135	4 154	4 233
2.49	4 092	4 010	4 126	4 208	4 144	4 172
2.50	4 102	4 078	4 041	4 005	3 995	4 051
2.51	3 902	3 966	3 973	3 889	3 976	3 986
2.52	3 848	3 876	3 870	3 843	3 932	3 856

2.53	3 862	3 873	3 799	3 800	3 898	3 778
2.54	3 769	3 727	3 695	3 722	3 695	3 674
2.55	3 872	3 676	3 804	3 656	3 766	3 814
2.56	3 640	3 733	3 698	3 576	3 705	3 737
2.57	3 632	3 682	3 733	3 643	3 616	3 704
2.58	3 684	3 598	3 627	3 541	3 528	3 600
2.59	3 569	3 545	3 616	3 481	3 410	3 563
2.60	3 446	3 574	3 426	3 361	3 554	3 473
2.61	3 363	3 537	3 497	3 508	3 520	3 503
2.62	3 375	3 320	3 487	3 539	3 382	3 577
2.63	3 407	3 383	3 276	3 352	3 302	3 386
2.64	3 392	3 314	3 352	3 363	3 385	3 358
2.65	3 366	3 313	3 256	3 350	3 325	3 368
2.66	3 207	3 286	3 299	3 313	3 289	3 201
2.67	3 254	3 222	3 200	3 199	3 210	3 194
2.68	3 215	3 183	3 267	3 092	3 165	3 184
2.69	3 238	3 200	3 165	3 252	3 147	3 130
2.70	3 064	3 106	3 085	3 256	3 104	3 141
2.71	3 178	3 042	3 060	3 120	3 073	3 178
2.72	2 971	2 970	2 996	3 055	3 103	3 109

Cuadro A.8: Aridades máximas registradas para un grafo de 1 000 000 nodos Parte 2.

τ	**7**	**8**	**9**	**10**	**Promedio**
1.92	38 573	38 508	38 407	38 514	38 460
1.93	35 942	35 526	35 348	35 994	35 776
1.94	33 213	32 810	33 227	33 026	33 265
1.95	31 251	30 752	30 890	30 974	30 897
1.96	29 021	28 853	28 811	28 906	28 903
1.97	27 024	27 126	26 809	26 925	26 978
1.98	25 098	25 031	25 122	25 370	25 190

1.99	23 446	23 780	23 460	23 851	23 644
2.00	22 520	22 516	22 307	22 341	22 360
2.01	20 665	20 780	20 541	20 619	20 807
2.02	19 548	19 655	19 561	19 461	19 640
2.03	18 563	18 608	18 600	18 525	18 496
2.04	17 336	17 584	17 592	17 307	17 418
2.05	16 630	16 612	16 450	16 534	16 523
2.06	15 485	15 692	15 351	15 366	15 507
2.07	14 900	14 661	14 733	14 891	14 778
2.08	14 015	14 081	13 938	13 948	13 989
2.09	13 282	13 159	13 204	13 276	13 258
2.10	12 812	12 749	12 667	12 686	12 722
2.11	12 272	12 081	11 980	12 006	12 092
2.12	11 489	11 574	11 599	11 404	11 504
2.13	10 966	11 016	10 752	11 018	10 965
2.14	10 721	10 606	10 490	10 288	10 521
2.15	10 026	10 140	10 081	10 154	10 066
2.16	9 739	9 598	9 584	9 865	9 661
2.17	9 319	9 203	9 500	9 280	9 295
2.18	8 949	9 094	9 013	8 852	8 968
2.19	8 437	8 639	8 699	8 740	8 621
2.20	8 387	8 209	8 386	8 389	8 325
2.21	8 027	7 970	7 995	8 064	8 049
2.22	7 622	7 749	7 800	7 762	7 688
2.23	7 417	7 535	7 371	7 380	7 453
2.24	7 315	7 254	7 137	7 239	7 248
2.25	7 086	7 009	6 994	7 033	7 049
2.26	6 689	6 820	6 841	6 703	6 778
2.27	6 637	6 493	6 542	6 607	6 589
2.28	6 393	6 411	6 392	6 366	6 391
2.29	6 226	6 269	6 344	6 229	6 257
2.30	6 129	6 189	6 112	6 058	6 120

2.31	5 888	5 938	5 975	5 909	5 914
2.32	5 712	5 847	5 770	5 772	5 770
2.33	5 731	5 655	5 669	5 785	5 633
2.34	5 435	5 455	5 541	5 578	5 486
2.35	5 365	5 326	5 441	5 376	5 373
2.36	5 397	5 334	5 185	5 316	5 287
2.37	5 226	5 167	5 234	5 064	5 141
2.38	5 080	4 949	5 037	4 925	5 007
2.39	4 922	4 788	4 909	4 863	4 889
2.40	4 910	4 815	4 797	4 974	4 828
2.41	4 777	4 682	4 749	4 657	4 712
2.42	4 636	4 631	4 485	4 572	4 599
2.43	4 660	4 546	4 597	4 430	4 512
2.44	4 385	4 478	4 455	4 429	4 464
2.45	4 407	4 385	4 307	4 426	4 366
2.46	4 212	4 189	4 191	4 317	4 277
2.47	4 152	4 398	4 100	4 217	4 204
2.48	4 111	4 252	4 127	4 200	4 173
2.49	4 151	4 079	4 048	4 098	4 113
2.50	3 907	4 019	4 008	4 049	4 026
2.51	4 000	4 021	3 978	3 885	3 958
2.52	3 939	3 853	3 902	3 791	3 871
2.53	3 885	3 856	3 876	3 841	3 847
2.54	3 893	3 804	3 686	3 767	3 743
2.55	3 665	3 772	3 865	3 689	3 758
2.56	3 604	3 803	3 766	3 633	3 690
2.57	3 607	3 721	3 603	3 615	3 656
2.58	3 582	3 629	3 601	3 609	3 600
2.59	3 441	3 461	3 495	3 479	3 506
2.60	3 519	3 411	3 435	3 428	3 463
2.61	3 396	3 500	3 461	3 324	3 461
2.62	3 381	3 445	3 391	3 337	3 423

2.63	3 412	3 438	3 359	3 372	3 369
2.64	3 372	3 372	3 233	3 403	3 354
2.65	3 343	3 262	3 414	3 289	3 329
2.66	3 241	3 400	3 337	3 333	3 291
2.67	3 178	3 190	3 267	3 127	3 204
2.68	3 251	3 226	3 243	3 077	3 190
2.69	3 121	3 188	3 215	3 197	3 185
2.70	3 011	3 027	3 169	3 038	3 100
2.71	3 083	3 034	3 108	2 997	3 087
2.72	3 006	3 096	3 030	3 085	3 042

Cuadro A.9: Aridades máximas registradas para un grafo de 10 000 000 nodos Parte 1.

τ	**1**	**2**	**3**	**4**	**5**	**6**
1.92	117 234	117 245	117 865	117 135	117 398	117 192
1.93	107 546	107 145	107 429	106 954	107 845	107 198
1.94	98 868	98 103	97 854	98 286	98 273	98 840
1.95	89 933	90 426	90 379	90 058	90 482	90 198
1.96	82 983	82 832	83 214	82 992	82 543	83 189
1.97	75 986	76 268	76 704	75 982	76 423	76 163
1.98	70 292	70 197	70 344	70 539	70 539	70 582
1.99	64 901	64 997	65 351	65 659	65 375	65 201
2.00	60 802	61 168	60 740	61 049	60 809	60 723
2.01	56 534	56 142	55 822	55 795	56 593	55 758
2.02	52 254	52 034	52 419	52 092	52 637	52 311
2.03	49 049	48 804	49 393	48 994	48 486	48 978
2.04	45 761	45 221	45 677	45 320	45 744	45 451
2.05	42 480	42 733	42 509	42 645	42 879	42 764
2.06	40 288	39 639	39 638	40 040	39 647	40 079
2.07	37 626	37 654	37 487	37 422	37 601	37 656
2.08	35 626	35 235	35 246	35 930	35 429	35 639

2.09	33 242	33 337	33 316	33 598	33 175	33 565
2.10	31 537	31 689	31 896	31 846	31 614	31 604
2.11	29 924	29 998	29 827	30 020	29 990	29 965
2.12	28 645	28 337	28 422	28 617	28 602	28 256
2.13	27 011	26 964	27 277	27 063	27 032	27 214
2.14	25 694	25 743	25 975	25 576	25 690	25 974
2.15	24 308	24 561	24 672	24 317	24 633	24 439
2.16	23 440	23 439	23 370	23 551	23 746	23 430
2.17	22 285	22 450	22 538	22 656	22 521	22 574
2.18	21 467	21 678	21 710	21 368	21 562	21 554
2.19	20 790	20 733	20 403	20 583	20 881	20 663
2.20	20 039	19 818	20 029	19 930	19 801	19 970
2.21	19 363	19 056	19 222	19 138	19 264	19 215
2.22	18 684	18 414	18 376	18 349	18 253	18 268
2.23	18 059	17 903	17 790	17 682	17 774	17 861
2.24	17 059	17 060	17 173	17 117	17 050	17 113
2.25	16 626	16 833	16 426	16 861	16 475	16 666
2.26	16 236	16 138	16 102	15 866	15 931	16 105
2.27	15 622	15 507	15 360	15 422	15 583	15 802
2.28	15 350	15 206	14 987	14 981	15 096	15 223
2.29	14 833	14 890	14 882	14 877	14 642	14 478
2.30	14 167	14 405	13 984	14 279	14 319	14 319
2.31	13 785	13 990	13 741	13 883	13 972	13 845
2.32	13 676	13 561	13 619	13 601	13 504	13 707
2.33	13 223	13 030	13 311	13 260	13 297	13 297
2.34	12 934	12 878	12 841	12 807	12 719	12 805
2.35	12 314	12 640	12 564	12 577	12 458	12 492
2.36	12 440	12 384	12 183	12 231	12 275	12 273
2.37	11 934	12 088	11 964	12 105	12 063	11 794
2.38	11 924	11 785	11 759	11 975	11 891	11 803
2.39	11 395	11 432	11 471	11 601	11 725	11 383
2.40	11 474	11 248	11 375	11 165	11 372	11 306

2.41	11 055	11 101	10 950	10 935	10 988	10 984
2.42	10 708	11 005	10 646	10 956	10 742	10 739
2.43	10 539	10 596	10 677	10 560	10 434	10 675
2.44	10 168	10 291	10 522	10 390	10 419	10 477
2.45	10 212	10 179	10 226	10 329	10 236	10 471
2.46	9 982	9 952	10 094	9 961	10 043	10 038
2.47	9 860	9 860	9 765	9 845	9 821	10 067
2.48	9 898	9 602	9 787	9 660	9 681	9 563
2.49	9 642	9 628	9 442	9 434	9 588	9 405
2.50	9 324	9 520	9 546	9 421	9 275	9 556
2.51	9 268	9 277	9 066	9 293	9 037	9 242
2.52	9 047	8 970	9 076	9 261	9 254	9 271
2.53	8 880	8 925	9 005	8 806	9 111	8 986
2.54	8 877	8 804	8 788	8 917	8 819	9 032
2.55	8 633	8 774	8 881	8 775	8 752	8 770
2.56	8 814	8 628	8 354	8 663	8 729	8 647
2.57	8 411	8 550	8 582	8 566	8 505	8 507
2.58	8 383	8 408	8 535	8 346	8 506	8 265
2.59	8 125	8 316	8 226	8 218	8 268	8 236
2.60	8 121	8 157	8 141	8 016	8 146	8 075
2.61	8 095	7 982	8 193	8 111	7 941	7 928
2.62	7 955	7 981	7 885	7 996	7 899	8 015
2.63	7 901	7 858	7 900	7 925	7 986	7 967
2.64	7 837	7 570	7 669	7 757	7 771	7 708
2.65	7 645	7 674	7 608	7 813	7 587	7 670
2.66	7 574	7 624	7 434	7 451	7 722	7 579
2.67	7 326	7 622	7 407	7 430	7 455	7 527
2.68	7 293	7 541	7 505	7 575	7 448	7 650
2.69	7 389	7 679	7 324	7 354	7 278	7 351
2.70	7 323	7 300	7 236	7 401	7 241	7 277
2.71	7 003	7 250	7 266	7 219	7 213	7 167
2.72	7 188	7 155	7 221	7 227	7 102	7 097

Cuadro A.10: Aridades máximas registradas para un grafo de 10 000 000 nodos Parte 2.

τ	**7**	**8**	**9**	**10**	**Promedio**
1.92	117 082	116 711	117 157	117 033	117 205
1.93	106 970	107 212	106 992	107 338	107 263
1.94	98 049	98 410	98 724	97 885	98 329
1.95	89 819	90 604	90 418	90 199	90 252
1.96	82 422	83 143	83 144	82 746	82 921
1.97	76 654	76 827	76 645	76 039	76 369
1.98	70 435	70 205	70 432	70 398	70 396
1.99	65 270	65 134	65 552	65 522	65 296
2.00	60 607	61 092	60 526	60 939	60 846
2.01	56 265	55 766	56 312	56 547	56 153
2.02	52 211	52 112	52 394	52 448	52 291
2.03	48 645	48 906	48 747	48 581	48 858
2.04	45 489	45 389	45 483	45 450	45 499
2.05	42 485	42 975	42 961	42 836	42 727
2.06	40 407	39 931	40 277	40 012	39 996
2.07	37 172	37 909	37 315	37 485	37 533
2.08	35 553	35 758	35 597	35 181	35 519
2.09	33 481	33 412	33 517	33 407	33 405
2.10	31 598	31 495	31 952	31 812	31 704
2.11	30 281	30 018	30 004	29 757	29 978
2.12	28 508	28 535	28 486	28 712	28 512
2.13	27 096	27 152	27 033	27 354	27 120
2.14	25 902	25 582	25 800	25 969	25 791
2.15	24 433	24 625	24 206	24 528	24 472
2.16	23 617	23 761	23 343	23 564	23 526
2.17	22 219	22 765	22 461	22 679	22 515
2.18	21 554	21 580	21 398	21 725	21 560
2.19	20 733	20 908	20 861	20 623	20 718
2.20	19 893	19 846	19 947	20 085	19 936

2.21	19 178	19 238	19 053	19 192	19 192
2.22	18 444	18 488	18 462	18 308	18 405
2.23	17 939	17 968	17 739	17 963	17 868
2.24	17 159	17 249	17 170	17 340	17 149
2.25	16 362	16 611	16 765	16 739	16 636
2.26	16 146	16 071	16 015	16 032	16 064
2.27	15 694	15 534	15 674	15 691	15 589
2.28	15 064	15 221	14 980	15 122	15 123
2.29	14 772	14 492	14 707	14 665	14 724
2.30	14 200	14 268	14 327	14 169	14 244
2.31	13 811	13 859	13 896	13 927	13 871
2.32	13 435	13 478	13 457	13 518	13 556
2.33	13 294	13 207	13 339	13 288	13 255
2.34	13 031	12 701	13 043	12 715	12 847
2.35	12 665	12 517	12 671	12 573	12 547
2.36	12 363	12 162	12 115	12 256	12 268
2.37	11 986	11 969	12 010	12 028	11 994
2.38	11 490	11 551	11 755	11 856	11 779
2.39	11 500	11 252	11 721	11 474	11 495
2.40	11 273	11 170	11 057	11 215	11 266
2.41	10 993	11 035	10 813	11 194	11 005
2.42	10 781	10 539	10 624	10 789	10 753
2.43	10 798	10 693	10 526	10 711	10 621
2.44	10 377	10 393	10 319	10 292	10 365
2.45	10 072	10 330	10 225	10 162	10 244
2.46	9 991	10 098	9 985	10 040	10 018
2.47	9 823	10 017	9 865	9 888	9 881
2.48	9 710	9 591	9 774	9 713	9 698
2.49	9 618	9 465	9 369	9 454	9 505
2.50	9 363	9 297	9 477	9 438	9 422
2.51	9 107	9 107	9 030	9 224	9 165
2.52	9 045	9 002	9 089	9 091	9 111

2.53	8 859	9 073	9 078	8 960	8 968
2.54	8 817	8 788	8 812	8 840	8 849
2.55	8 802	8 694	8 733	8 887	8 770
2.56	8 574	8 811	8 575	8 688	8 648
2.57	8 472	8 561	8 508	8 395	8 506
2.58	8 493	8 275	8 366	8 274	8 385
2.59	8 265	8 290	8 299	8 221	8 246
2.60	8 230	8 264	8 269	8 105	8 152
2.61	7 961	8 152	7 913	8 134	8 041
2.62	8 131	7 970	8 041	8 081	7 995
2.63	7 847	7 849	7 824	7 872	7 893
2.64	7 662	7 664	7 965	7 708	7 731
2.65	7 573	7 755	7 596	7 732	7 665
2.66	7 559	7 375	7 497	7 502	7 532
2.67	7 436	7 564	7 426	7 456	7 465
2.68	7 268	7 466	7 476	7 267	7 449
2.69	7 318	7 272	7 461	7 378	7 380
2.70	7 317	7 299	7 225	7 467	7 309
2.71	7 106	7 215	7 213	7 065	7 172
2.72	7 117	7 225	7 041	7 195	7 157

A.2.2. Aridades máximas promedio registrados por tamaño del grafo y por τ

Cuadro A.11: Aridades máximas promedio registrados por tamaño del grafo y por τ.

τ	**n=1k**	**n=10k**	**n=100k**	**n=1M**	**n=10M**
1.92	528.7	1 719.5	8 208.5	38 459.6	117 205.2
1.93	514.6	1 651.5	7 794.7	35 776.0	107 262.9
1.94	510.5	1 577.5	7 286.9	33 265.2	98 329.2
1.95	493.8	1 496.0	6 918.1	30 896.9	90 251.6

1.96	470.8	1 440.2	6 533.4	28 903.0	82 920.8
1.97	446.3	1 380.5	6 179.3	26 977.9	76 369.1
1.98	463.8	1 313.1	5 838.8	25 189.9	70 396.3
1.99	426.3	1 281.8	5 546.5	23 644.1	65 296.2
2.00	419.5	1 237.4	5 341.6	22 360.2	60 845.5
2.01	408.2	1 169.2	5 019.0	20 807.2	56 153.4
2.02	394.7	1 133.7	4 785.8	19 639.9	52 291.2
2.03	397.5	1 107.1	4 530.4	18 495.7	48 858.3
2.04	368.3	1 052.6	4 333.9	17 417.8	45 498.5
2.05	367.9	1 024.8	4 140.4	16 523.2	42 726.7
2.06	351.6	977.8	3 914.1	15 507.1	39 995.8
2.07	349.6	938.9	3 809.6	14 777.5	37 532.7
2.08	339.5	899.9	3 563.1	13 989.4	35 519.4
2.09	327.7	887.8	3 462.5	13 258.1	33 405.0
2.10	327.5	851.1	3 299.4	12 721.9	31 704.3
2.11	304.9	836.1	3 206.3	12 092.1	29 978.4
2.12	312.2	791.5	3 046.1	11 503.8	28 512.0
2.13	302.2	773.5	2 932.1	10 965.0	27 119.6
2.14	294.1	739.1	2 853.5	10 521.4	25 790.5
2.15	285.3	729.6	2 737.3	10 066.0	24 472.2
2.16	274.0	715.8	2 594.1	9 660.6	23 526.1
2.17	270.6	679.2	2 549.0	9 294.7	22 514.8
2.18	264.9	664.6	2 424.4	8 967.8	21 559.6
2.19	263.6	633.4	2 323.4	8 620.8	20 717.8
2.20	256.8	619.3	2 275.8	8 325.0	19 935.8
2.21	247.0	601.4	2 220.9	8 049.3	19 191.9
2.22	236.1	587.0	2 145.6	7 687.7	18 404.6
2.23	241.0	572.8	2 070.6	7 452.7	17 867.8
2.24	233.9	569.1	2 026.6	7 247.9	17 149.0
2.25	232.3	546.3	1 968.7	7 049.2	16 636.4
2.26	225.6	539.9	1 904.2	6 777.6	16 064.2
2.27	229.7	524.3	1 844.8	6 588.8	15 588.9

2.28	215.8	501.4	1 814.7	6 390.5	15 123.0
2.29	209.8	502.0	1 741.8	6 257.3	14 723.8
2.30	208.4	479.6	1 716.7	6 120.1	14 243.7
2.31	207.1	474.4	1 660.1	5 913.5	13 870.9
2.32	205.5	457.8	1 632.3	5 769.8	13 555.6
2.33	197.1	464.9	1 582.0	5 632.7	13 254.6
2.34	182.7	441.3	1 571.0	5 486.2	12 847.4
2.35	186.3	438.8	1 500.2	5 372.8	12 547.1
2.36	191.6	420.3	1 464.7	5 287.4	12 268.2
2.37	186.7	426.0	1 474.1	5 140.5	11 994.1
2.38	180.8	414.1	1 423.3	5 006.6	11 778.9
2.39	181.4	407.7	1 391.6	4 888.5	11 495.4
2.40	178.0	394.7	1 349.9	4 827.9	11 265.5
2.41	170.7	397.1	1 358.0	4 712.3	11 004.8
2.42	172.7	379.8	1 314.8	4 599.2	10 752.9
2.43	163.5	379.2	1 284.8	4 511.6	10 620.9
2.44	171.6	382.9	1 274.0	4 464.4	10 364.8
2.45	165.1	363.9	1 248.7	4 365.8	10 244.2
2.46	165.6	362.4	1 216.6	4 276.8	10 018.4
2.47	154.4	355.9	1 201.5	4 204.0	9 881.1
2.48	147.8	340.5	1 195.7	4 173.0	9 697.9
2.49	156.5	353.3	1 186.8	4 112.8	9 504.5
2.50	144.0	335.4	1 144.7	4 025.5	9 421.7
2.51	145.6	336.7	1 168.9	3 957.6	9 165.1
2.52	146.5	334.0	1 108.0	3 871.0	9 110.6
2.53	146.7	326.7	1 092.3	3 846.8	8 968.3
2.54	145.1	316.6	1 090.9	3 743.2	8 849.4
2.55	145.2	319.0	1 078.4	3 757.9	8 770.1
2.56	139.2	312.1	1 058.1	3 689.5	8 648.3
2.57	134.4	312.0	1 048.8	3 655.6	8 505.7
2.58	132.2	300.2	1 010.1	3 599.9	8 385.1
2.59	131.7	299.5	1 016.8	3 506.0	8 246.4

2.60	131.7	294.8	985.1	3 462.7	8 152.4
2.61	134.9	292.2	1 000.8	3 460.9	8 041.0
2.62	132.8	290.9	983.1	3 423.4	7 995.4
2.63	128.7	280.9	977.3	3 368.7	7 892.9
2.64	124.6	285.9	957.7	3 354.4	7 731.1
2.65	132.5	288.8	948.8	3 328.6	7 665.3
2.66	118.8	274.1	920.8	3 290.6	7 531.7
2.67	126.7	275.4	930.0	3 204.1	7 464.9
2.68	128.8	272.4	919.5	3 190.3	7 448.9
2.69	119.3	272.3	897.9	3 185.3	7 380.4
2.70	120.2	263.1	900.4	3 100.1	7 308.6
2.71	116.3	268.0	876.5	3 087.3	7 171.7
2.72	120.1	261.2	872.6	3 042.1	7 156.8

A.2.3. Registro de la Constate de Proporcionalidad y el Exponente de la potencia.

Cuadro A.12: Constante de proporcionalidad y exponente de potencia por cada valor de τ.

τ	**Constante de proporcionalidad**	**Exponente de la potencia**
1.92	7.6706	0.6041
1.93	7.8379	0.5974
1.94	8.1272	0.5893
1.95	8.1540	0.5839
1.96	8.0931	0.5794
1.97	7.9456	0.5758
1.98	8.6549	0.5645
1.99	8.2397	0.5636
2.00	8.4206	0.5580
2.01	8.4281	0.5527

2.02	8.4523	0.5483
2.03	8.9144	0.5402
2.04	8.3889	0.5402
2.05	8.7045	0.5337
2.06	8.4760	0.5312
2.07	8.6873	0.5259
2.08	8.5349	0.5231
2.09	8.5996	0.5191
2.10	8.7230	0.5146
2.11	8.3478	0.5146
2.12	8.6439	0.5084
2.13	8.5720	0.5058
2.14	8.4250	0.5039
2.15	8.4370	0.5007
2.16	8.1981	0.4998
2.17	8.1287	0.4977
2.18	8.0852	0.4951
2.19	8.0500	0.4925
2.20	7.9698	0.4909
2.21	7.7221	0.4907
2.22	7.4884	0.4901
2.23	7.7427	0.4854
2.24	7.7156	0.4835
2.25	7.6385	0.4821
2.26	7.5606	0.4804
2.27	7.7695	0.4763
2.28	7.2008	0.4797
2.29	7.1044	0.4788
2.30	7.0357	0.4775
2.31	7.1041	0.4748
2.32	7.0204	0.4739
2.33	6.8787	0.4739

2.34	6.2489	0.4789
2.35	6.4741	0.4745
2.36	6.6148	0.4712
2.37	6.6556	0.4697
2.38	6.3756	0.4710
2.39	6.4734	0.4683
2.40	6.2723	0.4690
2.41	6.1554	0.4693
2.42	6.1686	0.4672
2.43	5.8340	0.4701
2.44	6.3568	0.4629
2.45	5.9275	0.4665
2.46	6.0445	0.4635
2.47	5.5630	0.4685
2.48	5.2016	0.4722
2.49	5.8265	0.4633
2.50	5.1159	0.4711
2.51	5.3639	0.4668
2.52	5.3771	0.4651
2.53	5.3653	0.4644
2.54	5.2773	0.4643
2.55	5.3294	0.4633
2.56	5.0541	0.4659
2.57	4.9144	0.4671
2.58	4.7298	0.4683
2.59	4.8081	0.4662
2.60	4.7865	0.4653
2.61	4.9666	0.4624
2.62	4.8794	0.4630
2.63	4.6501	0.4654
2.64	4.5913	0.4655
2.65	5.0143	0.4586

2.66	4.2899	0.4684
2.67	4.7301	0.4606
2.68	4.7911	0.4593
2.69	4.3827	0.4651
2.70	4.3891	0.4639
2.71	4.3208	0.4642
2.72	4.4337	0.4617

A.2.4. Registro de las Aridades Máximas Obtenidas en la Simulación

Cuadro A.13: Aridades máximas obtenidas de la fórmula de la aridad máxima.

τ	**1 000**	**10 000**	**100 000**	**1 000 000**	**10 000 000**
1.92	506.593	2 034.288	8 168.948	32 803.471	131 726.586
1.93	487.234	1 924.261	7 599.597	30 013.535	118 534.223
1.94	469.420	1 824.806	7 093.690	27 575.772	107 197.127
1.95	452.978	1 734.600	6 642.355	25 435.764	97 401.914
1.96	437.757	1 652.511	6 238.153	23 548.739	88 895.406
1.97	423.625	1 577.568	5 874.819	21 877.665	81 471.826
1.98	410.467	1 508.934	5 547.055	20 391.753	74 962.952
1.99	398.181	1 445.889	5 250.360	19 065.284	69 230.506
2.00	386.678	1 387.804	4 980.895	17 876.669	64 160.214
2.01	375.877	1 334.136	4 735.375	16 807.707	59 657.163
2.02	365.710	1 284.410	4 510.976	15 842.989	55 642.132
2.03	356.114	1 238.211	4 305.263	14 969.412	52 048.691
2.04	347.034	1 195.173	4 116.129	14 175.792	48 820.885
2.05	338.421	1 154.976	3 941.748	13 452.549	45 911.380
2.06	330.231	1 117.339	3 780.526	12 791.445	43 279.971
2.07	322.425	1 082.011	3 631.075	12 185.371	40 892.371
2.08	314.968	1 048.772	3 492.177	11 628.173	38 719.232
2.09	307.829	1 017.426	3 362.765	11 114.509	36 735.342
2.10	300.980	987.799	3 241.898	10 639.724	34 918.966

2.11	294.396	959.736	3 128.748	10 199.753	33 251.305
2.12	288.056	933.099	3 022.582	9 791.035	31 716.049
2.13	281.939	907.765	2 922.751	9 410.443	30 299.001
2.14	276.027	883.625	2 828.678	9 055.224	28 987.771
2.15	270.305	860.579	2 739.851	8 722.948	27 771.516
2.16	264.758	838.540	2 655.814	8 411.468	26 640.718
2.17	259.374	817.428	2 576.160	8 118.880	25 587.004
2.18	254.141	797.173	2 500.524	7 843.492	24 602.990
2.19	249.048	777.711	2 428.581	7 583.799	23 682.150
2.20	244.089	758.985	2 360.039	7 338.461	22 818.700
2.21	239.253	740.944	2 294.636	7 106.280	22 007.506
2.22	234.534	723.541	2 232.138	6 886.185	21 244.003
2.23	229.926	706.736	2 172.333	6 677.216	20 524.118
2.24	225.423	690.491	2 115.030	6 478.512	19 844.219
2.25	221.021	674.771	2 060.058	6 289.299	19 201.054
2.26	216.715	659.548	2 007.262	6 108.882	18 591.708
2.27	212.501	644.794	1 956.504	5 936.633	18 013.569
2.28	208.377	630.484	1 907.655	5 771.987	17 464.284
2.29	204.338	616.597	1 860.603	5 614.431	16 941.737
2.30	200.383	603.112	1 815.243	5 463.505	16 444.020
2.31	196.510	590.012	1 771.482	5 318.790	15 969.410
2.32	192.716	577.280	1 729.235	5 179.905	15 516.346
2.33	189.001	564.902	1 688.425	5 046.506	15 083.417
2.34	185.362	552.864	1 648.981	4 918.279	14 669.341
2.35	181.799	541.155	1 610.839	4 794.939	14 272.956
2.36	178.310	529.763	1 573.941	4 676.225	13 893.202
2.37	174.894	518.679	1 538.233	4 561.900	13 529.114
2.38	171.551	507.894	1 503.667	4 451.746	13 179.811
2.39	168.281	497.399	1 470.197	4 345.563	12 844.488
2.40	165.081	487.187	1 437.782	4 243.169	12 522.408
2.41	161.953	477.250	1 406.384	4 144.395	12 212.894
2.42	158.895	467.584	1 375.968	4 049.086	11 915.325

2.43	155.907	458.180	1 346.501	3 957.098	11 629.127
2.44	152.989	449.034	1 317.953	3 868.299	11 353.773
2.45	150.139	440.141	1 290.296	3 782.565	11 088.775
2.46	147.359	431.495	1 263.503	3 699.781	10 833.679
2.47	144.647	423.092	1 237.549	3 619.841	10 588.066
2.48	142.002	414.927	1 212.411	3 542.643	10 351.544
2.49	139.425	406.996	1 188.066	3 468.095	10 123.749
2.50	136.914	399.295	1 164.494	3 396.108	9 904.339
2.51	134.470	391.818	1 141.675	3 326.597	9 692.996
2.52	132.092	384.563	1 119.589	3 259.485	9 489.417
2.53	129.779	377.526	1 098.217	3 194.696	9 293.320
2.54	127.531	370.702	1 077.543	3 132.159	9 104.438
2.55	125.347	364.088	1 057.548	3 071.805	8 922.517
2.56	123.225	357.679	1 038.216	3 013.570	8 747.316
2.57	121.167	351.473	1 019.531	2 957.390	8 578.608
2.58	119.170	345.465	1 001.476	2 903.205	8 416.174
2.59	117.234	339.651	984.037	2 850.957	8 259.806
2.60	115.358	334.027	967.199	2 800.591	8 109.304
2.61	113.541	328.590	950.945	2 752.050	7 964.476
2.62	111.783	323.336	935.263	2 705.284	7 825.138
2.63	110.082	318.261	920.137	2 660.240	7 691.113
2.64	108.437	313.361	905.553	2 616.868	7 562.229
2.65	106.848	308.633	891.497	2 575.119	7 438.320
2.66	105.312	304.072	877.955	2 534.946	7 319.227
2.67	103.830	299.674	864.914	2 496.303	7 204.795
2.68	102.401	295.436	852.361	2 459.144	7 094.871
2.69	101.022	291.353	840.281	2 423.424	6 989.311
2.70	99.693	287.422	828.662	2 389.100	6 887.971
2.71	98.413	283.639	817.491	2 356.130	6 790.714
2.72	97.180	280.001	806.755	2 324.471	6 697.405

A.2.5. Registro de la Diferencia entre la Simulación y los Valores Empíricos de la Aridad Máxima.

Cuadro A.14: Diferencia entre la aridad registrada por la simulación y los datos empíricos.

τ	**1 000**	**10 000**	**100 000**	**1 000 000**	**10 000 000**
1.92	-22.11	314.79	-39.55	-5 656.13	14 521.39
1.93	-27.37	272.76	-195.10	-5 762.46	11 271.32
1.94	-41.08	247.31	-193.21	-5 689.43	8 867.93
1.95	-40.82	238.60	-275.74	-5 461.14	7 150.31
1.96	-33.04	212.31	-295.25	-5 354.26	5 974.61
1.97	-22.68	197.07	-304.48	-5 100.23	5 102.73
1.98	-53.33	195.83	-291.75	-4 798.15	4 566.65
1.99	-28.12	164.09	-296.14	-4 578.82	3 934.31
2.00	-32.82	150.40	-360.70	-4 483.53	3 314.71
2.01	-32.32	164.94	-283.63	-3 999.49	3 503.76
2.02	-28.99	150.71	-274.82	-3 796.91	3 350.93
2.03	-41.39	131.11	-225.14	-3 526.29	3 190.39
2.04	-21.27	142.57	-217.77	-3 242.01	3 322.39
2.05	-29.48	130.18	-198.65	-3 070.65	3 184.68
2.06	-21.37	139.54	-133.57	-2 715.65	3 284.17
2.07	-27.18	143.11	-178.53	-2 592.13	3 359.67
2.08	-24.53	148.87	-70.92	-2 361.23	3 199.83
2.09	-19.87	129.63	-99.74	-2 143.59	3 330.34
2.10	-26.52	136.70	-57.50	-2 082.18	3 214.67
2.11	-10.50	123.64	-77.55	-1 892.35	3 272.91
2.12	-24.14	141.60	-23.52	-1 712.76	3 204.05
2.13	-20.26	134.27	-9.35	-1 554.56	3 179.40
2.14	-18.07	144.52	-24.82	-1 466.18	3 197.27
2.15	-14.99	130.98	2.55	-1 343.05	3 299.32
2.16	-9.24	122.74	61.71	-1 249.13	3 114.62
2.17	-11.23	138.23	27.16	-1 175.82	3 072.20

2.18	-10.76	132.57	76.12	-1 124.31	3 043.39
2.19	-14.55	144.31	105.18	-1 037.00	2 964.35
2.20	-12.71	139.69	84.24	-986.54	2 882.90
2.21	-7.75	139.54	73.74	-943.02	2 815.61
2.22	-1.57	136.54	86.54	-801.51	2 839.40
2.23	-11.07	133.94	101.73	-775.48	2 656.32
2.24	-8.48	121.39	88.43	-769.39	2 695.22
2.25	-11.28	128.47	91.36	-759.90	2 564.65
2.26	-8.88	119.65	103.06	-668.72	2 527.51
2.27	-17.20	120.49	111.70	-652.17	2 424.67
2.28	-7.42	129.08	92.95	-618.51	2 341.28
2.29	-5.46	114.60	118.80	-642.87	2 217.94
2.30	-8.02	123.51	98.54	-656.59	2 200.32
2.31	-10.59	115.61	111.38	-594.71	2 098.51
2.32	-12.78	119.48	96.94	-589.90	1 960.75
2.33	-8.10	100.00	106.43	-586.19	1 828.82
2.34	2.66	111.56	77.98	-567.92	1 821.94
2.35	-4.50	102.35	110.64	-577.86	1 725.86
2.36	-13.29	109.46	109.24	-611.17	1 625.00
2.37	-11.81	92.68	64.13	-578.60	1 535.01
2.38	-9.25	93.79	80.37	-554.85	1 400.91
2.39	-13.12	89.70	78.60	-542.94	1 349.09
2.40	-12.92	92.49	87.88	-584.73	1 256.91
2.41	-8.75	80.15	48.38	-567.90	1 208.09
2.42	-13.80	87.78	61.17	-550.11	1 162.42
2.43	-7.59	78.98	61.70	-554.50	1 008.23
2.44	-18.61	66.13	43.95	-596.10	988.97
2.45	-14.96	76.24	41.60	-583.24	844.57
2.46	-18.24	69.10	46.90	-577.02	815.28
2.47	-9.75	67.19	36.05	-584.16	706.97
2.48	-5.80	74.43	16.71	-630.36	653.64
2.49	-17.08	53.70	1.27	-644.70	619.25

2.50	-7.09	63.89	19.79	-629.39	482.64
2.51	-11.13	55.12	-27.22	-631.00	527.90
2.52	-14.41	50.56	11.59	-611.51	378.82
2.53	-16.92	50.83	5.92	-652.10	325.02
2.54	-17.57	54.10	-13.36	-611.04	255.04
2.55	-19.85	45.09	-20.85	-686.09	152.42
2.56	-15.97	45.58	-19.88	-675.93	99.02
2.57	-13.23	39.47	-29.27	-698.21	72.91
2.58	-13.03	45.26	-8.62	-696.70	31.07
2.59	-14.47	40.15	-32.76	-655.04	13.41
2.60	-16.34	39.23	-17.90	-662.11	-43.10
2.61	-21.36	36.39	-49.85	-708.85	-76.52
2.62	-21.02	32.44	-47.84	-718.12	-170.26
2.63	-18.62	37.36	-57.16	-708.46	-201.79
2.64	-16.16	27.46	-52.15	-737.53	-168.87
2.65	-25.65	19.83	-57.30	-753.48	-226.98
2.66	-13.49	29.97	-42.84	-755.65	-212.47
2.67	-22.87	24.27	-65.09	-707.80	-260.11
2.68	-26.40	23.04	-67.14	-731.16	-354.03
2.69	-18.28	19.05	-57.62	-761.88	-391.09
2.70	-20.51	24.32	-71.74	-711.00	-420.63
2.71	-17.89	15.64	-59.01	-731.17	-380.99
2.72	-22.92	18.80	-65.85	-717.63	-459.40

A.2.6. Registro de la Diferencia entre la Simulación y los Valores Empíricos de la Aridad Máxima en Porcentajes.

Cuadro A.15: Diferencia entre la aridad registrada por la simulación y los datos empíricos.

tau	**1 000**	**10 000**	**100 000**	**1 000 000**	**10 000 000**
1.92	-4 %	18 %	0 %	-15 %	12 %
1.93	-5 %	17 %	-3 %	-16 %	11 %

1.94	-8 %	16 %	-3 %	-17 %	9 %
1.95	-8 %	16 %	-4 %	-18 %	8 %
1.96	-7 %	15 %	-5 %	-19 %	7 %
1.97	-5 %	14 %	-5 %	-19 %	7 %
1.98	-11 %	15 %	-5 %	-19 %	6 %
1.99	-7 %	13 %	-5 %	-19 %	6 %
2.00	-8 %	12 %	-7 %	-20 %	5 %
2.01	-8 %	14 %	-6 %	-19 %	6 %
2.02	-7 %	13 %	-6 %	-19 %	6 %
2.03	-10 %	12 %	-5 %	-19 %	7 %
2.04	-6 %	14 %	-5 %	-19 %	7 %
2.05	-8 %	13 %	-5 %	-19 %	7 %
2.06	-6 %	14 %	-3 %	-18 %	8 %
2.07	-8 %	15 %	-5 %	-18 %	9 %
2.08	-7 %	17 %	-2 %	-17 %	9 %
2.09	-6 %	15 %	-3 %	-16 %	10 %
2.10	-8 %	16 %	-2 %	-16 %	10 %
2.11	-3 %	15 %	-2 %	-16 %	11 %
2.12	-8 %	18 %	-1 %	-15 %	11 %
2.13	-7 %	17 %	0 %	-14 %	12 %
2.14	-6 %	20 %	-1 %	-14 %	12 %
2.15	-5 %	18 %	0 %	-13 %	13 %
2.16	-3 %	17 %	2 %	-13 %	13 %
2.17	-4 %	20 %	1 %	-13 %	14 %
2.18	-4 %	20 %	3 %	-13 %	14 %
2.19	-6 %	23 %	5 %	-12 %	14 %
2.20	-5 %	23 %	4 %	-12 %	14 %
2.21	-3 %	23 %	3 %	-12 %	15 %
2.22	-1 %	23 %	4 %	-10 %	15 %
2.23	-5 %	23 %	5 %	-10 %	15 %
2.24	-4 %	21 %	4 %	-11 %	16 %
2.25	-5 %	24 %	5 %	-11 %	15 %

2.26	-4 %	22 %	5 %	-10 %	16 %
2.27	-7 %	23 %	6 %	-10 %	16 %
2.28	-3 %	26 %	5 %	-10 %	15 %
2.29	-3 %	23 %	7 %	-10 %	15 %
2.30	-4 %	26 %	6 %	-11 %	15 %
2.31	-5 %	24 %	7 %	-10 %	15 %
2.32	-6 %	26 %	6 %	-10 %	14 %
2.33	-4 %	22 %	7 %	-10 %	14 %
2.34	1 %	25 %	5 %	-10 %	14 %
2.35	-2 %	23 %	7 %	-11 %	14 %
2.36	-7 %	26 %	7 %	-12 %	13 %
2.37	-6 %	22 %	4 %	-11 %	13 %
2.38	-5 %	23 %	6 %	-11 %	12 %
2.39	-7 %	22 %	6 %	-11 %	12 %
2.40	-7 %	23 %	7 %	-12 %	11 %
2.41	-5 %	20 %	4 %	-12 %	11 %
2.42	-8 %	23 %	5 %	-12 %	11 %
2.43	-5 %	21 %	5 %	-12 %	9 %
2.44	-11 %	17 %	3 %	-13 %	10 %
2.45	-9 %	21 %	3 %	-13 %	8 %
2.46	-11 %	19 %	4 %	-13 %	8 %
2.47	-6 %	19 %	3 %	-14 %	7 %
2.48	-4 %	22 %	1 %	-15 %	7 %
2.49	-11 %	15 %	0 %	-16 %	7 %
2.50	-5 %	19 %	2 %	-16 %	5 %
2.51	-8 %	16 %	-2 %	-16 %	6 %
2.52	-10 %	15 %	1 %	-16 %	4 %
2.53	-12 %	16 %	1 %	-17 %	4 %
2.54	-12 %	17 %	-1 %	-16 %	3 %
2.55	-14 %	14 %	-2 %	-18 %	2 %
2.56	-11 %	15 %	-2 %	-18 %	1 %
2.57	-10 %	13 %	-3 %	-19 %	1 %

2.58	-10 %	15 %	-1 %	-19 %	0 %
2.59	-11 %	13 %	-3 %	-19 %	0 %
2.60	-12 %	13 %	-2 %	-19 %	-1 %
2.61	-16 %	12 %	-5 %	-20 %	-1 %
2.62	-16 %	11 %	-5 %	-21 %	-2 %
2.63	-14 %	13 %	-6 %	-21 %	-3 %
2.64	-13 %	10 %	-5 %	-22 %	-2 %
2.65	-19 %	7 %	-6 %	-23 %	-3 %
2.66	-11 %	11 %	-5 %	-23 %	-3 %
2.67	-18 %	9 %	-7 %	-22 %	-3 %
2.68	-20 %	8 %	-7 %	-23 %	-5 %
2.69	-15 %	7 %	-6 %	-24 %	-5 %
2.70	-17 %	9 %	-8 %	-23 %	-6 %
2.71	-15 %	6 %	-7 %	-24 %	-5 %
2.72	-19 %	7 %	-8 %	-24 %	-6 %

B. Evaluación Experimental

En este anexo se presentan resultados de la Evaluación Experimental, principalmente gráficos que por tema de espacio no se han colocado en el Capítulo 5.

B.1. Realismo

B.1.1. Algoritmo 1

Gráficos de distribución de grados para grafos dirigidos y no dirigidos de 1 000 y 1 000 000 de nodos generados por el Algoritmo 1.

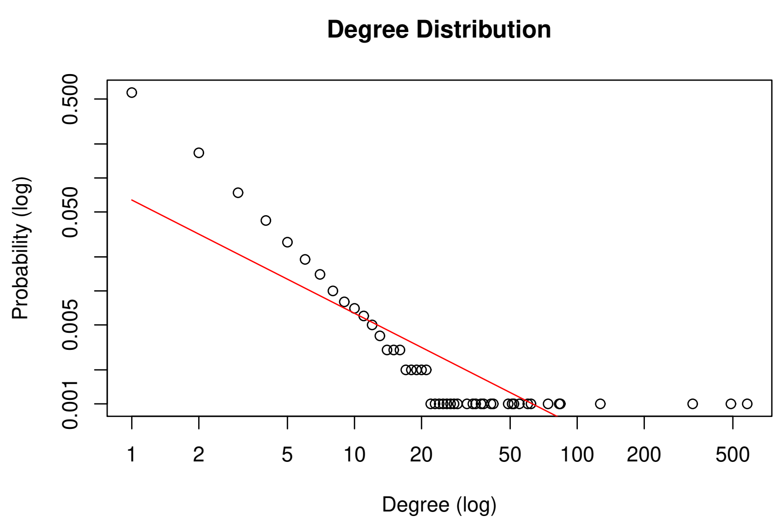


Figura B.1: Distribución de grados, en escala logarítmica, de un grafo no dirigido de mil nodos generados con el Algoritmo 1.

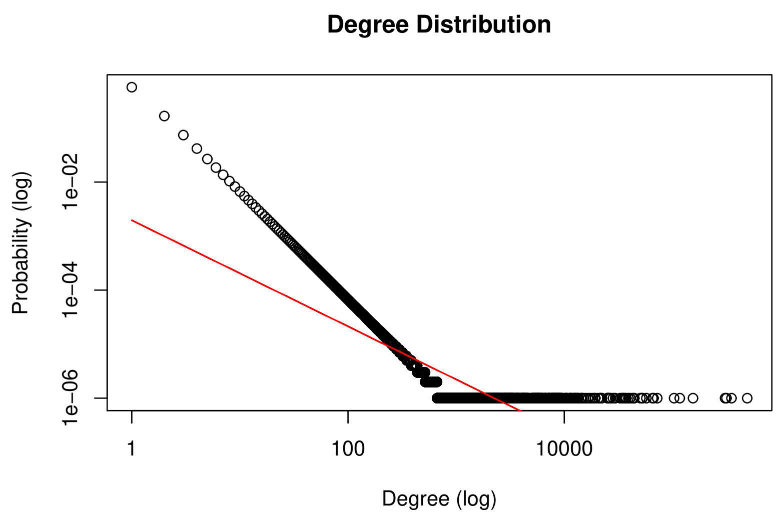


Figura B.2: Distribución de grados, en escala logarítmica, de un grafo no dirigido de un millón de nodos generados con el Algoritmo 1.

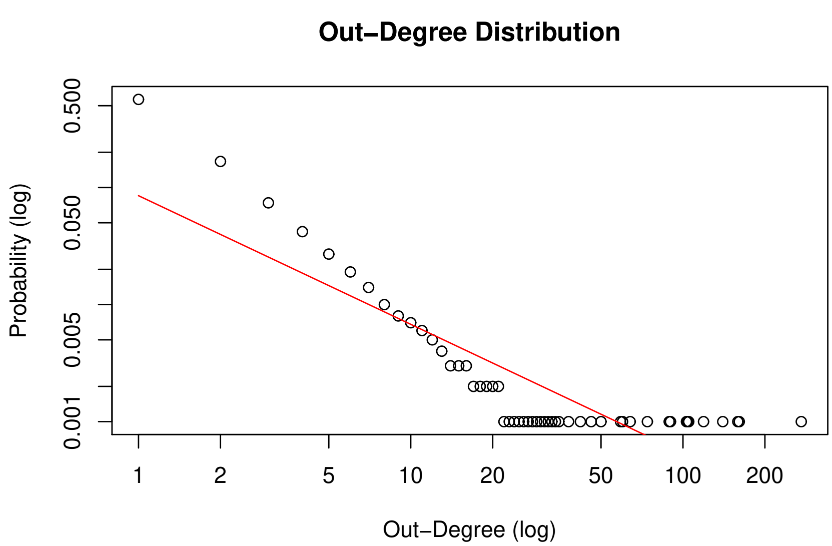


Figura B.3: Distribución de out-degree, en escala logarítmica, de un grafo dirigido de mil nodos generados con el Algoritmo 1.

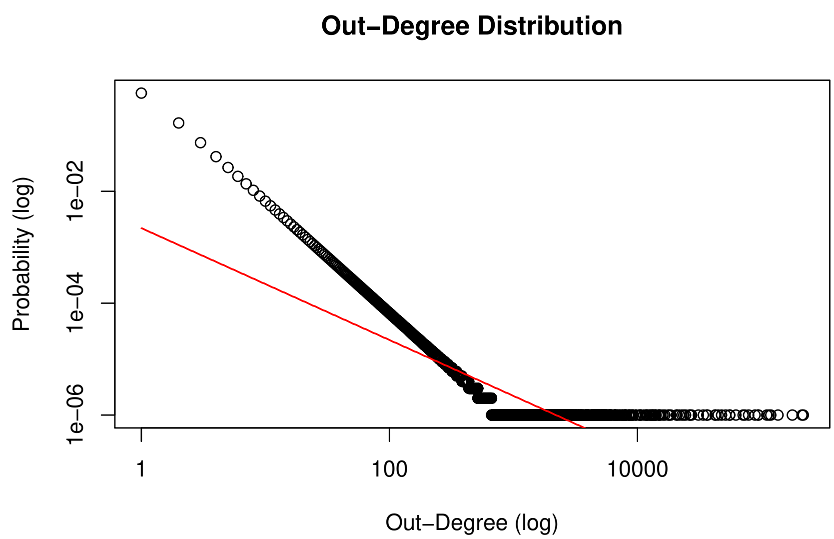


Figura B.4: Distribución de out-degre, en escala logarítmica, de un grafo no dirigido de un millón de nodos generados con el Algoritmo 1.

B.1.2. Algoritmo 2

Gráficos de distribución de grados para grafos dirigidos y no dirigidos de 1 000 y 1 000 000 de nodos generados por el Algoritmo 2.

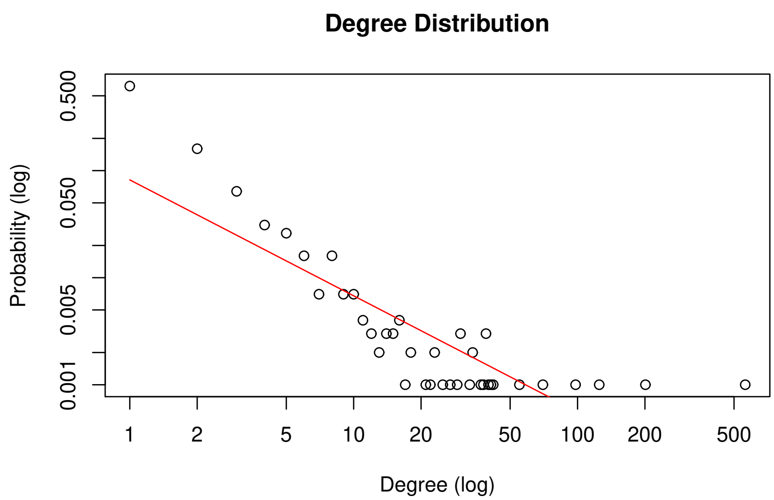


Figura B.5: Distribución de grados, en escala logarítmica, de un grafo no dirigido de mil nodos generados con el Algoritmo 2.

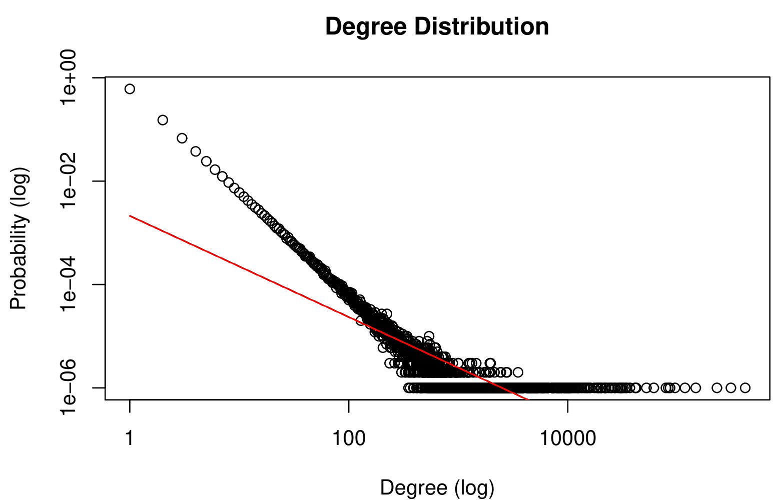


Figura B.6: Distribución de grados, en escala logarítmica, de un grafo no dirigido de un millón de nodos generados con el Algoritmo 2.

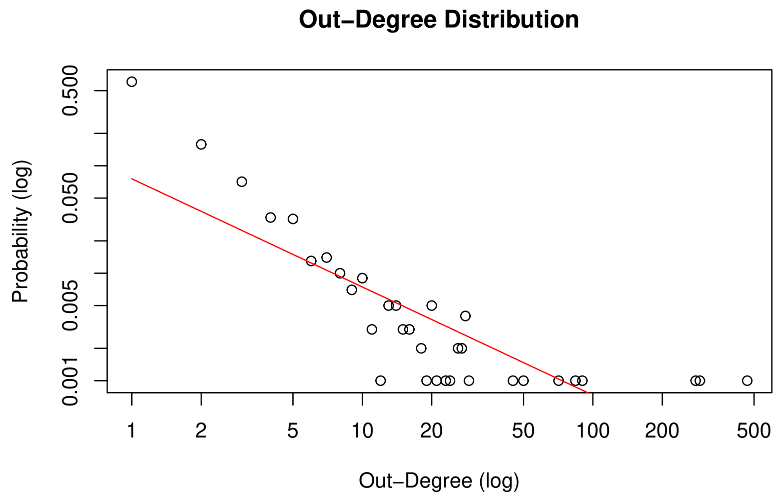


Figura B.7: Distribución de out-degree, en escala logarítmica, de un grafo dirigido de mil nodos generados con el Algoritmo 2.

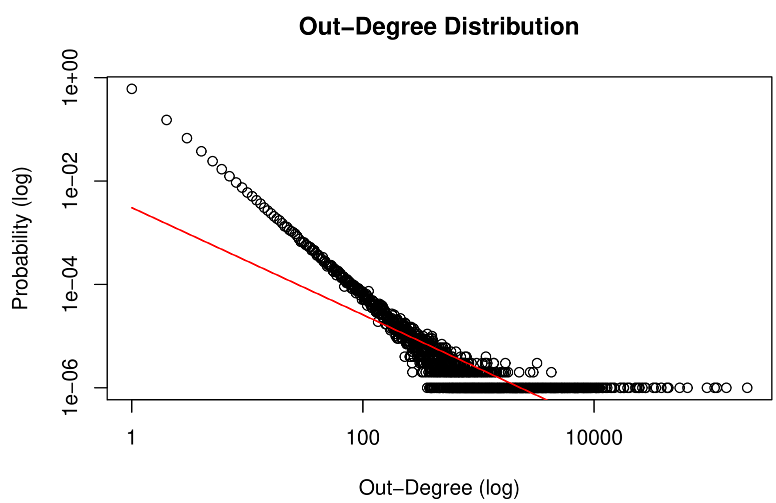


Figura B.8: Distribución de out-degre, en escala logarítmica, de un grafo no dirigido de un millón de nodos generados con el Algoritmo 2.

B.1.3. Algoritmo 3

Gráficos de distribución de grados para grafos dirigidos y no dirigidos de 1 000 y 1 000 000 de nodos generados por el Algoritmo 3.

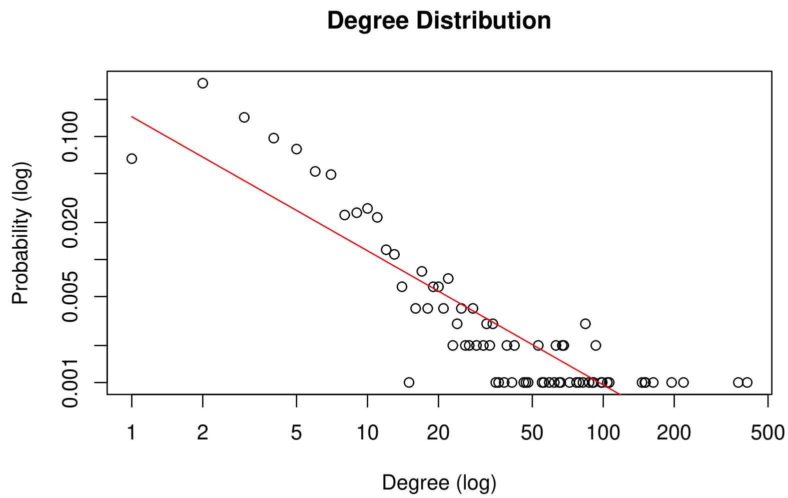


Figura B.9: Distribución de grados, en escala logarítmica, de un grafo no dirigido de mil nodos generados con el Algoritmo 3.

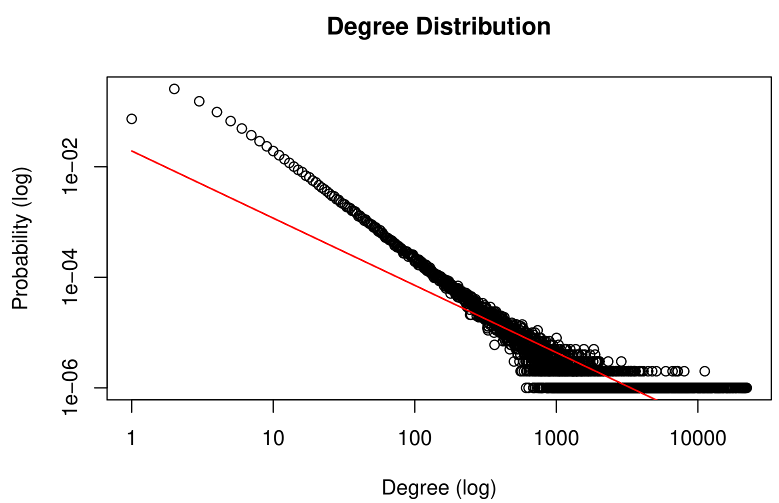


Figura B.10: Distribución de grados, en escala logarítmica, de un grafos no dirigido de un millón de nodos generados con el Algoritmo 3.

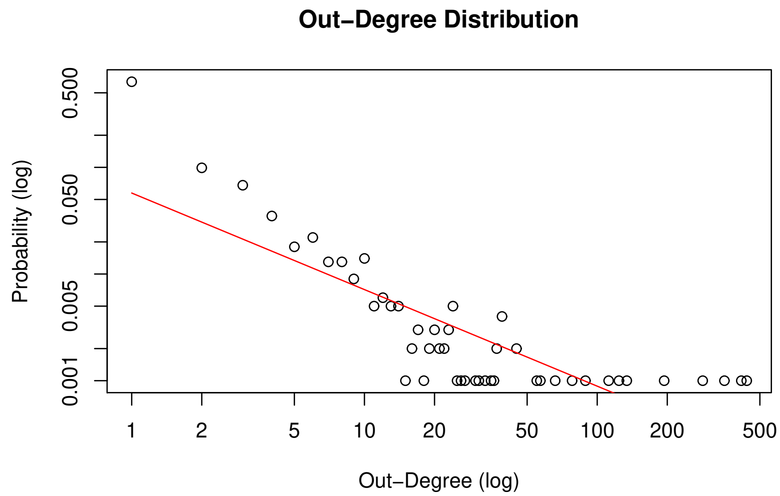


Figura B.11: Distribución de out-degree, en escala logarítmica, de un grafo dirigido de mil nodos generados con el Algoritmo 3.

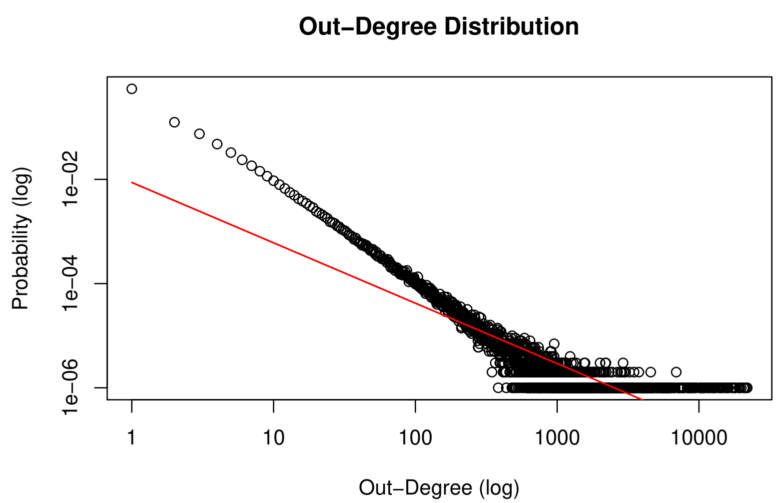
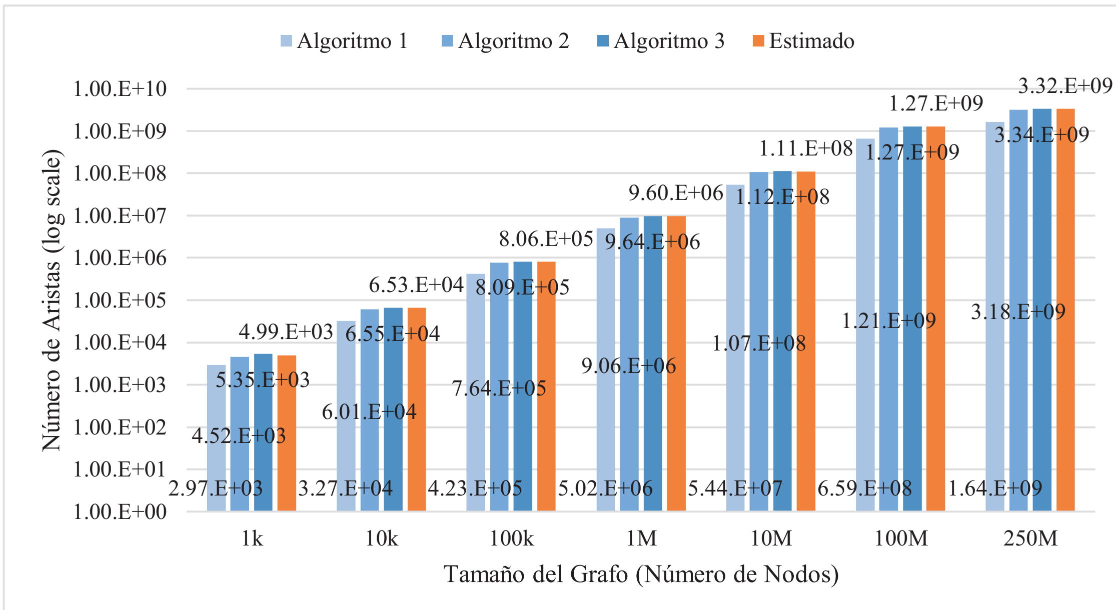


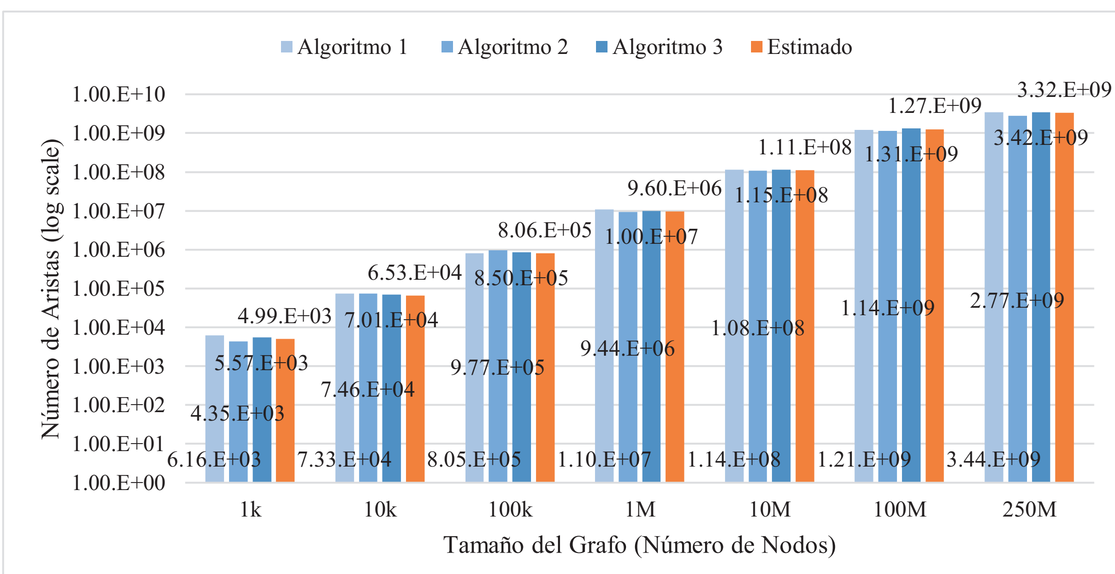
Figura B.12: Distribución de out-degre, en escala logarítmica, de un grafo no dirigido de un millón de nodos generados con el Algoritmo 3.

B.2. Eficiencia

Gráficos con la comparación entre los Algoritmos 1, 2 y 3 en la generación de aristas para grafos dirigidos y



(a) Grafos no dirigidos



(b) Grafos dirigidos

Figura B.13: Número de aristas generado por cada método por tamaño de grafos.