

UNIVERSIDAD MAYOR DE SAN ANDRÉS
FACULTAD DE CIENCIAS PURAS Y NATURALES
CARRERA DE INFORMÁTICA



PERFIL DE TESIS DE GRADO

MODELO GENERADOR DE GRAFOS ALEATORIOS

Para obtener el Título de Licenciatura en Informática
Mención Ingeniería de Sistemas Informáticos

POR: CARLOS MIJAEL TOLA APAZA

TUTOR: LIC. JORGE TERAN

LA PAZ - BOLIVIA

Noviembre, 2021

Índice

1. Introduccion	3
2. Antecedentes	3
3. Planteamiento del problema	4
3.1. Problema central	5
4. Hipótesis	5
4.1. Operacionalización de variables	5
5. Definicion de objetivos	5
5.1. Objetivo general	5
5.2. Objetivos secundarios	5
6. Justificación	6
6.1. Justificación económica	6
6.2. Justificación social	6
6.3. Justificación científica	6
7. Alcances y limites	6
7.1. Alcances	6
7.2. Límites	6
8. Metodología	6
9. Bibliografía	7
10. Anexos	9
10.1. Cronograma de avance	9
10.2. Indice de trabajo	9

1. INTRODUCCION

Una red o grafo es un sistema compuesto por objetos discretos (vértices) y una relación (arista) entre algunos pares de ellos. Son herramientas de modelado ubicuas en muchos campos de investigación. Algunos ejemplos comunes incluyen redes sociales, neuronales, de proteínas y de comunicación. Hoy en día con el poder de computo disponible, han hecho que el estudio de redes y herramientas para analizarlas sean temas en alza dentro de la comunidad académica.

Los grafos aleatorios son construcciones probabilísticas que permiten definir realizaciones aleatorias de grafos. El uso generalizado de grafos aleatorios se ha destacado en el contexto de las aplicaciones de bases de datos y análisis de algoritmos durante varios años. Esto se debe a que tales estructuras de datos resultan ser muy útiles en muchas aplicaciones, desde el análisis de redes complejas hasta el estudio de algoritmos aleatorios. Generar estos grafos de manera aleatoria tiene un nivel de dificultad bastante alta, ya que al momento de generar un grafo se tienen que controlar algunas propiedades dependiendo el tipo de grafo que deseamos generar, por ejemplo: si el objetivo es construir un Arbol (Grafo conexo con E vértices y $E - 1$ aristas) se debe tener control sobre donde se genera un nuevo nodo y arista, verificando que no existan ciclos, caso contrario el modelo no cumpliría con su objetivo. También podríamos mencionar que construir un Grafo Completo (Grafo con E vértices y $\frac{E*(E-1)}{2}$ aristas) tiene una complejidad No Polinomial NP, haciendo que sea solo posible generar grafos completos con un límite de 20 nodos aproximadamente. Otro factor importante es que al ser este un modelado de grafos aleatorios, cada vértice a ser agregado, tiene cierta probabilidad de ser tomado en cuenta en el grafo generado o no.

Existen varios algoritmos generadores de grafos, como ser los modelos de grafos de Erdős-Rényi, y algunas variantes de dicho modelado, que no generan grafos representativos (no se asemejan a redes del mundo real), no pueden generar grafos con grandes cantidades de vértices y se toman bastante tiempo en el generado [B2]. Dados estos argumentos, el objetivo de este trabajo es crear un modelo generador de grafos aleatorios que sean representativos de forma eficiente. Cuando decimos representativos, nos referimos a Redes que se asemejen al mundo real y cuando decimos eficiente, nos referimos a: mejores tiempos de ejecución para el generado y el mínimo uso de CPU, GPU y memoria comparando a modelos ya existentes.

Debido a que en Teoría de Grafos existe una infinidad de tipos de grafos, nos enfocaremos simplemente en un conjunto de grafos que son: Grafos directos acíclicos (DAG), Grafos Bipartitos, Grafos completos, Grafos no dirigidos conexos y no conexos, y árboles.

2. ANTECEDENTES

El primer modelo de grafos aleatorios fue introducidos por P. Erdős en 1959 como una herramienta para aplicar el método probabilístico para probar una relación asintótica entre la cintura y el número cromático de grafos generales. Más tarde, junto con A. Rényi, continuó el estudio de este modelo. Este es uno de los dos modelos conocidos hoy en día como modelos de Erdős-Rényi. Consiste en la construcción de grafos finitos con un número fijo de aristas distribuidas uniformemente a lo largo del conjunto de pares de vértices; mientras que el segundo modelo que comparte el mismo nombre, se debe a E. Gilbert, y consiste en grafos finitos obtenidos al establecer, independientemente y con la misma probabilidad, una arista entre cada par de vértices. A partir de ahí, se desarrollaron muchos modelos de grafos aleatorios diferentes. Estos probaron ser no sólo construcciones interesantes, sino también muy útiles en aplicaciones: como herramientas para estudiar escenarios promedio en

redes complejas y como una forma de explorar posibles mecanismos que dan lugar a muchas de las características distintivas de las redes observadas en el mundo real. Por ejemplo, incluso un modelo tan simple como el de Erdős-Rényi exhibe la propiedad del mundo pequeño (prevalencia de distancias relativamente cortas entre la mayoría de los pares de vértices) observada en muchas redes sociales y de comunicación.

Al respecto de las investigaciones relacionadas con el tema podemos destacar las siguientes:

- **Titulo:** Desempeño asintótico de algoritmos secuenciales en grafos aleatorios - Tesis presentada para optar al título de Doctor.

Autor: Lic. Manuel Sáenz

Año: 2019

Institucion: Universidad de Buenos Aires

En el mencionado trabajo, se estudia la evolucion y convergencia de algoritmos generadores de grafos, asi de esa forma lograr delimitar el algoritmo de modelado de grafos Erdős-Rényi. Para demostrar sus limitaciones se demuestra que para los grafos generados por el modelo Erdős-Rényi es posible construir un conjunto independiente maximo en tiempo polinomial. Pese a que el problema de construir Conjuntos Independientes Maximos es del tipo NP-difícil.

- **Titulo:** Generación rápida de gráficos aleatorios

Autor: Sadegh Heyrani, Nobari Xuesong Lu

Año: 2011

Institucion: Universidad Nacional de Singapore

En el mencionado trabajo, se plante un modelo generador de grafos, usando programacion paralela y gpu junto a su algoritmo ER y ZER (creado por ellos). Los grafos que pueden ser generados por este modelado son: Grafos dirigidos y no dirigidos.

- **Titulo:** Implementacion de un generador de topologias aleatorias en emulador de red mininet.

Autor: Rosa Alejandra García Grisales, Juan Carlos Agudelo Calderón

Año: 2015

Institucion: Universidad Católica De Pereira

En el mencionado trabajo, se plante implementar un generador de grafos para el emulador Mininet, que es una aplicacion desarrollada para la investigacion de prototipos SDN en tiempo real. Este método usa un Modelo de Waxmax (variante de Erdős-Rényi), que construye grafos en tiempo lineal, pero usando bastante memoria limitando el generado a aproximadamente 1000 nodos solamente.

- **Titulo:** Modelado de grafos aleatorios: una revisión de los conceptos

Autor: Mikhail Drobyshevskiy, Denis Turdakov

Año: 2019

Institucion: Instituto de Física y Tecnología de Moscú

En este trabajo, se hace una analisis a profundidad de el Algoritmo Erdős-Rényi. Se indentifico seis direcciones donde este modelos tienen sus aplicaciones: comprensión de redes, análisis, extrapolación, puntos de referencia, modelos nulos y aleatorización.

3. PLANTEAMIENTO DEL PROBLEMA

El modelado de grafos juega un rol importante en el análisis de redes complejas. Nos ayudan a entender, controlar predecir fenómenos que pueden ocurrir en redes sociales, redes biológicas, internet y muchos mas. Aunque ya existen, varios modelos generadores de grafos, tales como el Modelado

Erdős-Rényi, esto no suelen ser representativos y al mismo tiempo son costosos en tiempo y memoria [B2].

Otros autores han afirmado que:

En la era del BIG DATA, hay muchas redes masivas que deben explotarse y analizarse. Ya que tales redes no se pueden manejar en la memoria de una sola computadora, nuevos métodos de agrupamiento se han introducido para modelos avanzados de computación (Buzun et al., 2014; Zeng & Yu, 2016). Estos algoritmos utilizan representaciones de entrada jerárquicas, lo que implica que los experimentos realizados en Grafos de referencia de tamaño pequeño o mediano no se pueden utilizar para predecir el rendimiento en instancias mucho más grandes (Emmons et al., 2016). Desafortunadamente, muchos puntos de referencia de agrupamiento de gráficos Los generadores actualmente disponibles no pueden generar los Grafos del tamaño necesario (Bae & Howe, 2015; Buzun et al., 2014). (*Artificial Benchmark for Community Detection (ABCD)—Fast random graph model with community structure, 2021*)

3.1. PROBLEMA CENTRAL

¿Con que grado de fiabilidad los modelos generadores de grafos aleatorios Erdős-Rényi pueden construir grafos representativos en tiempo y coste de memoria minimos?

4. HIPÓTESIS

El modelo para la generacion de grafos aleatorios usando GPU (Unidad de procesamiento gráfico) y programación paralela representa estos Grafos en tiempo y memoria eficiente con una fiabilidad del 90 %.

4.1. OPERACIONALIZACIÓN DE VARIABLES

Se pudieron identificar las siguientes variables:

Concepto	Dimension	Indicador	Escala de medicion
Eficiencia	Tiempo de ejecucion		Segundos
	Memoria usada en el proceso del generado		Megabytes
Representacion eficaz	Grado		Numero Entero
	Numero de componentes		Numero Entero
Aleatoriedad	Entropia		Numero Real

5. DEFINICION DE OBJETIVOS

5.1. OBJETIVO GENERAL

Desarrollar un modelo generador de grafos aleatorios representativos y eficiente en tiempo y memoria.

5.2. OBJETIVOS SECUNDARIOS

- Analizar y diseñar un modelo generador de grafos aleatorios para: Grafos directos acíclicos (DAG), Grafos Bipartitos, Grafos completos, Grafos no dirigidos conexos y no conexos, y árboles.

- Implementar el modelo generador de grafos.
- Verificar que los tiempos al generar los grafos sean eficientes experimentalmente.
- Verificar que el generador de grafos sea eficiente según su complejidad algorítmica.

6. JUSTIFICACIÓN

6.1. JUSTIFICACIÓN ECONÓMICA

Para el analisis de diferentes redes, se requiere tener acceso a distintas redes de prueba, para su simulacion y estudio. Generar estas redes o grafos de manera eficiente significa que se reduzcan los costos de energia y se eliminara el tiempo invertido en dicho modelado.

6.2. JUSTIFICACIÓN SOCIAL

El impacto que tienen las redes sociales hoy en día en la sociedad ha marcado un hito en la forma de comunicarnos, debido a su crecimiento actual estas redes sociales siempre esta a la mejora de sus algoritmos. Para esto es necesario investigar y probar dichos algoritmos en grafos aleatorios.

6.3. JUSTIFICACIÓN CIENTÍFICA

El presente trabajo combinara tecnicas algorítmicas, mascarar de bits, con tecnicas matematicas, Hashing, de tal manera que el modelado de las redes o grafos sea eficiente, esto se puede verificar usando notaciones de complejidad algorítmicas.

7. ALCANCES Y LIMITES

7.1. ALCANCES

- El presente trabajo diseñara un modelo generador de grafos aleatorios para DAGs, Grafos bipartitos, Grafos completos, Grafos no dirigidos conexos y no conexos, y arboles de manera eficiente.

7.2. LÍMITES

- Algunos tipos de grafos con construcciones y solucion no polinomial NP no seran tomados en cuenta.

8. METODOLOGÍA

Para el desarrollo del presente trabajo se utilizara un **Método de investigacion experimental** para cumplir con los objetivos propuestos para el modelo generador de grafos aleatorios. Esta se dividiran en las siguientes etapas.

1. Recopilacion de informacion.
2. Elaboracion del Modelo generador de grafos aleatorios.
3. Pruebas de funcionamiento y/o presentación de mejoras para el modelo.
4. Comparaciones y analisis de los datos.

5. Conclusiones.

La primera etapa se recopilara la informacion necesaria y se estudiara temas relacionadas al mismo.

La segunda etapa se realizara el Modelo generado de grafos tomando encuesta los alcances y limites.

La tercera etapa se verificara la funcionalidad del modelo.

En la cuarta etapa se analizaran los datos y se los comparara con otros modelos.

En la ultima etapa, se presentaran los resultados finales, las conclusiones y recomendaciones con respecto al proyecto

ETAPAS	ACTIVIDADES
Recopilacion de informacion	<ul style="list-style-type: none">- Investigar la existenciaa de modelos generadores de grafos.- Estudiar el proceso de generado de dichos modelos investigados.
Elaboracion del modelo	<ul style="list-style-type: none">- Analizar y diseñar las etapas que necesitaran los modelados.- Modelar el proceso generico, para construir los grafos de forma aleatoria (Dirigidos y no Dirigidos).- Modelaro el proceso para los distintos grafos descritos en los alcances (DAG, Completos y Arboles).
Pruebas de funcionamiento y mejoras.	<ul style="list-style-type: none">- Verificar el funcionamiento pleno del modelo.- Aplicar las mejoras encontradas y propuestas.
Comparacion y Analisis	<ul style="list-style-type: none">- Comparar el modelado propuesto con respector al modelos ya existentes.- Analizar de formar estadisticaa los datos devuelto de las comparaciones.- Generalizar y encontrar la complejidad Algoritmica.
Conclusiones	<ul style="list-style-type: none">- Presentar las conclusions obtenias al finalizar el trabajo.- Compartir algunas recomendaciones obtenidas del trabajo realizado.

9. BIBLIOGRAFÍA

[B1] Manuel Zaens. (2019). Desempeño asintotico de algoritmos secuenciales en grafos aleatorios. 2019, de Universidad de Buenos Aires Sitio web:
[http://cms.dm.uba.ar/academico/carreras/doctorado/Tesis_Saenz_corregida\(1\).pdf](http://cms.dm.uba.ar/academico/carreras/doctorado/Tesis_Saenz_corregida(1).pdf)

[B2] Sadegh Heyrani, Nobari Xuesong Lu. (2011). Generación rápida de gráficos aleatorios. 2019, de Universidad Nacional de Singapore Sitio web:
<https://www.cs.au.dk/~karras/rg.pdf>

[B3] Rosa Alejandra García Grisales, Juan Carlos Agudelo Calderón. (2015). Implementacion de un generador de topologias aleatorias en emulador de red mininet. 2019, de Universidad Católica De Pereira Sitio web:
<https://repositorio.ucp.edu.co/bitstream/10785/3678/1/CDMIST122.pdf>

[B4] Mikhail Drobyshevskiy, Denis Turdakov. (2019). Modelado de grafos aleatorios: una revisión de los conceptos. 2019, de Mikhail Drobyshevskiy, Denis Turdakov Sitio web:
<https://dl.acm.org/doi/pdf/10.1145/3369782>

[B5] Bogumił Kaminski, Paweł Prałat², (2021). Artificial Benchmark for Community Detection(ABCD)—Fast random graph model with community structure. 2021 de Universidad de Cambridge web:
<https://www.cambridge.org/core/journals/network-science/article/artificial-benchmark-for-community-detection-abcdfast-random-graph-model-with-community-structure/453FE29A1FA3C1798B0EC116587FE422>

10. ANEXOS

10.1. CRONOGRAMA DE AVANCE

ACTIVIDADES	DURACION (DIAS)	DEL 1 DE FEBRERO A 30 DE JUNIO																							
		Diciembre				Enero				Febrero				Marzo				Abril				Mayo			
		1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
Redacción del Capitulo II Marco Teórico	14																								
Desarrollo del Capítulo III Marco Aplicativo	21																								
Modelo generador de grafos DAG	28																								
Modelo generador de grafos Bipartitos	35																								
Modelo generador de grafos completos	28																								
Algoritmo validador del modelo	21																								
Redacción del Capitulo III Marco Aplicativo	21																								
Redacción del Capitulo IV Estado de la Hipótesis	21																								
Redacción del Capitulo V Conclusiones y Recomendaciones	14																								

10.2. INDICE DE TRABAJO

1. CAPÍTULO 1: MARCO REFERENCIAL

- 1.1 Introduccion
- 1.2 Antecedentes
- 1.3 Planteamiento del problema
 - 1.3.1 Problema central
 - 1.3.2 Problemas secundarios
- 1.4 Objetivo
 - 1.4.1 Objetivo principal

- 1.4.2 objetivos secundarios
- 1.5 Hipótesis
- 1.6 Justificación
 - 1.6.1 Justificación económica
 - 1.6.2 Justificación social
 - 1.6.3 Justificación científica
- 1.7 Alcances y Límites
 - 1.7.1 Alcances
 - 1.7.2 Límites
- 1.8 Metodología
- 2. CAPITULO 2: MARCO TEORICO
 - 2.1 Mascaras de bits
 - 2.1.1 Operaciones con mascarar de bits
 - 2.2 Programación dinámica (DP)
 - 2.2.1 DP con mascarar de bits
 - 2.3 Teoría de grafos
 - 2.3.1 Grafos acíclicos no diridos (DAG)
 - 2.3.2 Grafos Bipartitos
 - 2.3.3 Grafos completos
 - 2.3.4 Caminos y ciclos hamiltonianos
 - 2.4 Grafos aleatorios
 - 2.4.1 Modelo de Erdős-Rényi
 - 2.4.2 Modelo de Configuraciones
 - 2.4.3 Otros modelos de grafos aleatorios
 - 2.5 Cadenas de Markov
- 3. CAPITULO 3: MARCO APLICATIVO
 - 3.1 Modelo generador de grafos DAG
 - 3.2 Modelo generador de grafos Bipartitos
 - 3.3 Modelo generador de grafos completos
 - 3.4 Algoritmo validdor del modelo
- 4. CAPITULO 4: ANÁLISIS Y RESULTADOS
 - 4.1 Complejidad algoritmica.
 - 4.2 Demostracion de la mejora con respecto a otros modelos.
- 5. CAPITULO 5:
 - 5.1 Conclusiones
 - 5.2 Recomendaciones
- 6. Bibliografía
- 7. Anexos