

TABLA DE CONTENIDOS

	página
Dedicatoria	I
Agradecimientos	II
Tabla de Contenidos	III
Índice de Figuras	VII
Índice de Tablas	IX
Resumen	XII
1. Introducción	13
1.1. Contexto	13
1.2. Problema	14
1.3. Hipótesis	15
1.4. Objetivos	15
1.4.1. Objetivo General	16
1.4.2. Objetivos Específicos	16
1.5. Propuesta de Solución	16
1.6. Alcances	18
1.7. Estructura del Documento	18
2. Antecedentes	20
2.1. Funciones de Probabilidad	20
2.1.1. Función de Densidad de Probabilidad	20
2.1.2. Ley de Potencia	21
2.1.3. Números Aleatorios con Distribución de la Ley de Potencia	21
2.2. Grafos	22
2.2.1. Sobre la adyacencia de los nodos	23
2.2.2. Aridad de un Nodo	23
2.3. Representación de Grafos	24

2.3.1.	Lista de Aristas	24
2.3.2.	Matriz de Adyacencia	24
2.3.3.	Arreglo de Aridades	26
2.4.	Algoritmo R-MAT	26
2.4.1.	Estadística de Kolmogorov-Smirnov (KS)	28
2.5.	Trabajo Relacionado	28
2.6.	Tecnologías Computacionales	30
3.	Metodología	32
3.1.	Fase Conceptual	32
3.2.	Fase de Planeación y Diseño de Investigación	33
3.3.	Fase Empírica y Analítica	33
3.3.1.	Método de Evaluación Experimentos Preliminares	34
3.3.2.	Configuración Experimental para Experimentos Preliminares	34
3.4.	Fase de Difusión	34
4.	Generación de Grafos acorde a la Ley de Potencia	35
4.1.	Algoritmo 1	35
4.1.1.	Generación de la Distribución de Aridades	37
4.1.2.	Experimentos Preliminares Algoritmo 1	38
4.2.	Algoritmo 2	41
4.2.1.	Generación de Probabilidades Acumuladas	41
4.2.2.	Generación de Arreglo de Frecuencias	43
4.2.3.	Generación del Arreglo de Aridades	45
4.2.4.	Experimentos Preliminares Algoritmo 2	47
4.3.	Distribución de Números Aleatorios	49
4.4.	Modelado de la Aridad Máxima	54
4.4.1.	¿Por qué Modelar una Función Para Obtener la Aridad máxima?	55
4.4.2.	Experimento para Modelar la Función de Aridad Máxima	57
4.4.3.	Función que Determina la Constante de Proporcionalidad	62
4.4.4.	Función que Determina el Exponente de Potencia	63
4.4.5.	Función que Determina la Aridad Máxima	65
4.5.	Algoritmo 3	67
4.5.1.	Generación de Probabilidades Acumuladas	67

4.5.2.	Generación del Arreglo de Frecuencias	69
4.5.3.	Generación del Arreglo de Aridades	72
4.5.4.	Experimentos Preliminares Algoritmo 3	74
5.	Evaluación Experimental	77
5.1.	Fase Experimental 1	77
5.1.1.	Configuración Experimental	78
5.1.2.	Resultados Experimentales Fase 1	78
5.2.	Fase Experimental 2	86
5.2.1.	Resultados Experimentales Fase 2	86
5.2.2.	Comparación de R ³ MAT y Algoritmo 3	87
6.	Conclusiones	90
	Bibliografía	94
	Anexos	
A:	Experimentos Preliminares	98
A.1.	Experimentos Preliminares Algoritmo 1	98
A.2.	Experimento para la Aridad Máxima	109
A.2.1.	Aridades Máximas Clasificadas por Tamaño del Grafo.	109
A.2.2.	Aridades máximas promedio registrados por tamaño del grafo y por τ	133
A.2.3.	Registro de la Constate de Proporcionalidad y el Exponente de la potencia.	136
A.2.4.	Registro de las Aridades Máximas Obtenidas en la Simulación	139
A.2.5.	Registro de la Diferencia entre la Simulación y los Valores Empíricos de la Aridad Máxima.	142
A.2.6.	Registro de la Diferencia entre la Simulación y los Valores Empíricos de la Aridad Máxima en Porcentajes.	144
B:	Evaluación Experimental	148
B.1.	Realismo	148
B.1.1.	Algoritmo 1	148
B.1.2.	Algoritmo 2	151

B.1.3. Algoritmo 3	153
B.2. Eficiencia	155

ÍNDICE DE FIGURAS

	página
2.1. Grafo dirigido.	23
2.2. Grafo no dirigido.	25
2.3. Ejemplo de matriz de adyacencia generada por R-MAT.	28
4.1. Distribución de la aridad en escala logarítmica de base 10 para un grafo de 100 000 nodos.	49
4.2. Mil números aleatorios generados con <code>Math.random()</code>	51
4.3. Los mil números generados clasificados por aridad.	52
4.4. Mil números aleatorios generados con distribución de ley de potencia.	54
4.5. Mil números generados con la Ecuación 4.13 clasificados por aridad.	55
4.6. Distribución de la aridad de un grafo de 100 000 nodos Algoritmo 1.	56
4.7. Distribución de la aridad de un grafo de 100 000 nodos Algoritmo 2.	56
4.8. Distribución de la aridad de un grafo de 100 000 nodos R ³ MAT.	57
4.9. Gráfico de dispersión para la aridad máxima versus la cantidad de nodos.	60
4.10. Gráfico de dispersión para la constante de proporcionalidad versus τ	61
4.11. Gráfico de dispersión para el exponente de la potencia versus τ	61
4.12. Gráfico de una función exponencial cuyo exponente negativo es un logaritmo natural al cuadrado.	62
4.13. Gráfico de dispersión para la constante de proporcionalidad versus τ	63
4.14. Gráfico de una función exponencial con exponente negativo.	64
4.15. Gráfico de dispersión para del exponente de la potencia versus τ	64
4.16. Distribución de la aridad de un grafo de 100 000 nodos Algoritmo 3.	76
5.1. Tiempo de ejecución de los métodos para generar grafos de diferentes tamaños.	82
5.2. Distribución de Out Degree de grafos de un millón de nodos con PaR-MAT (a), SNAP (b), R ³ MAT _a (c) y Algoritmo 3 (d).	88
B.1. Distribución de grados, en escala logarítmica, de un grafo no dirigido de mil nodos generados con el Algoritmo 1.	148

B.2. Distribución de grados, en escala logarítmica, de un grafo no dirigido de un millón de nodos generados con el Algoritmo 1.	149
B.3. Distribución de out-degree, en escala logarítmica, de un grafo dirigido de mil nodos generados con el Algoritmo 1.	149
B.4. Distribución de out-degree, en escala logarítmica, de un grafo no dirigido de un millón de nodos generados con el Algoritmo 1.	150
B.5. Distribución de grados, en escala logarítmica, de un grafo no dirigido de mil nodos generados con el Algoritmo 2.	151
B.6. Distribución de grados, en escala logarítmica, de un grafo no dirigido de un millón de nodos generados con el Algoritmo 2.	151
B.7. Distribución de out-degree, en escala logarítmica, de un grafo dirigido de mil nodos generados con el Algoritmo 2.	152
B.8. Distribución de out-degree, en escala logarítmica, de un grafo no dirigido de un millón de nodos generados con el Algoritmo 2.	152
B.9. Distribución de grados, en escala logarítmica, de un grafo no dirigido de mil nodos generados con el Algoritmo 3.	153
B.10. Distribución de grados, en escala logarítmica, de un grafos no dirigido de un millón de nodos generados con el Algoritmo 3.	153
B.11. Distribución de out-degree, en escala logarítmica, de un grafo dirigido de mil nodos generados con el Algoritmo 3.	154
B.12. Distribución de out-degree, en escala logarítmica, de un grafo no dirigido de un millón de nodos generados con el Algoritmo 3.	154
B.13. Número de aristas generado por cada método por tamaño de grafos. .	155

ÍNDICE DE TABLAS

	página
1.1. Distribución de aridades.	17
1.2. Arreglo de aridades.	17
2.1. Aridad del grafo dirigido de la Figura 2.1.	24
2.2. Matriz de adyacencia.	25
2.3. Arreglo de aridades para el grafo de la Figura 2.1.	26
2.4. Arreglo de aridades para el grafo de la Figura 2.2.	26
4.1. Arreglo de aridades Algoritmo 1.	38
4.2. Grafos generados en la evaluación preliminar del Algoritmo 1.	40
4.3. Cuadro resumen de las probabilidades generadas para cada aridad posible en un grafo de 1 000 nodos.	41
4.4. Arreglo de probabilidades Algoritmo 2.	42
4.5. Arreglo de probabilidades acumuladas Algoritmo 2.	43
4.6. Arreglo de aridades Algoritmo 2.	47
4.7. Grafos generados en la evaluación preliminar del Algoritmo 2.	47
4.8. Cantidad de nodos con aridad 1 para los diferentes grafos generados.	48
4.9. Registro de los mil números aleatorios obtenidos con Math.random().	50
4.10. Registro de los mil números aleatorios obtenidos con distribución de la ley de potencia.	53
4.11. Aridades máximas registradas para un grafo de 1 000 nodos.	58
4.12. Aridades máximas registradas para un grafo de 10 000 000 nodos.	58
4.13. Aridad máxima promedio dada una cantidad de nodos y un valor de τ para un grafo.	59
4.14. Constante de proporcionalidad y exponente de potencia por cada valor de τ	60
4.15. Aridades máximas obtenidas de la fórmula de la Ecuación 4.16.	65
4.16. Diferencia entre la aridad registrada en el Cuadro 4.15 y el Cuadro 4.13.	66
4.17. Diferencia entre la aridad registrada en el Cuadro 4.15 y el Cuadro 4.13 en porcentaje.	66
4.18. Arreglo de probabilidades Algoritmo 3.	69

4.19. Arreglo de probabilidades acumuladas Algoritmo 3. 69

4.20. Arreglo de aridades Algoritmo 3. 74

4.21. Grafos generados en la evaluación preliminar del Algoritmo 3. 74

4.22. Aridad máxima de los diferentes grafos generados. 75

5.1. Número de aristas estimadas para $\tau = 2$ 79

5.2. Grafos generados en la evaluación experimental con el Algoritmo 1. . . 79

5.3. Grafos generados en la evaluación experimental con el Algoritmo 2. . . 80

5.4. Grafos generados en la evaluación experimental con el Algoritmo 3. . . 80

5.5. Resultados de los tiempos de ejecución de cada algoritmo. 81

5.6. Número de aristas generada por los métodos para generar grafos no dirigidos de diferentes tamaños. 83

5.7. Número de aristas generada por los métodos para generar grafos dirigidos de diferentes tamaños. 83

5.8. Puntuaciones KS-stat para grafos no dirigidos. 84

5.9. Puntuaciones KS-stat para grafos dirigidos all-degree. 85

5.10. Puntuaciones KS-stat para grafos dirigidos in-degree. 85

5.11. Puntuaciones KS-stat para grafos dirigidos out-degree. 85

5.12. Comparación del Algoritmo 3 con PaRMAT, SNAP y R³MAT. 87

5.13. Registro del tiempo de generación del arreglo de aridades en R³MAT y el Algoritmo 3. 89

A.1. Probabilidades para las aridades de un grafo de 1 000 nodos. 98

A.2. Aridades máximas registradas para un grafo de 1 000 nodos. 109

A.3. Aridades máximas registradas para un grafo de 10 000 nodos Parte 1. 111

A.4. Aridades máximas registradas para un grafo de 10 000 nodos Parte 2. 114

A.5. Aridades máximas registradas para un grafo de 100 000 nodos Parte 1. 117

A.6. Aridades máximas registradas para un grafo de 100 000 nodos Parte 2. 119

A.7. Aridades máximas registradas para un grafo de 1 000 000 nodos Parte 1.123

A.8. Aridades máximas registradas para un grafo de 1 000 000 nodos Parte 2.125

A.9. Aridades máximas registradas para un grafo de 10 000 000 nodos Parte 1. 128

A.10. Aridades máximas registradas para un grafo de 10 000 000 nodos Parte 2. 131

A.11. Aridades máximas promedio registrados por tamaño del grafo y por τ . 133

A.12. Constante de proporcionalidad y exponente de potencia por cada valor de τ	136
A.13. Aridades máximas obtenidas de la fórmula de la aridad máxima. . . .	139
A.14. Diferencia entre la aridad registrada por la simulación y los datos empíricos.	142
A.15. Diferencia entre la aridad registrada por la simulación y los datos empíricos.	144