
GENERACIÓN DE GRAFOS INDUCIDOS POR LA LEY DE POTENCIA

DIEGO ALEJANDRO ITURRIAGA PEÑA
INGENIERO CIVIL EN COMPUTACIÓN

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 R3MAT, 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, $P(x = k) = C \times k^{-t}$ 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 probó 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.