

Título del Proyecto: Análisis-Algoritmos y Visualización del Grafo de la Red 'X'

Objetivo

El objetivo principal de este proyecto es analizar y visualizar la estructura de un gran grafo de red social para descubrir patrones interesantes y obtener información sobre la conectividad social, las estructuras de la comunidad y las propiedades de la red.

Conjunto de Datos

- Nombre: Red Social 'X'
- Fuente: <https://drive.google.com/drive/folders/1XvzgZ3Nko3EruGOHDirM6bQwfc8fejpl?usp=sharing>
 - Descripción: Este conjunto de datos es un subconjunto de todos los usuarios de la red social, junto con sus conexiones y ubicaciones. El grafo creado considera a los usuarios como nodos y sus conexiones como aristas, y estos forman un componente conectado del grafo total. Recuerden que la relación seguidor-seguido no es mutua. "X sigue a Y" no implica "Y sigue a X".
 - Se tiene dos archivos .txt con 10 millones de usuarios: **ubicación** y **usuario**.
 - El archivo de **ubicación** ("10_million_location.txt.zip") contiene la latitud y longitud de cada usuario. El formato del archivo es:
 - lat_1, long_1
 - lat_2, long_2
 - lat_3, long_3
 - donde lat_1 y long_1 son la latitud y longitud del usuario número 1 respectivamente, y así sucesivamente.
 - El archivo de **usuario** ("10_million_user.txt.zip") contiene la lista de adyacencia de cada usuario, donde la i-ésima fila representa la lista de adyacencia del i-ésimo usuario. La indexación comienza en 1. El formato del archivo es:
 - 2, 3, 4, 5, 6, . . .
 - 1, 27188, 50986, . . .
 - 1, 34590, 13474, . . .

Esquema del Proyecto

1. Introducción

- Motivación: Explicar por qué es importante entender las estructuras de las redes sociales.
- Objetivos: Definir los objetivos y resultados esperados del proyecto.

3. Preprocesamiento de Datos

- Limpieza de Datos: Manejar cualquier dato faltante o inconsistente.

Construcción del Grafo:

Construir el grafo usando cualquier lenguaje de programación(no herramienta, se revisará el código y se hará preguntas directas), donde los nodos representan usuarios y las aristas representan amistades.

4. Análisis Exploratorio de Datos (EDA)

- Estadísticas Básicas: Calcular estadísticas básicas como el número de nodos, aristas, etc.
- Visualización: Crear una visualización simple del grafo de la red.

5. Propiedades y Métricas de la Red

- Detección de Comunidades: Aplicar algoritmos como Girvan-Newman o Louvain para detectar comunidades dentro de la red.

6. Análisis Avanzado

- **Análisis de Camino Más Corto:** Calcular la longitud promedio del camino más corto.
- **Árboles de expansión Mínima**

7. Visualización

- Visualizaciones Interactivas: Crear visualizaciones interactivas usando Plotly o Gephi para explorar la red.
- Visualización de Comunidades: Visualizar las comunidades detectadas con diferentes colores para resaltar la estructura de la comunidad.

8. Conclusión

- Resumen de hallazgos: Resumir las ideas y hallazgos clave del análisis.
- Trabajo Futuro: Sugerir posibles mejoras y direcciones futuras para investigaciones adicionales.

9. Referencias

Citar todas las fuentes, bibliotecas y conjuntos de datos utilizados en el proyecto.

Todo el esquema deberá ser presentado en un informe tipo paper+presentación+demostración de funcionamiento en laboratorio.