Informe N° 3: Gephi

Gianfranco Pupiales

2024-05-18

Tabla de Contenidos

| 1. Objetivos |
|---|
| 2. Introducción |
| 3. Desarrollo |
| 3.1 Enunciado |
| 3.2 Importación de datos |
| 3.3 Aplicación de algoritmo de distribución $\dots \dots \dots$ |
| 3.4 Personalización de apariencia |
| 3.5 Grado de los nodos |
| 3.6 Detección de comunidades |
| 3.8 Conclusiones del grafo $\dots \dots \dots$ |
| 4. Conclusiones |
| 5. Declaración del Uso de IA |
| 6. Referencias Bibliográficas |

1. Objetivos

- Entender los conceptos fundamentales de los grafos, incluyendo su estructura, tipos y características.
- Utilizar el software Gephi para la representación visual de grafos y para la visualización de algoritmos de distribución, comunidades y métricas.

2. Introducción

Un grafo es una estructura de datos no lineal compuesta por nodos o vértices, así como aristas que conectan dos nodos distintos dentro del grafo. Estas aristas pueden estar ponderadas, lo que significa que tienen un valor asociado que indica algún tipo de medida entre los nodos conectados. Los grafos pueden ser dirigidos, lo que implica que las aristas tienen una dirección

específica que va desde un nodo de origen a un nodo de destino, o no dirigidos, donde las aristas no tienen una dirección definida. Además, los grafos pueden clasificarse en diversos tipos, como grafos simples, grafos ponderados, grafos dirigidos, grafos no dirigidos, grafos cíclicos y grafos acíclicos, entre otros. Estas estructuras son fundamentales en la representación y análisis de relaciones entre elementos en diversas áreas, como la informática, las redes sociales, la logística y la biología, entre otras disciplinas («What is Graph in Data Structure & Types of Graph?—simplilearn.com»).

Gephi, un software de código abierto, ofrece una plataforma para la visualización y análisis de redes. Este software ayuda a los analistas a identificar patrones, tendencias y valores atípicos en los datos de manera intuitiva. Con un motor de renderizado 3D, permite la exploración en tiempo real de grandes gráficos. Gephi facilita funciones para explorar, analizar, filtrar, agrupar, manipular y exportar diversos tipos de redes. Su diseño se centra en la visualización y manipulación de datos, lo que permite a los usuarios descubrir propiedades de las redes de manera accesible («About — gephi.org»).

3. Desarrollo

3.1 Enunciado

Cree un grafo (o use un dataset de tipo grafo) que tenga al menos 20 nodos y 40 aristas. Visualice la data en Gephi, aplique un algoritmo de distribución, personalice la apariencia, calcule el grado (con o sin pesos, dependiendo del grafo), detecte comunidades y de las diferentes métricas obtenga dos o tres conclusiones a partir de su grafo.

3.2 Importación de datos

Para este caso se va a utilizar el siguiente .csv para la creación de los nodos y aristas en Gephi. Esto representa a un usuario de X que hace una publicación vinculada a un hashtag, donde el peso indica la cantidad de publicaciones que ha realizado con ese hashtag.

```
Source, Target, Type, Weight
@luisgomez, diversion, Directed, 5.0
@anacastillo, baile, Directed, 8.0
@javierlopez, motivacion, Directed, 6.0
@carlosramirez, comedia, Directed, 7.0
@patriciamontoya, belleza, Directed, 9.0
@sofiaperez, viaje, Directed, 6.0
@juancastro, fitness, Directed, 8.0
@camilagonzalez, cocina, Directed, 7.0
@pabloalvarez, musica, Directed, 8.0
@laurasoto, arte, Directed, 7.0
```

```
@luisgomez,amigos,Directed,6.0
@anacastillo,fotografia,Directed,7.0
@javierlopez,mascotas,Directed,5.0
@carlosramirez,humor,Directed,8.0
@patriciamontoya,moda,Directed,9.0
Osofiaperez, entrenamiento, Directed, 7.0
@juancastro,aventura,Directed,6.0
@camilagonzalez,desafio,Directed,8.0
@pabloalvarez,inspiracion,Directed,7.0
@laurasoto,creatividad,Directed,6.0
@luisgomez,risas,Directed,7.0
@anacastillo,belleza,Directed,8.0
@javierlopez,consejos,Directed,6.0
@carlosramirez,amor,Directed,7.0
@patriciamontoya,salud,Directed,8.0
Osofiaperez, relax, Directed, 6.0
@juancastro,musica,Directed,7.0
@camilagonzalez,estilo,Directed,8.0
@pabloalvarez,bromas,Directed,6.0
@laurasoto, vida, Directed, 7.0
@luisgomez,diversion,Directed,5.0
@anacastillo,baile,Directed,8.0
@javierlopez,motivacion,Directed,6.0
@carlosramirez,comedia,Directed,7.0
Opatriciamontoya, belleza, Directed, 9.0
Osofiaperez, viaje, Directed, 6.0
@juancastro,fitness,Directed,8.0
@camilagonzalez,cocina,Directed,7.0
Opabloalvarez, musica, Directed, 8.0
@laurasoto,arte,Directed,7.0
@luisgomez,amigos,Directed,6.0
@anacastillo,fotografia,Directed,7.0
@javierlopez,mascotas,Directed,5.0
@carlosramirez,humor,Directed,8.0
Opatriciamontoya, moda, Directed, 9.0
Osofiaperez, entrenamiento, Directed, 7.0
@juancastro,aventura,Directed,6.0
@camilagonzalez,desafio,Directed,8.0
Opabloalvarez, inspiracion, Directed, 7.0
@laurasoto,creatividad,Directed,6.0
@luisgomez,risas,Directed,7.0
@anacastillo,belleza,Directed,8.0
```

```
@javierlopez,consejos,Directed,6.0
@carlosramirez,amor,Directed,7.0
@patriciamontoya,salud,Directed,8.0
Osofiaperez, relax, Directed, 6.0
@juancastro,musica,Directed,7.0
@camilagonzalez,estilo,Directed,8.0
@pabloalvarez,bromas,Directed,6.0
@laurasoto, vida, Directed, 7.0
@juancastro,risas,Directed,8.0
@camilagonzalez,viaje,Directed,7.0
Opabloalvarez, entrenamiento, Directed, 6.0
@laurasoto,belleza,Directed,8.0
@luisgomez,motivacion,Directed,7.0
@anacastillo,amigos,Directed,6.0
@javierlopez,baile,Directed,7.0
@carlosramirez,arte,Directed,8.0
@patriciamontoya,humor,Directed,6.0
Osofiaperez, musica, Directed, 7.0
@juancastro,moda,Directed,8.0
@camilagonzalez,consejos,Directed,6.0
@pabloalvarez,vida,Directed,7.0
@laurasoto,amor,Directed,8.0
```

Hay que destacar que la cabecera inicial de *Source*, *Target*, *Type*, *weight* es muy importante para la importación del conjunto de datos en Gephi para que reconozca a qué se le atribye cada campo. Al iniciar Gephi e importar los datos, se presenta una pantalla similar a la siguiente.

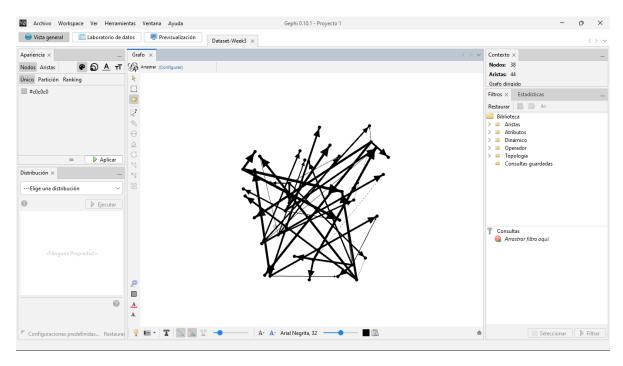


Figura 1: Pantalla al importar un .CSV

3.3 Aplicación de algoritmo de distribución

Para aplicar un algoritmo de distribución, se debe acceder a la pestaña "Vista General" y dirigirse a la sección "Distribución" ubicada en la parte inferior. Allí, se puede seleccionar la distribución deseada, en este caso, "Yifan Hu Proporcional".

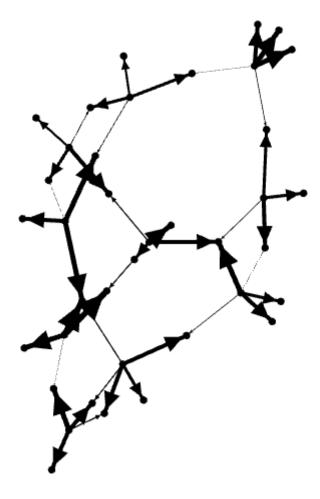


Figura 2: Distribución Yifan Hu Proporcional

3.4 Personalización de apariencia

Para personalizar la apariencia de los nodos o aristas se tiene que ir a la ventana de "Vista General", en la parte superior izquierda se encuentra la pestaña "Apariencia". En ese apartado se puede realizar tanto la modificación para nodos como para aristas. Se puede configurar color, tamaño, color de etiqueta y tamaño de etiqueta.

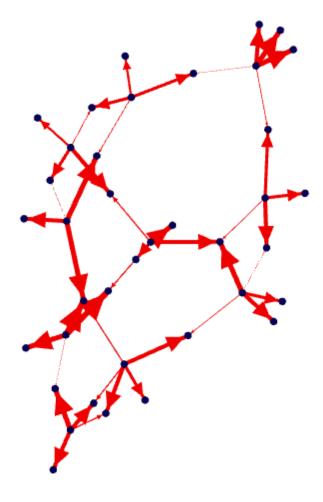


Figura 3: Personalización inicial de grafo

3.5 Grado de los nodos

Para obtener el grado de los nodos, se tiene que acceder a la ventana "Vista general", después a la pestaña de "Estadísticas" ubicada en la parte derecha. Y se tiene que ejecutar la opción de "Grado medio" y "Grado medio con pesos". Al ejecutar esas opciones se puede observar en la ventana de "Laboratorio de datos" la creación de dos columnas con estos nombres respectivamente.

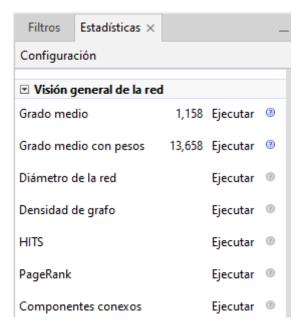


Figura 4: Pestaña de ejecución del cálculo de grados

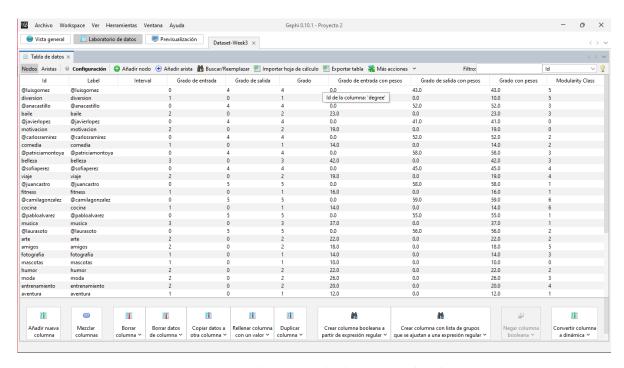


Figura 5: Laboratorio de datos actualizado

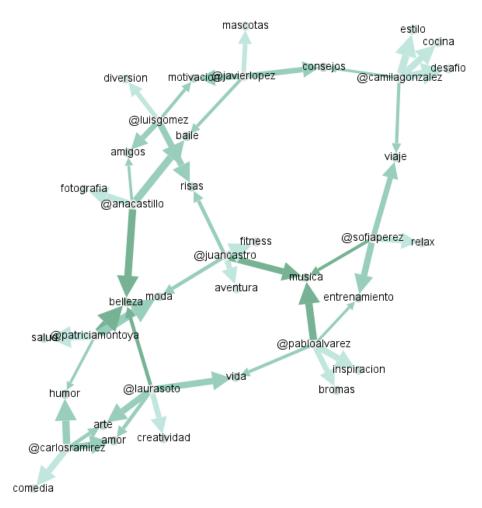


Figura 6: Visualización grado de entrada

3.6 Detección de comunidades

Para poder detectar las comunidades en necesario acceder a la ventana "Vista General", en la pestaña "Estadísticas" se debe dar clic en ejecutar en la parte de "Modularidad". Para poder visualizar los cambios se debe seleccionar la opción de "Modularity Class" ubicada en la pestaña "Nodos" en "Apariencia" en "Vista general".

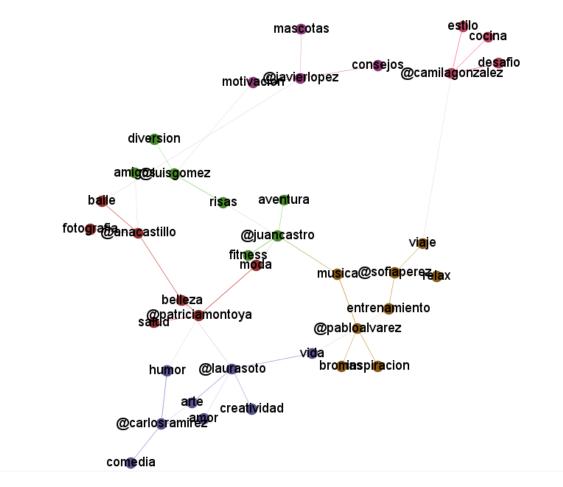


Figura 7: Visualización de comunidades

Si se desea agregar filtros hay que dirigirse a la ventana "Vista General", pestaña "Filtros", seleccionar la opción de "Partición" y seleccionar "Modularity Class (Nodo)". Esto despliega un apartado para poder filtrar por comunidad.

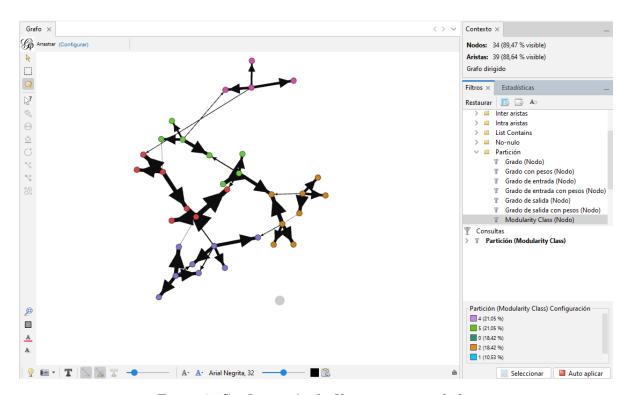


Figura 8: Configuración de filtros en comunidades



Figura 9: Apicación de filtro en comunidades

3.8 Conclusiones del grafo

a) El hashtag más utilizado es "belleza".



Figura 10: Hashtag más utilizado

b) Los hashtags menos utilizados son "diversion" y "mascotas".



Figura 11: Hashtags menos utilizados

c) El usuario de X con más publicaciones es "camilagonzalez".

| ld | Label | Interval | Grado de entrada | Grado de salida | Grado | Grado de entrada con pesos | Grado de salida con pesos ∨ |
|-----------------|-----------------|----------|------------------|-----------------|-------|----------------------------|-----------------------------|
| @camilagonzalez | @camilagonzalez | | 0 | 5 | 5 | 0.0 | 59.0 |

Figura 12: Usuario de X con más publicaciones

d) El usuario de X con menos publicaciones es "javierlopez".

| ld | Label | Interval | Grado de entrada | Grado de salida | Grado | Grado de entrada con pesos | Grado de salida con pesos ∨ |
|--------------|--------------|----------|------------------|-----------------|-------|----------------------------|-----------------------------|
| @javierlopez | @javierlopez | | 0 | 4 | 4 | 0.0 | 41.0 |

Figura 13: Usuario de X con menos publicaciones

4. Conclusiones

- Los grafos son una estructura de datos esencial que representa relaciones y conexiones entre entidades, ya que permiten representar conjuntos de nodos y aristas de manera eficiente y flexible, además su capacidad para manejar tanto estructuras simples como complejas, incluyendo grafos no dirigidos, dirigidos y ponderados, resalta su versatilidad.
- Gephi es una herramienta extremadamente útil para la visualización de conjuntos de grafos pequeños y medianos. Permite aplicar filtros, calcular grados, personalizar la apariencia y explorar entre diferentes distribuciones.

5. Declaración del Uso de IA

En este informe se utilizó IA para mejorar la coherencia, cohesión y sintaxis de las conclusiones. Además el dataset generado fue creado con el uso de IA en base al dataset presentado en clase.

6. Referencias Bibliográficas

- «About gephi.org». https://gephi.org/about/#:~:text=Gephi%20is%20an%20open% 2Dsource,tells%20stories%20with%20their%20data.
- $\begin{tabular}{ll} $\tt www.simplilearn.com/tutorials/data-structure-tutorial/graphs-in-data-structure\#:$$\sim:$$ $$ text=A\%20graph\%20is\%20a\%20non, are\%20also\%20known\%20as\%20vertices. \\ \end{tabular}$