1. Objetivo del laboratorio

Aprender el uso de la herramienta de Social Network Analysis Gephi.

2. Elementos a utilizar:

- Gephi

3. Práctica 1 (Análisis de redes sociales)

Objetivo

1) Abrir Gephi
2) Abrir archivo nodes_got.csv
3) Abrir archivo edges_got.csv. Justo al final marcar la opción "Añadir a especio existente"
4) Pulsa en la T mayúscula esquina de abajo a la izquierda.
5) Elegir la distribución "Force Atlas 2". Marcar "Prevención de solapamiento" y "Disuadir Hubs". Poner escalado a 100. Pulsa Ejecutar. Luego cambiar escalado a 10 y volver a ejecutar.
6) Elegir la distribución "Ajuste de etiquetas", velocidad 2 y ejecutar.
7) Pulsar en Nodos. Elegir un color (paleta), tamaño de nodo en Ranking entre 10 y 100 (círculos concéntricos), color de etiquetas (A) y tamaño etiquetas en Rankings entre 1 y 4.
8) Asignar un color a las aristas. Aplicando Ranking.
9) Ejecutar distribución Fruchterman Reingold con Area 50,000.
10) Volver a ejecutar "Force Atlas 2".
11) Ejecutar en Estadísticas (barra derecha): grado medio, grado medio con pesos, diámetro, densidad del grafo, modularidad, componentes conexos y coeficiente medio de clústering.
12) Volvemos al estilo de los nodos y elegimos el color según el grado de los pesos.
13) Por último, en Nodos, partición elegimos Modularity class y ejecutamos.

Copia el grafo final en un documento Word. Finalmente teniendo en cuenta las métricas conseguidas en el apartado 11 y teniendo en cuenta las definiciones dadas abajo. Aporta gráficas y valores de los resultados de cada una e interpreta estos.

Definition 3. A connected component, Feng et al (2016), is a maximal subgraph of a graph in which any two vertices are connected to each other by a path.

Connected components let us to discover local communities in networks.

Definition 4. A directed graph is strongly connected if there is a path between all the pairs of nodes. If we have a maximal strongly connected subgraph, we can consider it a Strongly Connected Component (SCC) of a graph. It can be calculated by using Tarjan's algorithm, Tarjan (1972), with Nuutila's modifications, Nuutila and Soisalon-Soininen (1999).

Definition 5. A weakly connected graph is when avoiding the directions of the edge it becomes a strongly connected graph.

Definition 6. Diameter, proposed in Tauro et al (2001). Effective diameter or eccentricity is the minimum number of hops in which some fraction (say, 90%) of all connected pairs of nodes can reach each other.

Diameter gives us an idea of how easy is the information to be expanded over all the nodes of a network. A low diameter means that is easier to reach all the nodes of the graph starting from a particular one.

Definition 7: Density the proportion of edges compared to the maximum edges of the graph if it were complete.

Density can be used to know how fast the information is spreading among the network.

Definition 8: Reachability measures the number of nodes to go from one to another, no matter how many you have to pass through.

This measure tells us if a node is more isolated than others drawing possible divisions in the network.

Definition 7. Degree centrality, Opsahl, Agneessens and Skvoretz (2010), taking into account that the degree of a node is its number of connections, was computed as the number of ties or neighbours of a node.

If we talk about undirected graphs, a node with a lot of edges has more possibilities to obtain the information that is flowing in the network. That means that the node will be less dependent to the rest of the network. In directed graphs we have to differentiate between edges reaching a node and those leaving it, we call it in-degree and out-degree respectively. In-degree has to be interpreted as popularity and out-degree as influence.

Definition 8. Closeness was the inverse of the sum of all shortest paths to others or the smallest number of ties to go through to reach all others individually.

The closeness centrality emphasizes in the distance of a node to reach the others. These nodes having a lower closeness centrality are considered a reference point in the network, so spreading information starting from this point will cost less.

Definition 9. Betweenness centrality, introduced by Freeman (1977), as a way to measure how a node can control the relations between other nodes in a social network.

Betweenness centrality of a certain node (its actor centrality) will be given by the proportion of times it is between other nodes for sending information and the number of falls in pathways between other nodes.

Definition 10. Eigenvector centrality. It based on the idea that if a node influenced another node and this one is influencing other, the nodes at the end of the chain will be highly influential.

It is used when there is an interest in ranking the nodes of a network in terms of popularity. Taking into account that a node is popular not only if it has a lot of friends that could be reached in one step. These friends also have to be popular and have to be connected to a lot of nodes. It is similar to how Google ranks the Websites.

4. Práctica 2. MEMORIA DEL LABORATORIO

Obligatorio redacción clara y correcta ortográfica/gramaticalmente. **Cada paso que se haga tiene que estar justificado.**