

## Lab activity: analysis of music networks

### P3

**1. Indicate the number of nodes shared by each pair of graphs (gB, gD, gB', gD', gwD, gwB) and all the graphs together. Use the function num common nodes.**

**(a) How many nodes are shared between gB and gD? What information does this tell us about the importance of the algorithm used by the crawler (i.e. the scheduler) to decide next nodes to crawl?**

Comparteixen 53 nodes. Aquesta informació ens assegura que els artistes que son comuns als dos grafs mantenen algun grau de similitud o connexió.

**(b) How many nodes are shared between gB and gB'? What information does this tell us about the reciprocity of gB? And about the Spotify's artist related algorithm?**

Comparteixen 100 nodes. Donat que els 100 nodes son artistes que com a mínim tenen una connexió mutua amb un altre (bidirectional edge) i que tots els nodes es troben en gB, es pot concloure que aquests nodes representen aquells artistes que podrien arribar a haver fet una col·laboració o interaccions mútues pel que fa a l'algoritme de relació de Spotify.

**(c) How many nodes are shared between gB' and gwB? What information does this tell us about the similarity metric used to construct gwB.**

Comparteixen 83 nodes. El fet que es comparteixin 83 artistes en els dos grafs suposa que aquest nodes (artistes), no només estan units bidireccionalment sinó que també tenen una alta similitat, en termes de característiques, basat en la mètrica utilitzada

**2. Calculate the 25 most central nodes in the graph gB' using both degree centrality and betweenness centrality. How many nodes are there in common between the two sets? Explain what information this gives us about the analyzed graph.**

En comú en tenen 14. És probable que els 14 nodes comuns entre els 25 nodes més centrals identificats mitjançant *degree centrality* i la *betweenness* en gB' siguin nodes rellevants i influents al gràfic.

El fet que hi hagin 14 artistes en comú suggereix que certs artistes juguen un paper central tant pel que fa al nombre de connexions (centralitat de grau) com en la seva posició en l'estructura global de la xarxa (centralitat intermèdia). És probable que aquests nodes comuns siguin centres importants a la xarxa d'artista, que poden servir com a connectors clau entre diferents parts del graf.

**3. Find cliques of size greater than or equal to min size clique in the graphs  $gB'$  and  $gD'$ . The value of the variable min size clique will depend on the graph. Choose the maximum value that generates at least 2 cliques. Indicate the value you chose for min size clique and the total number of cliques you found for each size. Calculate and indicate the total number of different nodes that are part of all these cliques and compare the results from the two graphs.**

En el cas del bfs el valor màxim per generar com a mínim 2 cliques és 7, mentres que en el dfs és 5 i retorna 4 cliques.

El nombre de nodes únics en el cas de  $gB'$  és 15 , mentres que el de  $gD'$  és 16. Comparatvament els dos valors s'aproximen molt, tot i que els cliques en el  $gB'$  son de més tamany que en  $gD'$ .

**4. Choose one of the cliques with the maximum size and analyze the artists that are part of it. Try to find some characteristic that defines these artists and explain it.**

Una vegada s'han analitzat tots els artistes amb les seves respectives caracteristiques es pot concloure amb que el clique analitzat recull artistes que generalment tenen una alta *danceability*. Com a característica que defineix el conjunt seria la **danceability**.