

UNIVERSIDAD DE BUENOS AIRES

INFORME 3:

Comunidades

Raúl Barriga
Mariela Celis
Jimmy Masías
Sebastián Pinto

8 de noviembre de 2016

Índice

1. Introducción	1
2. Partición en clusters	1
3. Relación entre género y estructura de comunidades	2

1. Introducción

La red de este informe trata de una comunidad de delfines de Doubtful Sound, Nueva Zelanda. La comunidad, que se constituye de 62 ejemplares identificados por una marca en la aleta dorsal, fue fotografiada entre 1995 y 2001. A partir de esos datos se construyó la red que contiene 159 links, donde se establece que existe un link entre aquellos individuos que fueron vistos juntos de forma más frecuente que la esperada aleatoriamente, es decir, por un criterio de “compañía preferida”. [D. Lusseau, The emergent properties of a dolphin social network, Proc. R. Soc. London B (suppl.) 270, S186-S188 (2003).]

2. Partición en clusters

Implementamos diferentes algoritmos para la detección de comunidades en la red de delfines. La asignación de cada algoritmo se puede observar en la figura 1, en la cual se puede observar que todos los algoritmos detectan entre 4 y 6 comunidades presentes en la red, aunque en algunos casos aparecen comunidades compuestas por solo dos delfines, es decir, de un tamaño considerablemente menor que el las restantes, lo cual se podría considerar la absorción de esta pequeña comunidad por parte de otra de mayor tamaño.

En la tabla I calculamos la modularidad y el silhouette dada por cada algoritmo. Para cuantificar qué tan pertinente es la partición en clusters de esta red, estudiamos estos mismos observables recableado en forma aleatoria la red original, manteniendo la membresía a las particiones inalterable. Eso nos resultó en una distribución de valores de modularidad y silhouette de la figura 2. Si se compara esta distribución con los valores de la tabla I, se puede concluir que tanto la modularidad como el silhouette es significativamente mayor en la red actual que la esperada por azar (hipótesis nula), con lo cual decimos que esta red se describe correctamente como una red compuesta por comunas.

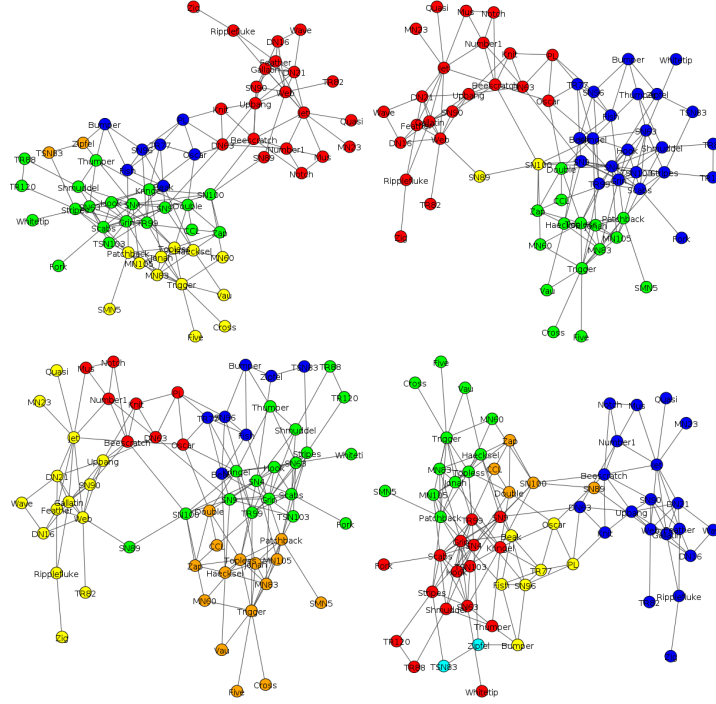


Figura 1: Layouts indicando el cluster asignado por cada algoritmo. De izquierda a derecha, y arriba hacia bajo: Edge betweenness, Fast greedy, Louvain e Infomap.

Algoritmo	Modularidad	Silhouette
Edge-betweenness	0.519	0.338
Fast greedy	0.495	0.184
Louvain	0.519	0.294
Infomap	0.529	0.328

Tabla I: Modularidad y silhouette de las particiones dadas por diferentes algoritmos.

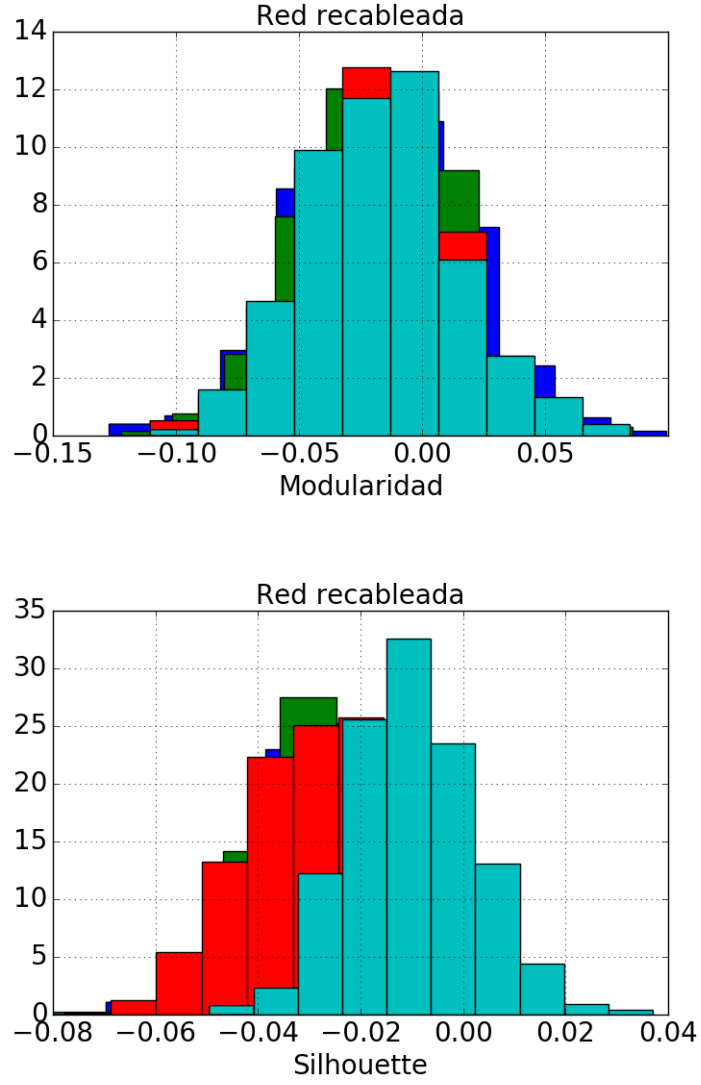


Figura 2: Modularidad y silhouette recableando en forma aleatoria, manteniendo la pertenencia a cada cluster dada por los algoritmos de detección de particiones. Se puede observar que la probabilidad de obtener los valores de la tabla I dada una reconexión aleatoria es prácticamente nula.

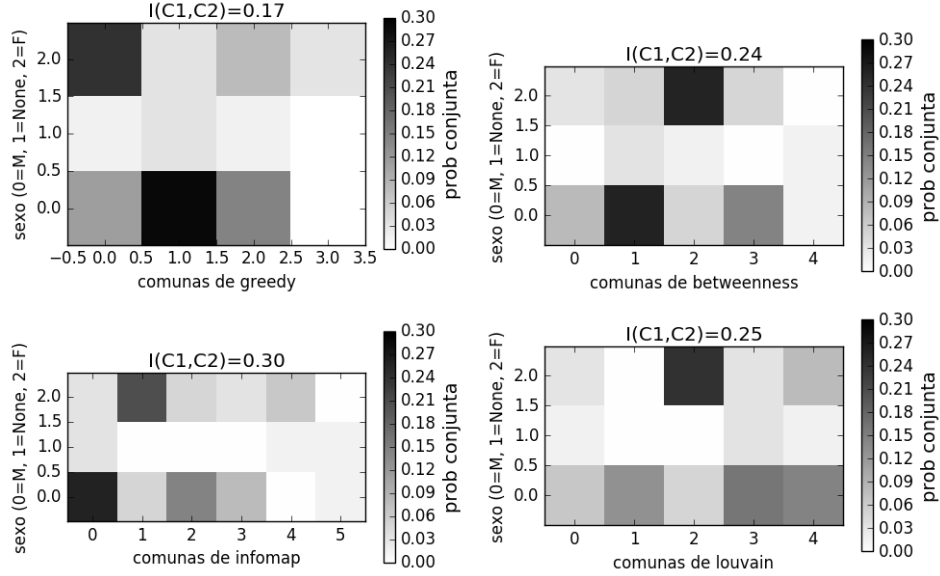


Figura 3: Valores de las matrices de probabilidad conjunto para los algoritmos *greedy* (izquierda, arriba) *betweenness* (derecha, arriba), *infomap* (izquierda, abajo) y *louvain* (derecha, abajo).

3. Relación entre género y estructura de comunidades

Para cuantificar la relación entre el género de los delfines y la estructura de comunidades que fueron deducidas por los diferentes algoritmos (e.g. *Greedy*), empleamos la definición de *Informacion Mutua*:

$$I(\{c1\}, \{c2\}) = \sum_{c1, c2} p(c1, c2) \log \frac{p(c1, c2)}{p(c1)p(c2)}, \quad (1)$$

donde $c1$ y $c2$ son etiquetas de las comunas del grafo. Las comunas construidas en el presente grafo fueron deducidas usando los algoritmos *greedy*, *betweenness*, *infomap* y *louvain*. La definición 1 cuantifica cuanto se departa la información real del grafo, respecto de la cantidad de información que brinda una grafo cuyas comunas estan descorrelacionadas del género. (JM: che Seba, a ver si esto tiene sentido)