



Lección 5. Subgrupos y grupos en el Análisis de Redes Sociales



Análisis de Redes Sociales

Subgrupos y grupos en el análisis de Redes Sociales

UNIVERSIDAD D CORDOBA

El Análisis de Redes Sociales (ARS) estudia la interacción entre los actores en redes sociales. Es un área de investigación interdisciplinar en la que están implícitas la Sociología, la Psicología, la Antropología, la Física, las Matemáticas y la Informática, entre otras.

Los orígenes del ARS, para el desarrollo de conceptos sociológicos, puede situarse a comienzos de la década de 1930, donde el enfoque sociométrico se transforma como una forma de conceptualizar la estructura de las relaciones sociales establecidas entre pequeños grupos de individuos.

En este tema veremos subestructuras mayores que las diadas y triadas y cómo los diferentes elementos de una red interactúan para formar grupos entre ellos.

1. Introducción

El ARS implica tener en cuenta múltiples niveles de análisis (Figura 1). Las diferencias entre los actores se pueden interpretar en base a las limitaciones y oportunidades que surgen de la forma en que éstos están inmersos en las redes; la estructura y el comportamiento de las redes están basados en las interacciones locales entre los actores.

En esta lección examinaremos el análisis estructural de las *subestructuras* que pueden estar presentes en una red. Las diadas y triadas pueden ser pensadas como subestructuras. Las combinaciones de diadas y triadas en estructuras mayores, las cuales están conectadas a su vez, desarrollan o constituyen las redes. Muchas de las aproximaciones para comprender la estructura de una red enfatizan la forma en la que las conexiones están compuestas y se extienden para desarrollar subestructuras más grandes o subagrupaciones. Definir estos grupos, nos permite identificar subpoblaciones en la red y la cohesión que existen entre ellos.

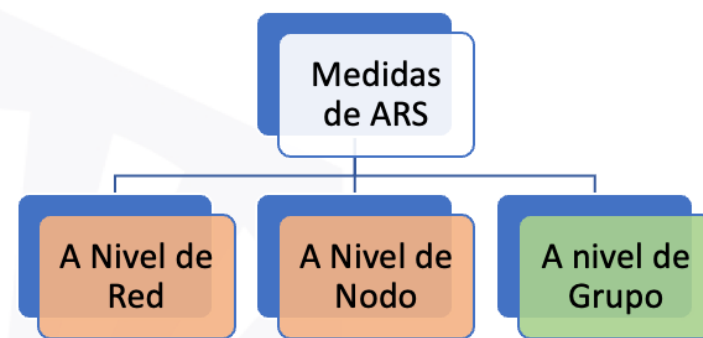


Figura 1. Diferentes niveles de análisis de redes sociales.

Recordamos que en un ARS podemos valorar diferentes medidas a distintos niveles que iremos viendo a lo largo del curso. En esta lección veremos las medidas a nivel de nodo [1]:

- **Medidas globales (a nivel de red):** proporcionan información más compacta que permite evaluar la estructura global de la red, proporcionando información sobre propiedades importantes de los fenómenos sociales subyacentes.
- **Medidas locales (a nivel de nodos):** proporcionan información más específica evaluando la estructura y relaciones de cada nodo en la red. Están basadas en el concepto general de centralidad (redes no dirigidas) o prestigio (redes dirigidas), una medida general de la posición de un nodo en la estructura global de la red social. Se usan para identificar los nodos claves de la red. Muestran como las relaciones se concentran en unos pocos individuos, dando una idea de su poder social.
- **Medidas a nivel de grupo:** nos permiten analizar las diferentes agrupaciones que se pueden realizar entre los nodos y cómo de cohesionados están esos grupos con respecto al resto.

2. La importancia de los grupos

La visión de la estructura social centra la atención en cómo la conexión de grandes estructuras puede ser construidas a partir de componentes pequeños y cohesionados: un tipo de aproximación de abajo hacia arriba. Los analistas de redes han desarrollado un conjunto de definiciones y algoritmos para identificar los componentes más pequeños en las grandes estructuras: cliques, n-cliques, n-clans y k-núcleos sirven para ver las redes desde esa perspectiva.

La división de actores en subgrupos puede ser un aspecto muy importante de la estructura social. Nos permite comprender el comportamiento de la red en su conjunto. Por ejemplo, suponga que los actores de una red forman dos subgrupos, pero que existe alguna yuxtaposición de miembros, esto es, que algunos actores pertenecen a ambos subgrupos. Donde los grupos coinciden, podemos esperar que el conflicto entre ellos sea menos probable que en aquellos dónde no coinciden. Donde los grupos coinciden, la movilización y difusión pueden producirse más rápidamente a través de toda la red; donde no hay yuxtaposición de los grupos, pueden observarse rasgos de un grupo que no se difunden al otro.

Saber cómo un individuo está inmerso en la estructura de grupos en una red puede ser un aspecto crítico para la comprensión de su conducta. Por ejemplo, algunos pueden actuar como puentes entre grupos (ampliadores de fronteras). Otros pueden tener todas sus relaciones dentro de un único grupo (los locales o internos). Algunos actores pueden ser parte de una élite cerrada y densamente conectadas, mientras que otros están completamente aislados de ese grupo.

La idea de subestructuras, o grupos en una red es una poderosa herramienta para comprender la estructura social de los individuos. La mayoría de los algoritmos computacionales para definir los diferentes subgrupos operan sobre datos binarios simétricos. Así, en el caso de trabajar con redes dirigidas, se hace necesario simetrizar los datos. Existen diferentes formas de realizar la simetrización; por ejemplo, contar los lazos recíprocos, es decir, un vínculo existe sólo si xy e yx están presentes. Esto se hace para reducir la densidad de la matriz sustancialmente. Las matrices con una densidad muy alta, casi por definición, tienen pocos subgrupos distinguibles.

De nuevo, utilizaremos la red de amistad del club de kárate para aplicar las medidas que se estudian en esta lección. Al ser un grafo no dirigido, no es necesario simetrizarlo. En la figura 2 se muestra la red social de amistades entre los 34 integrantes de un club de kárate en una universidad estadounidense en la década de 1970. Recordamos que se trata de una red pequeña, con 34 nodos y 2 grupos diferenciados. Los dos grupos eran representados por los dos colores de los nodos y enlaces. Además, el tamaño de los nodos estaba en función del grado de cada actor, esto es, del número de conexiones que tienen. Así, en cada grupo nos encontramos actores más representativos por tener un mayor número de conexiones. Esta red será la que se use de ejemplo para cada una de las medidas que se estudien.

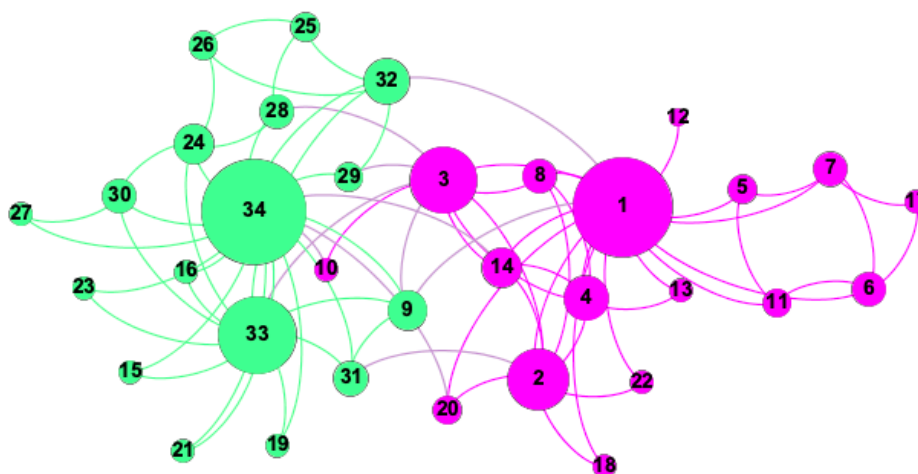


Figura 2. Red social de amistades de un club

Los principales rasgos de un grafo, en términos de subgrafos, se pueden identificar por observación directa: ¿Cómo de separados están los subgrafos (se yuxtaponen y comparten miembros, o dividen y crean facciones en la red)? ¿De qué tamaño son los subgrafos conectados? Hay pocos grupos grandes o un gran número de grupos pequeños? ¿Hay actores particulares que parecen desempeñar roles en la red? Por ejemplo, son nodos que conectan el grafo o hay quienes están aislados de los grupos?

3. Aproximaciones de abajo hacia arriba

Todas las redes están compuestas de grupos. Cuando dos nodos tienen un enlace, forman un *grupo*. Una aproximación para pensar acerca de la estructura grupal de una red comienza con este grupo básico y busca cuánto se puede extender esta relación cercana. Un clique extiende la diada añadiendo a ella miembros que están vinculados con todos los miembros de un grupo. Se puede relajar esta definición estricta de modo que sea posible incluir nodos adicionales que no están muy fuertemente vinculados (n-cliques, n-clans y k-núcleos). La noción, sin embargo, es partir de los vínculos simples para construir la red. Un mapa de toda la red puede ser construido examinando los tamaños de los distintos cliques y agrupaciones de tipo clique notando sus tamaños y yuxtaposiciones.

Este tipo de aproximación, acerca de las subestructuras de las redes, tienden a enfocar nuestra atención primero en individuos y en entender cómo están inmersos en la estructura mayor de la red a partir de los grupos yuxtapuestos.

3.1 Cliques

Un clique es un subconjunto de una red en el cual los nodos están más cercanos y fuertemente conectados que con el resto de los integrantes de la red. En lazos de amistad, por ejemplo, es usual que se formen cliques en función de la edad, el género, la raza o la ideología. Los cliques más pequeños están formados por dos actores: la diada. Pero la diada puede ser extendida para considerar un número mayor de nodos.

La definición más fuerte de un clique es un determinado número de nodos (más de dos, normalmente se usa tres como valor mínimo), que tienen todos los enlaces posibles presente entre ellos. Una agrupación de este tipo se conoce como *subgrafo máximo completo*, el cual se expande hasta incluir a la mayor cantidad de actores posibles.

En este caso, los resultados que se muestran son usando el software UCINET [3] para calcular los grupos y subestructuras, ya que Gephi no las tiene disponible. También, es posible encontrar librerías en Python [4] y R [5] que permiten obtener estos grupos.

En la Figura 3 se muestran los resultados de cliques de tamaño 3 o mayores que se pueden encontrar en la red de amistades del club de kárate (Figura 2). Hay siete 25 subgrafos máximos completos presente en estos datos. El mayor está compuesto por cuatro de los 34 actores. El nodo 1 presenta yuxtaposición con 11 de los 25 grupos, y el nodo 24 presenta también yuxtaposición otros 11 de los 25 nodos. Además, tanto el nodo 1 como el nodo 34 no comporten ningún grupo en común, con lo que son centrales y muy independientes entre ellos. También, se puede apreciar que todos los nodos pertenecen a algún grupo y que hay nodos que solamente pertenecen a uno de los grupos. Esto nos puede dar información de los actores que son más "centrales" y cuáles están más aislados de los cliques. Se aprecia también que el nodo 1 y 2 comparten la participación en 5 grupos, y los nodos 33 y 34 comparten 8 grupos.

25 cliques found.

```

1:  24 30 33 34
2:  16 33 34
3:  19 33 34
4:  21 33 34
5:  23 33 34
6:  15 33 34
7:  9 31 33 34
8:  32 33 34
9:  27 30 34
10: 24 28 34
11: 29 32 34
12: 1 2 3 4 14
13: 1 2 3 4 8
14: 1 2 18
15: 1 2 20
16: 1 2 22
17: 1 3 9
18: 3 9 33
19: 1 4 13
20: 1 5 11
21: 1 5 7
22: 6 7 17
23: 1 6 7
24: 1 6 11
25: 25 26 32

```

Figura 3. Cliques de tamaño 3 o superior (red de amistad del club de kárate, figura 2).

3.2 N-Cliques

La definición estricta de subgrupo máximo completamente conectado puede resultar demasiado fuerte. Requiere que todo miembro de un subgrupo tenga un lazo directo con todos y cada uno de los otros miembros. Se pueden pensar casos de cliques en los que al menos algunos de sus miembros no estén conectados tan fuertemente. En N-clique se relaja la definición de pertenencia de los clique para hacerla más general. Un actor es miembro de un N-clique si está conectado con todos los otros miembros del grupo a una distancia N. Usualmente, se usa la distancia de $N=2$. Esto es equivalente a ser amigo de un amigo.

Cuando aplicamos la definición N-clique a la red de la figura 2, obtenemos la información que se muestra en la figura 4. Hay 12 grupos de tamaño 8 el más pequeño y tamaño 18 el más grande. Como hemos relajado la exigencia de cercanía para ser considerado miembros, hay menos grupos y son más grandes. No obstante, los nodos 1 y 34 siguen siendo críticos, y ahora comparten varios grupos junto con otros nodos.

12 2-cliques found.

```

1:  3 9 10 14 15 16 19 20 21 23 24 28 29 30 31 32 33 34
2:  1 2 3 4 9 10 14 20 28 29 31 32 33 34
3:  1 2 3 4 8 9 10 14 20 28 29 31 32 33
4:  1 2 3 4 8 9 14 18 20 22 31 32
5:  1 2 3 4 5 6 7 8 9 11 12 13 14 18 20 22 32
6:  3 24 25 28 29 32 33 34
7:  1 3 25 28 29 32 33 34
8:  24 26 28 29 30 32 33 34
9:  24 25 26 28 29 32 33 34
10: 1 25 26 28 29 32 33 34
11: 9 10 14 15 16 19 20 21 23 24 27 28 29 30 31 32 33 34
12: 1 5 6 7 11 17

```

Figura 4. N-Cliques de tamaño 3 o superior con $N=2$ (red de amistad del club de kárate, figura 2).

3.3 N-Clanes

El enfoque N-clique, tiende a encontrar agrupaciones grandes y difusas en lugar de las más fuertes y discretas del enfoque máximo. En algunos casos, se pueden encontrar N-cliques que tienen la propiedad de la conectividad con actores que no son en sí mismos miembros de un clique. Para evitar esta situación, los N-clanes definen la pertenencia por medio de exigir que la distancia entre cualquier par de miembros del grupo debe realizarse por medio de miembros del grupo.

Cuando aplicamos la definición N-clan a la red de la figura 2, obtenemos la información que se muestra en la figura 5. Hay 8 grupos de tamaño 6 el más pequeño y tamaño 18 el más grande. Algunos grupos son compartidos con los grupos que se obtuvieron en n-clique (4 y 7). En este caso el nodo 34 aparece en un mayor número de grupos que el nodo 1, y también comparten grupos.

8 2-clans found.

```

1:  1 2 3 4 9 10 14 20 28 29 31 32 33 34
2:  1 2 3 4 5 6 7 8 9 11 12 13 14 18 20 22 32
3:  3 24 25 28 29 32 33 34
4:  1 3 25 28 29 32 33 34
5:  24 26 28 29 30 32 33 34
6:  24 25 26 28 29 32 33 34
7:  9 10 14 15 16 19 20 21 23 24 27 28 29 30 31 32 33 34
8:  1 5 6 7 11 17

```

Figura 5. N-Clan con N=2 (red de amistad del club de kárate, figura 2).

4. Enfoques de arriba-abajo

Los enfoques anteriores comienzan con la díada y examinan si este tipo de estructura cohesiva se puede extender hacia afuera. En los enfoques que vemos en esta sección se comienza con la totalidad de la red como su marco de referencia e identifican las subestructuras como partes que son localmente más densas. Veremos algunas de las técnicas que nos permiten la búsqueda de estas subestructuras.

4.1 Componentes

Los componentes de un grafo son partes que están internamente conectadas, pero desconectadas entre los subgrafos. Si un grafo contiene uno o más nodos aislados, estos actores son componentes. Más interesantes son aquellos que dividen los grafos en partes separadas, en las que cada una tiene diversos actores que están recíprocamente conectados (no nos importa cuán cercanamente ligados estén).

En la red de la figura 2 solamente hay una componente. Esto es, todos los actores están conectados. La noción de componente es normalmente muy estricta para encontrar significativos puntos débiles, agujeros o subpartes locales densas en un grafo grande siendo la mayoría de ellas de una sola componente.

4.2 Conjuntos lambda y puentes

En este caso se buscan si hay relaciones claves, las cuales si se eliminasen formarían una estructura desconectada y cada una de ellas se denominan conjuntos lambda. Se buscan relaciones (puentes) y no actores claves. El enfoque del conjunto Lambda ordena cada relación en la red en términos de su importancia por medio de la valoración del flujo entre actores que pasa a través de cada nexo. Luego identifica el conjunto de actores que, si fuesen desconectados, perturbarían significativamente el flujo entre todos los actores.

Cuando aplicamos esta definición a la red de la figura 2, obtenemos la información que se muestra en la figura 7. Se puede ver que el vínculo entre los nodos 33 y 34 es el que se identifica como más importante, en el sentido de que transportan una gran cantidad de tráfico, y si se eliminasen la comunicación se vería muy afectada. El siguiente enlace más crítico es el que los une con el nodo 1 y el nodo 3, posteriormente el nodo 2, y así se puede ir comprobando la relevancia a través del gráfico.

Lambda	1	1	1	1	1	1	1	1	2	2	2	2	1	2	2	2	2	2	3	3	1	2	3					3	3					
	2	0	3	5	6	7	8	9	1	2	3	7	5	1	0	5	6	9	6	7	8	8	0	1	9	4	4	4	2	2	1	3	3	4
12	XXX
10	XXXXXXXXX
9	XXXXXXXXXX
6	XXXXXXXXXXXXXXXXX
5	XXXXXXXXXXXXXXXXXXXXX
4	XXXXXXXXXXXXXXXXXXXXX
3	XXXXXXXXXXXXXXXXXXXXX
2	.	XX																																
1	XX																																	

Figura 7. Conjuntos lambda de la red de amistad del club de kárate (figura 2)

4.3 Facciones

Las facciones buscan nodos que sean equivalentes en la medida que tengan el mismo perfil de relaciones con otros nodos. Se definen particiones de la red sobre la base de agrupar juntos a nodos por su similitud en cuanto a con quién están vinculados. En un extremo, podríamos tener un grafo que fuese divisible en componentes y que cada componente es un clique (esto es, en un grupo, todos están vinculados con todos) y entre grupos no hay lazos. Algunas divisiones menos estrictas permitirían que hubiese algunos vínculos entre grupos y una densidad no completa dentro de los grupos.

Para obtener las facciones se debe indicar el número de grupo o divisiones que se quiere realizar. Para la red de la figura 2 y utilizando la opción de 4 grupos se obtiene las facciones que se muestran en la Figura 8. Se puede ver que seleccionando 4 grupos nos encontramos con grupos más fuertes y débiles en la estructura. Existen grupos entre los que no existe ninguna relación, existen muy pocas y otros que están altamente conectados. Las relaciones entre las facciones también nos dan información de potenciales aliados y enemigos.

```

1: 5 6 7 10 11 12 17
2: 24 25 26 27 28 29 30 32
3: 1 2 3 4 8 13 14 18 20 22
4: 9 15 16 19 21 23 31 33 34

```

	1 1 1 1 7 0 1 2 5 6 7	2 2 2 2 2 3 3 2 5 6 8 7 4 0 2 9	1 1 2 4 8 3 4 1 2	2 2 1 8 0 3	2 1 1 2 3 9 6 1 9 1	3 1 3 3 5 3 4
17		1 1				
10			1			1
11		1 1		1		
12				1		
5		1		1		
6	1	1		1		
7	1	1 1		1		
25		1 1				
26		1				
28		1	1			1
27						1
24		1 1				1 1
30			1 1			1 1
32		1 1		1		1 1
29			1	1		1
2			1 1 1 1 1 1 1 1		1	
14			1	1 1 1		1
18			1	1		
3	1	1	1 1	1 1	1	1
4			1 1	1 1	1 1	
1	1 1 1 1 1		1 1 1 1 1	1 1 1 1	1	
22			1	1		
8			1	1 1 1		
20			1	1		1
13				1 1		
23						1 1
19						1 1
16						1 1
21						1 1
9			1	1	1	1 1
31			1		1	1 1
15						1 1
33		1 1 1	1		1 1 1 1 1 1	1
34	1	1 1 1 1 1 1	1	1	1 1 1 1 1 1 1 1	

Figura 8. Facciones de la red de amistad del club de kárate (figura 2)

En la figura 9 se muestra las facciones utilizando la opción de 2 grupos. Es interesante observar que se formaría los grupos que están representados en la Figura 2 con diferente color. Se mostraría así dos grupos muy conectados entre ellos y poco conectados con el otro. Estos grupos representan los que están a favor del instructor del club y los que están a favor del presidente, ambos lideran cada grupo y no están relacionados entre ellos.

Group Assignments:

1: 1 2 3 4 5 6 7 8 10 11 12 13 14 17 18 20 22
2: 9 15 16 19 21 23 24 25 26 27 28 29 30 31 32 33 34

Grouped Adjacency Matrix

	1	2	3	4	5	6	7	8	7	0	1	2	3	4	0	8	2	1	5	9	9	1	6	3	4	5	6	7	8	9	0	1	2	3	3	3	3
1		1	1	1	1	1	1	1			1	1	1	1	1	1	1	1	1																		1
2		1		1	1			1						1	1	1	1		1															1			
3		1	1		1					1					1				1										1	1						1	
4		1	1	1				1					1	1																							
5		1						1							1																						
6		1						1		1					1																						
7		1				1	1			1																											
8		1	1	1	1																																
17								1	1																												
10				1																																	1
11		1						1	1																												
12		1																																			
13		1				1																															
14		1	1	1	1																																1
20		1	1																																		1
18		1	1																																		
22		1	1																																		
15																																					1
9		1		1																																	1
19																																					1
21																																					1
16																																					1
23																																					1
24																																					1
25																																					1
26																																					1
27																																					1
28																																					1
29																																					1
30																																					1
31																																					1
32																																					1
33																																					1
34																																					1

Figura 9. Facciones de la red de amistad del club de kárate (figura 2)

5. Resumen

Uno de los aspectos más interesantes acerca de las estructuras sociales son las agrupaciones. El número, tamaño y las conexiones entre los grupos de una red nos dicen mucho acerca de la conducta probable de la red en su conjunto, por ejemplo, si la información va a circular rápido por esa red o si es más probable que los conflictos involucren a múltiples grupos o no. Un estudio de cómo se combinan y mezclan los grupos, es decir, si los grupos comparten diferentes grupos entre sí, es nos permite predecir comportamientos en la conducta de la red completa.

La ubicación de los individuos en redes puede también ser pensada en lo relevante que son para dividir grupos. Algunos individuos pueden actuar como puentes entre grupos, otros como aislados; algunos actores pueden estar en todos los grupos y otros locales, en términos de sus afiliaciones grupales. Tales variaciones en las formas que los individuos se conectan a grupos puede tener consecuencias para su comportamiento como individuos.

Referencias

- [1] Borgatti, S. P., Everett, M. G., & Johnson, J. C. (2018). *Analyzing social networks*. Sage.
- [2] Dey, N., Borah, S., Babo, R., & Ashour, A. S. (2018). *Social network analytics: computational research methods and techniques*. Academic Press.
- [3] Ucinet Software. <https://sites.google.com/site/ucinetsoftware/home> [último acceso, junio, 2022]
- [4] NetworkX. Análisis de redes en Python. <https://networkx.org/> [último acceso, junio, 2022]
- [5] Igraph. Análisis de redes en R. <https://igraph.org/r/> [último acceso, junio, 2022]