

UNIVERSIDAD COMPLUTENSE DE MADRID
FACULTAD DE INFORMÁTICA
ANÁLISIS DE REDES SOCIALES

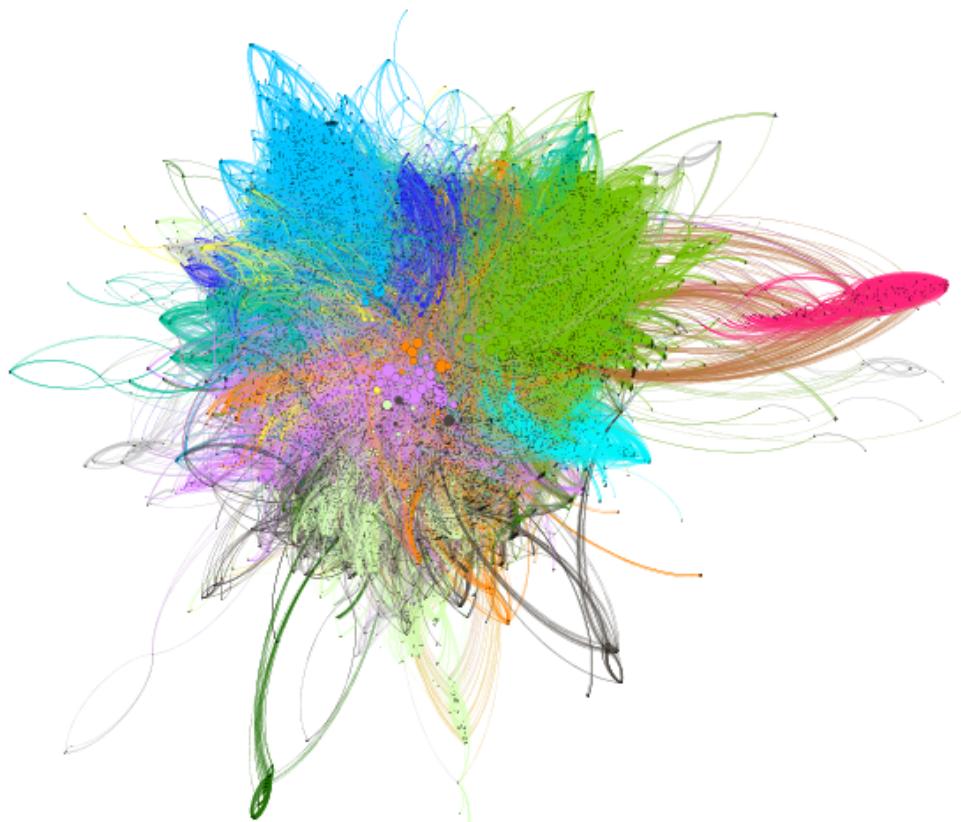


Fig 1: Vista general de la red distribuida en comunidades

ANÁLISIS DE UNA RED
PRÁCTICA 1

Belén García Puente
Ela Katherine Shepherd Arévalo
Víctor Santamaría Gredilla

1. Introducción	3
2. Métricas generales	5
2.1 Componentes conexas	5
2.2 Grado	6
2.3 Centralidades	10
2.3.1 Centralidad de intermediación	10
2.3.2 Centralidad de cercanía	12
2.4 HITS	14
2.5 PageRank	16
2.6 Clustering	18
3. Análisis de comunidades	20
3.1 X-Men	20
3.2 Los Vengadores	22
3.3 Spider-Verse	23
3.4 Dr. Strange y Hulk	25
3.5 Fantastic Four	27
3.6 Comportamientos especiales	28
3.6.1 Mutantes X	28
3.6.2 Capitán América	29
4. Bibliografía	31

1. Introducción

Para esta práctica se ha escogido una red social relacionada con los superhéroes del universo de Marvel. Cada nodo es un personaje, y las aristas se establecen entre aquellos nodos que han entrado en contacto en algún punto de este universo. Durante todo este proyecto, estamos asumiendo que la red habla tan solo del universo Tierra-616.

Esta red consta de 10469 nodos y de 178115 aristas que los relacionan entre ellos, indicando que dos personajes se conocen.

La red se obtuvo a partir de este link: [LINK](#)

En primer lugar, dado que el archivo base era ".gephi", hemos comenzado a analizar la red con éste. Una vez obtenida la información relevante, hemos exportado a NetworkX la red en formato ".graphml" para realizar un análisis más completo y comparar algunos de los resultados obtenidos para cumplimentar las conclusiones.

En el caso de Gephi, ha resultado especialmente útil para las tareas de visualización de la red, pudiendo aplicar diversos filtros de tamaño y color para la visualización de los nodos y las diferentes métricas.

Respecto a NetworkX, hemos utilizado la herramienta principalmente con el fin de facilitar la visualización en lista de los elementos más importantes de las métricas tomadas. Esto nos ha permitido no solo corroborar los hallazgos obtenidos en Gephi sino también mostrar, de manera cuantitativa y no sólo gráfica, aquellos nodos de la red más relevantes.

Información general

```
In [5]: print(nx.info(G))
```

```
Name:  
Type: Graph  
Number of nodes: 10469  
Number of edges: 178115  
Average degree: 34.0271
```

```
In [51]: # Número de componentes conexas  
print("Componentes conexas:", nx.number_connected_components(G))
```

```
Componentes conexas: 8
```

Fig 2: Información general de la red

2. Métricas generales

2.1 Componentes conexas

Results:

Number of Weakly Connected Components: 8

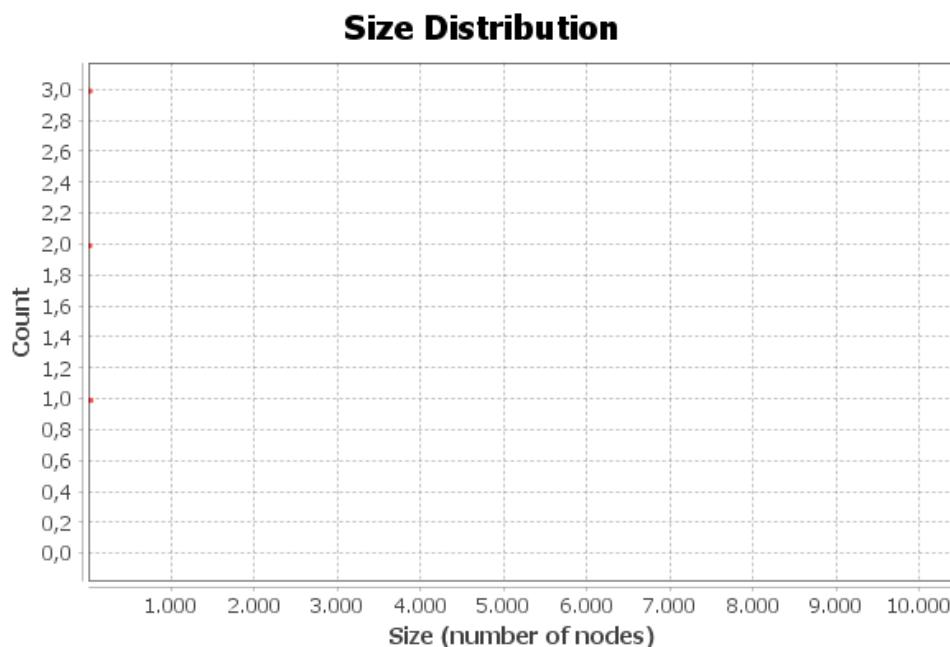


Fig 3: Distribución de componentes conexas

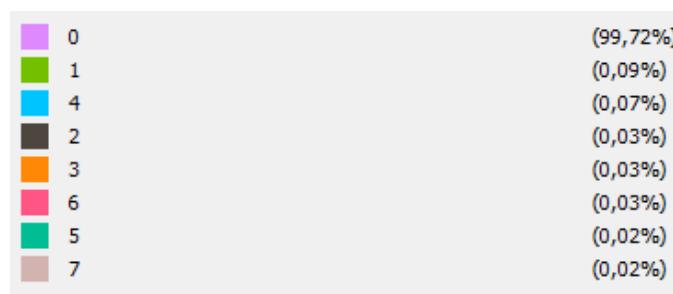


Fig 4: Cantidad de nodos de cada componente conexa

En la red encontramos un total de 8 componentes conexas, pero es importante recalcar el peso que tienen, puesto que la principal ocupa el 99.7% de los nodos y el 99.96% de las aristas, siendo las otras 7 residuales. Entendemos éste resultado de dos maneras:

- La primera, que existen componentes conexas no relacionadas con la red principal.

- La segunda, que debido a la naturaleza del dataset y del archivo de origen, existe un margen de error donde algunos datos están poblados con erratas.

Nodos: 10440 (99,72% visible)
Aristas: 178049 (99,96% visible)
Grafo no dirigido

Fig 5: Datos de la componente conexa más grande de la rd

2.2 Grado

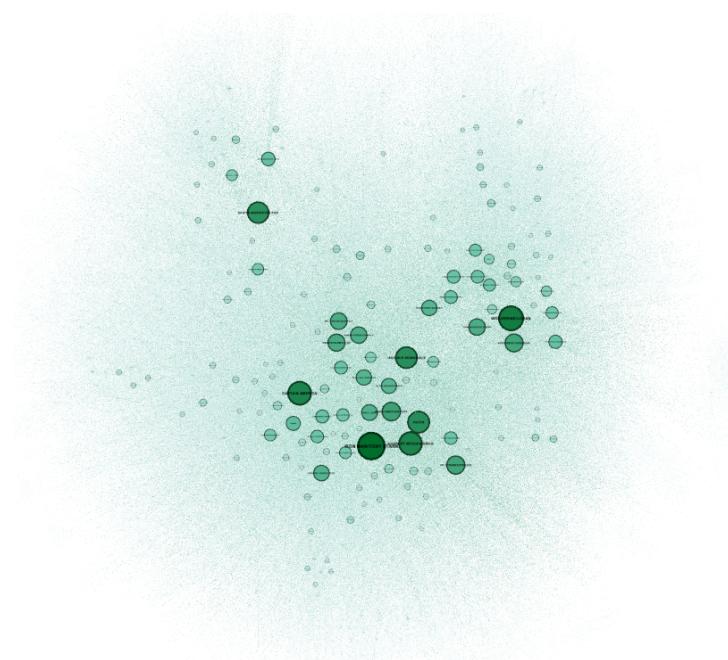


Fig 6

La imagen muestra una vista general de la red, con todos sus nodos. Cuanto más grande sea el nodo, y más oscuro, mayor grado tiene. Así, podemos observar que la gran mayoría de nodos son pequeños, y que tan solo unos pocos tienen una gran cantidad de conexiones con distintos personajes del MCU (*Marvel Comic Universe*).

En Gephi, el grado medio obtenido ha sido de 34.027, coincidiendo con el obtenido en NetworkX, con un valor de 34,0271.

Considerando el elevado número de personajes poco relevantes y/o de sagas poco relacionadas con el MCU, podemos decir que el grado medio por nodo es bastante elevado.

Profundicemos ahora en los nodos más destacables. En la siguiente imagen se puede observar que hemos eliminado los nodos menos significativos, dejando aquellos que cuentan con mil o más conexiones en la red. Dichos nodos se dividen en comunidades según los colores.

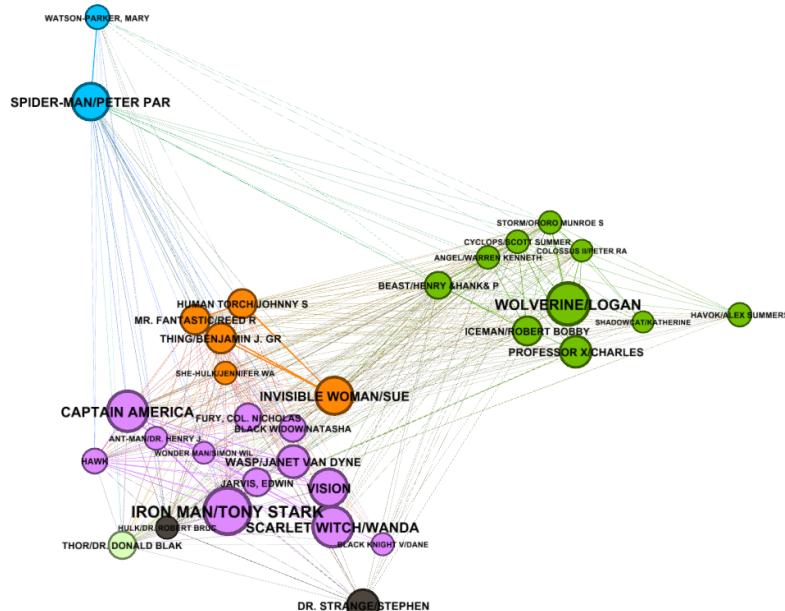


Fig 7: Nodos con un grado igual o mayor que 1000

```

aux = sorted(G.degree, key=lambda x: x[1], reverse=True)[0:10]
for x in aux:
    print(G.nodes[x[0]]["label"], "= ", x[1], " nodos")

IRON MAN/TONY STARK = 2189 nodos
WOLVERINE/LOGAN = 1984 nodos
CAPTAIN AMERICA = 1896 nodos
SCARLET WITCH/WANDA = 1883 nodos
VISION = 1765 nodos
INVISIBLE WOMAN/SUE = 1750 nodos
SPIDER-MAN/PETER PAR = 1730 nodos
WASP/JANET VAN DYNE = 1517 nodos
DR. STRANGE/STEPHEN = 1489 nodos
PROFESSOR X/CHARLES = 1471 nodos

```

Fig 8: Nodos con mayor grado

Los nodos más importantes son también los personajes más famosos tales como Capitán América, Spider-Man, Wolverine y Iron-Man. Centrémonos un poco más en ellos.

Si indagamos un poco más en la historia de Capitán América, encontramos su nombre en casi ochocientos cómics de la industria. Además, sus numerosas apariciones no se encuentran solo en su propia saga, sino que también tiene cameos en cómics como Daredevil, Iron Fist y en el universo de los X-Men. Al no especificar que se trata de Steve Rogers, estamos asumiendo que el nodo de esta red toma como referencia a todos aquellos que han llevado el nombre del Capitán América.

Iron-Man, por su parte, no se queda atrás. Su número de apariciones asciende hasta casi setecientos cincuenta, quedando justo detrás del Capitán América. Su nombre aparece en diferentes comunidades, siendo así un nexo entre varios universos (como, por ejemplo, el de los Cuatro Fantásticos). Por otra parte, y como curiosidad, Tony Stark tiene unos cincuenta alias, lo que quiere decir que existen al menos cincuenta grupos de personas diferentes que lo conocen de cada una de esas formas.

A continuación, Wolverine aparece en un total de mil ciento veintitrés cómics. Al igual que Capitán América, este personaje une comunidades relacionándose con personajes que no pertenecen a la suya.

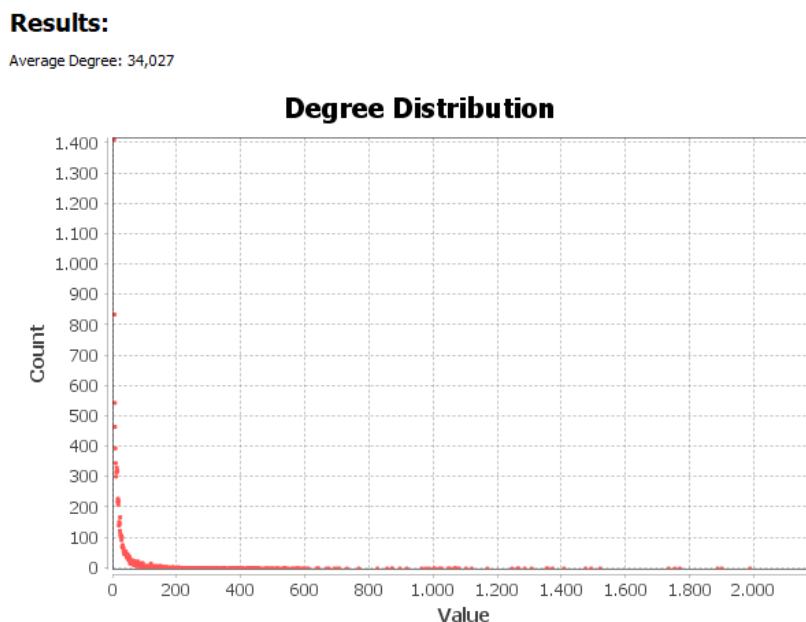


Fig 9

Como podemos observar en la tabla, el grado medio es de treinta y cuatro. ¿Por qué? Anteriormente hemos analizado nodos que cuentan con más de mil conexiones. Esto quiere decir que existen muchos más nodos pequeños que hacen que la media baje hasta ese punto. Este es el caso de personajes secundarios que tan solo aparecen en uno o dos cómics y similares.

2.3 Centralidades

2.3.1 Centralidad de intermediación

La centralidad de intermediación es una medida que sirve para observar qué nodos sirven como puente entre otros. Es decir, qué nodos conectan dos personajes diferentes que de por sí no se conocen.

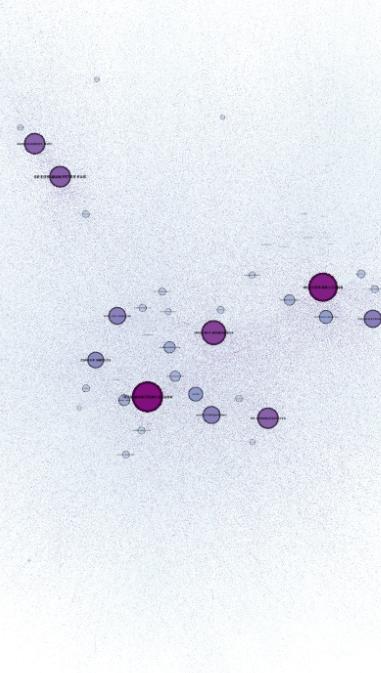


Fig 10: Vista general de la red por centralidad de intermediación

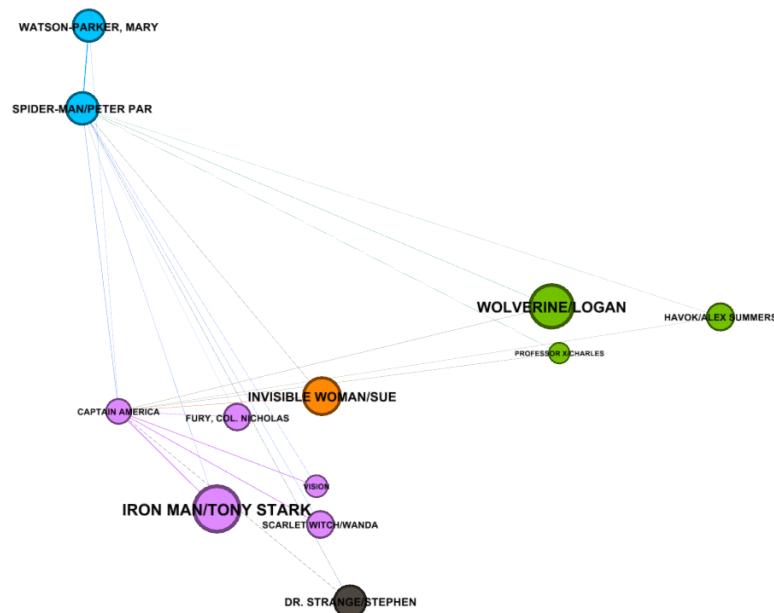


Fig 11: Nodos con un valor a partir de 0.02

```

aux = sorted(btwnCent.items(), key=operator.itemgetter(1), reverse=True)[0:10]
for x in aux:
    print(G.nodes[x[0]]["label"], "= ", x[1], " centralidad de intermediación")

IRON MAN/TONY STARK = 0.06734558824881197 centralidad de intermediación
WOLVERINE/LOGAN = 0.06179017101313348 centralidad de intermediación
INVISIBLE WOMAN/SUE = 0.055014508270724384 centralidad de intermediación
SPIDER-MAN/PETER PAR = 0.051093932898699505 centralidad de intermediación
DR. STRANGE/STEPHEN = 0.04693775558860187 centralidad de intermediación
WATSON-PARKER, MARY = 0.046006234007496705 centralidad de intermediación
FURY, COL. NICHOLAS = 0.042633076680272414 centralidad de intermediación
SCARLET WITCH/WANDA = 0.041805474056975316 centralidad de intermediación
HAVOK/ALEX SUMMERS = 0.040344771407915626 centralidad de intermediación
CAPTAIN AMERICA = 0.03870756731226559 centralidad de intermediación

```

Fig 12

Como podemos observar en la imagen, los nodos con mayor centralidad de intermediación son, más o menos, también aquellos que tenían mayor grado. Esto es porque, al tener tantas conexiones, es mucho más probable que ellos sean el puente entre dos personajes más aislados.

La centralidad de intermediación evalúa en cuántos caminos aparece un nodo a la hora de unir otros dos.

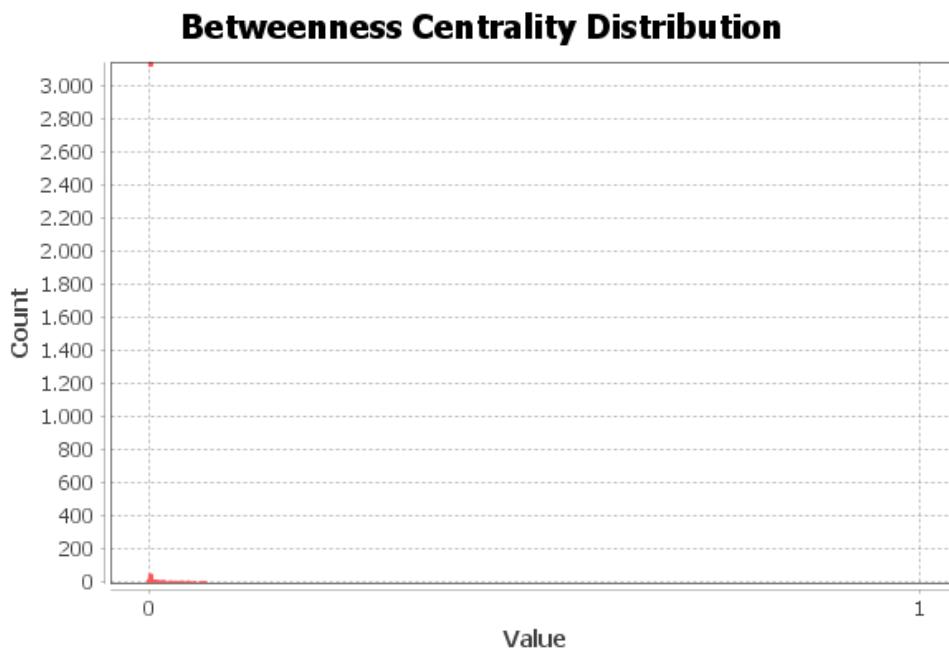


Fig 13

2.3.2 Centralidad de cercanía

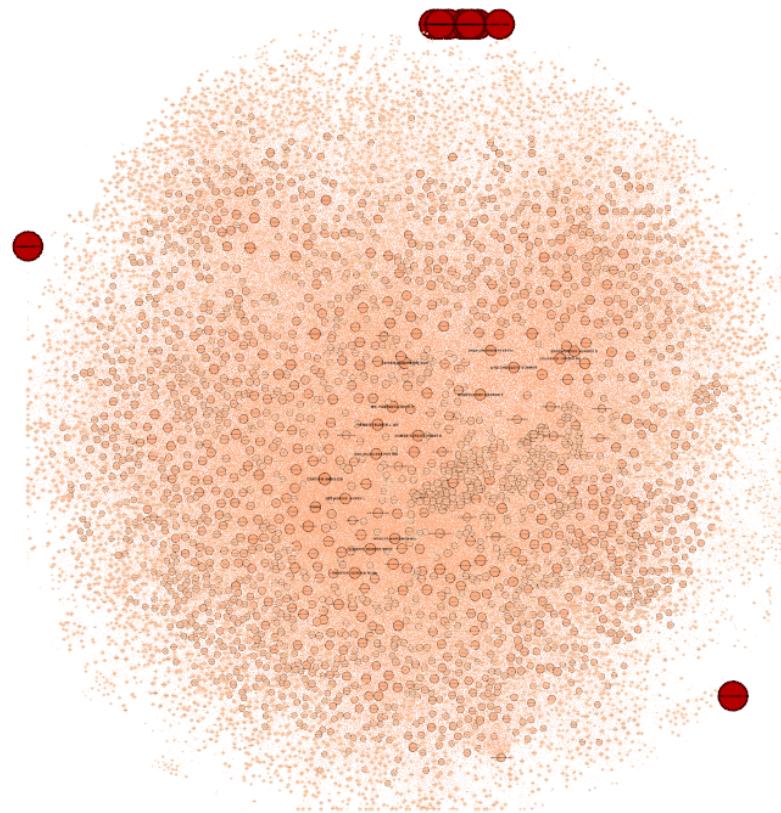


Fig 14

La centralidad de cercanía es el máximo camino mínimo que puede recorrer un nodo, es decir, es la distancia más pequeña que tendría que recorrer para llegar al nodo más lejano. En la imagen superior podemos observar que, por lo general, las centralidades de cercanía son bastante pequeñas. Esto es porque los nodos están muy bien conectados entre ellos. Por otra parte, los nodos grandes y rojos representan personajes muy aislados de Marvel, los cuales no tienen demasiadas conexiones y que tienen que recorrer grandes caminos para llegar a otros puntos de la red.

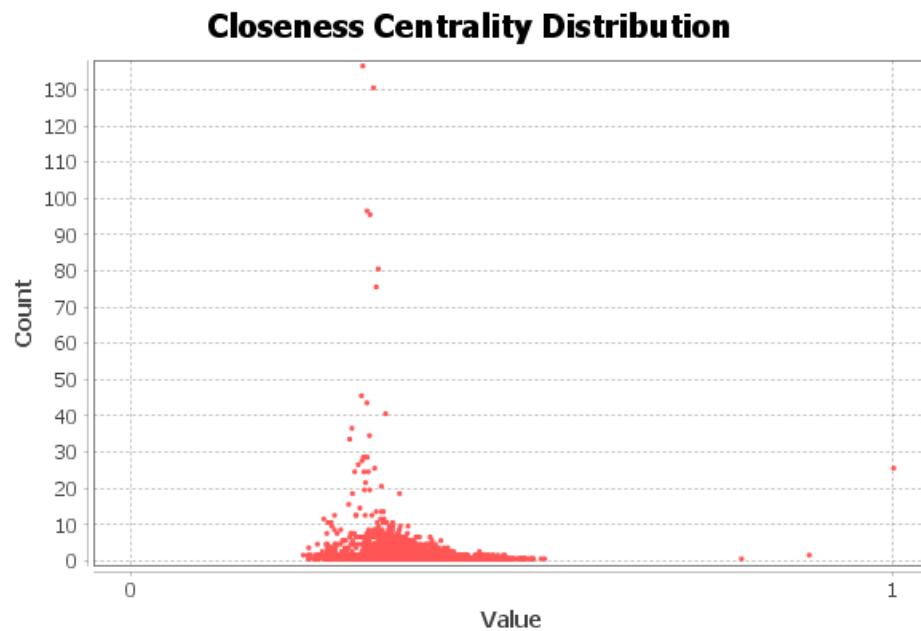


Fig 15

2.4 HITS

```
aux = sorted(hits[0].items(), key=operator.itemgetter(1), reverse=True)[0:10]
for x in aux:
    print(G.nodes[x[0]]["label"], "=" ,x[1], "Hits")
```

CAPTAIN AMERICA = 0.01385525714087428 Hits
 INVISIBLE WOMAN/SUE = 0.012592623462237693 Hits
 THING/BENJAMIN J. GR = 0.012383237118130826 Hits
 HUMAN TORCH/JOHNNY S = 0.012060936938322244 Hits
 MR. FANTASTIC/REED R = 0.011889278258958467 Hits
 IRON MAN/TONY STARK = 0.011069636393192849 Hits
 THOR/DR. DONALD BLAK = 0.010058669989142436 Hits
 SCARLET WITCH/WANDA = 0.00996971908651309 Hits
 VISION = 0.009724857728888108 Hits
 WASP/JANET VAN DYNE = 0.009512837447923545 Hits

Fig 16

Authority

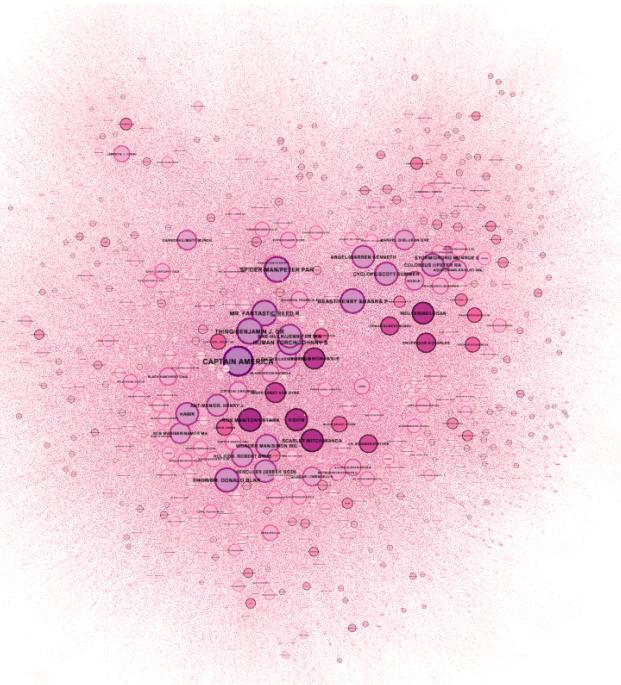


Fig 17: Vista general de la red con ranking de Hits- Authority

Hubs

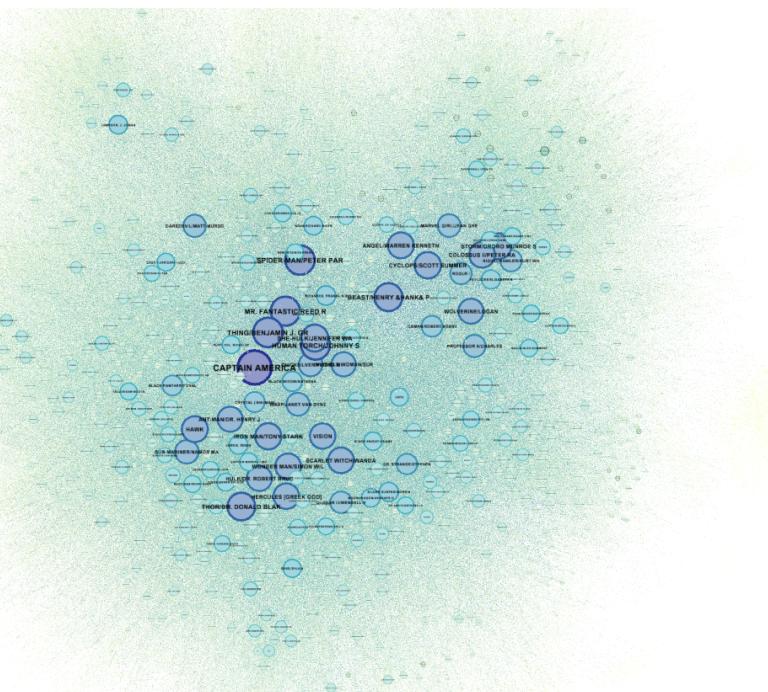


Fig 18: Vista general de la red con ranking de Hits- Hubs

El valor de HITs, obtenido mediante el de Hubs y Authority, vuelve a coincidir mayormente con personajes principales. Al igual que con pagerank, la relevancia de personajes principales se justifica con una alta relevancia de estos, junto con un número alto de conexiones con otros nodos relevantes, por lo que las conclusiones a sacar son esencialmente las mismas.

2.5 PageRank

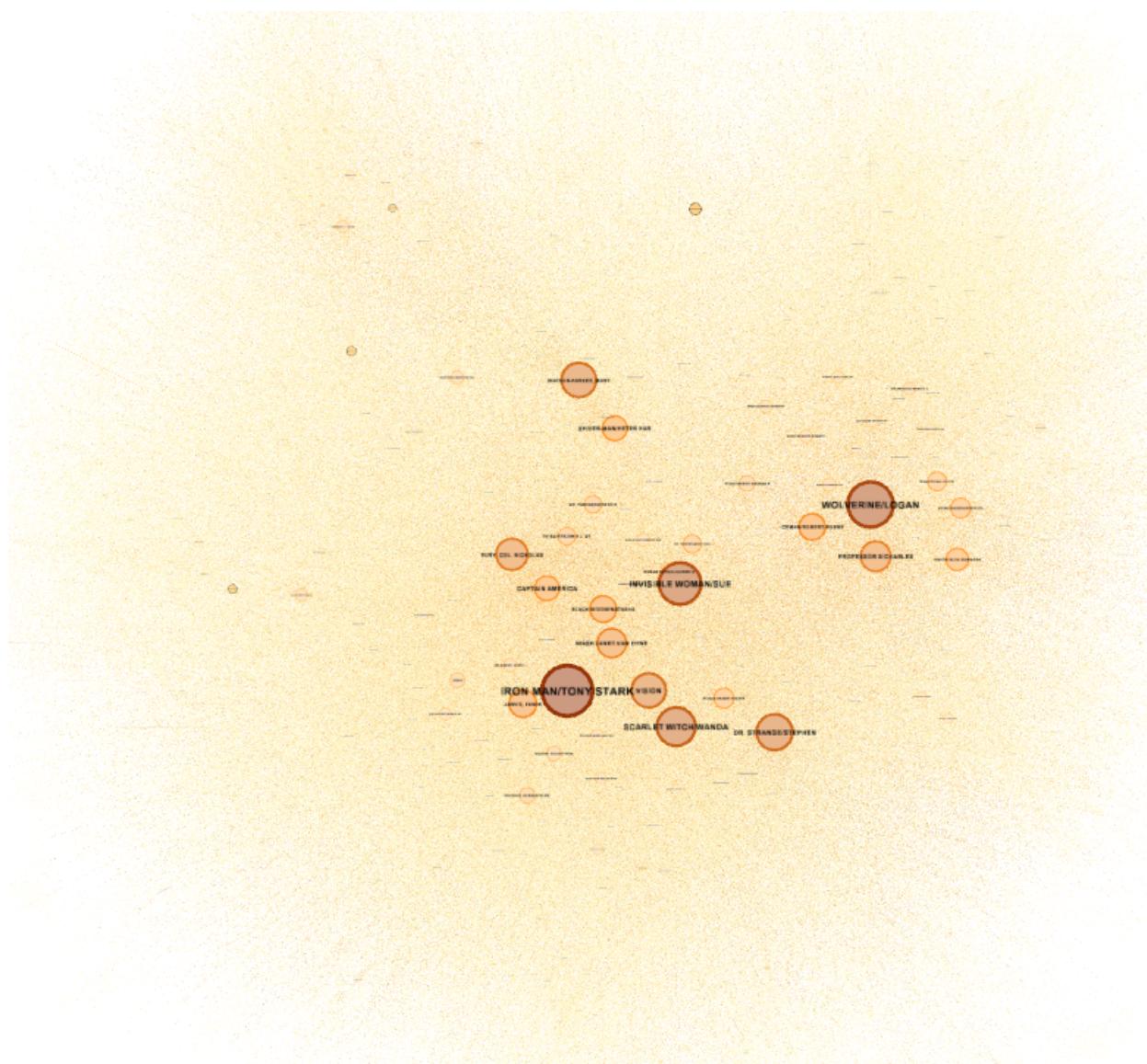


Fig 19

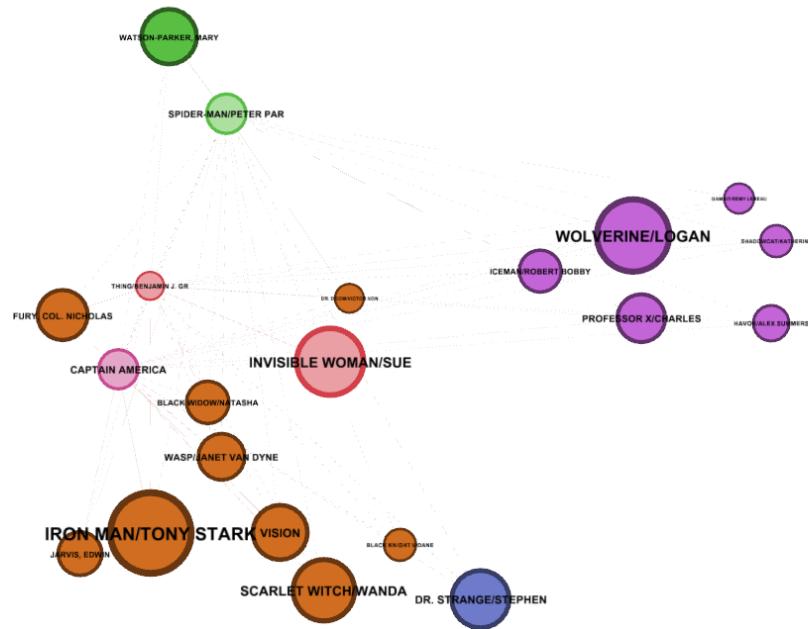


Fig 20: Nodos de la red con un valor de PageRank a partir de 0.03

El PageRank de la red determina quiénes son los nodos más importantes en base a la cantidad de aristas y la importancia de los nodos conectados por ellas.

En las capturas podemos observar que, como era de esperar por su importancia en el MCU, los personajes principales de las distintas sagas son aquellos con mayor importancia en cuanto a la métrica dada su alto grado.

Por ejemplo nodos como el de Tony Stark está directamente relacionado con muchos personajes de su propio universo, y algunos de ellos están a su vez conectados con nodos importantes como el de Thor o Capitán América, convirtiéndose así también en nodos importantes, y aumentando la posición en el pagerank de Tony Stark.

2.6 Clustering

Con el coeficiente de clustering nos topamos con un problema: Gephi no era capaz de analizar bien los coeficientes de clustering de cada nodo, ya que le asignaba a nodos como Iron Man 0 de coeficiente de clustering, cuando esta métrica indica cómo de bien conectada está a otros nodos, y este personaje es de los más centrales. Es por esto que nos salía un valor erróneo de 0.528 de clustering medio.

Results:

Average Clustering Coefficient: 0,528
Total triangles: 2794309
The Average Clustering Coefficient is the mean value of individual coefficients.

Clustering Coefficient Distribution

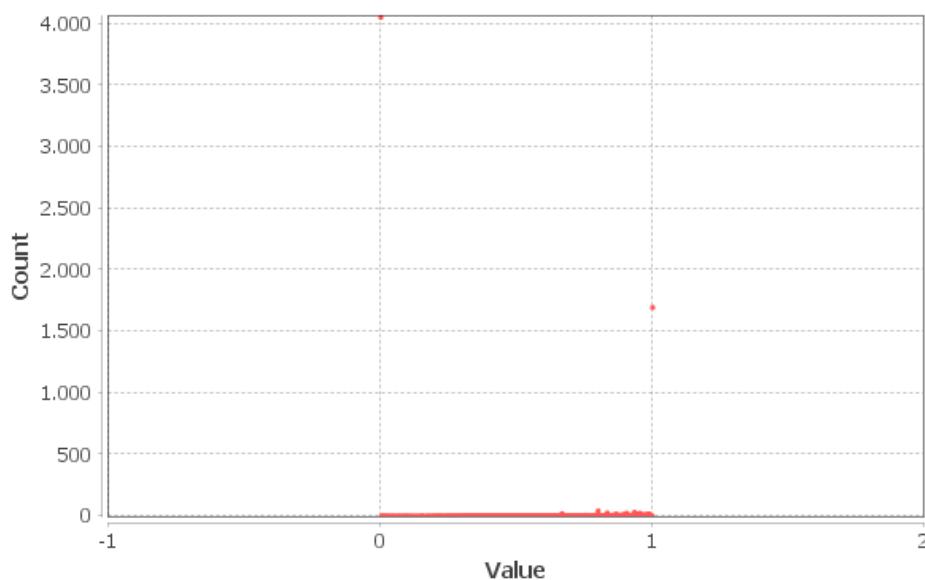


Fig 21

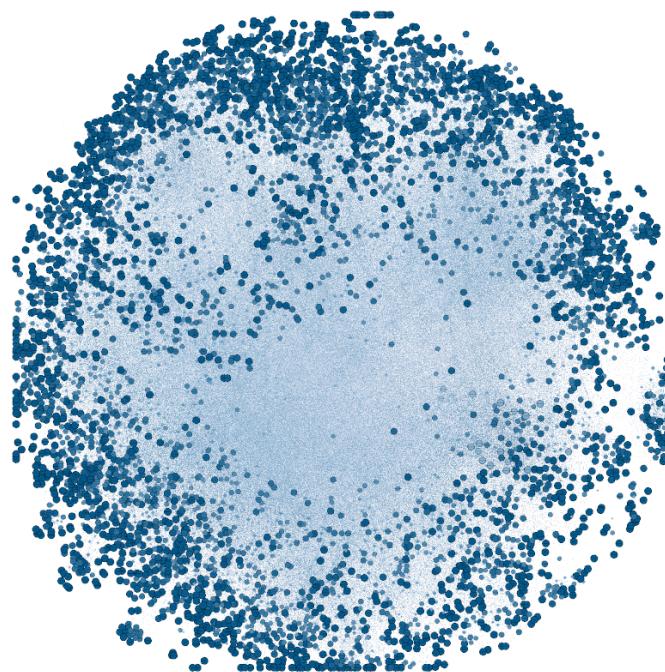


Fig 22: Vista general (errónea) de Gephi por clustering

Como podemos ver, por este problema los nodos más periféricos tienen mayor valor de clustering.

```
G_cluster = nx.average_clustering(G, count_zeros=False)  
  
print("Clustering average:", G_cluster)  
Clustering average: 0.7486268944775251
```

Fig 23

Con NetworkX logramos calcular que el clustering medio de la red realmente es 0.74.

3. Análisis de comunidades

Esta red presenta muchas comunidades. Sin embargo, nos vamos a centrar en las más grandes para su análisis. Estas serían los X-Men, los Vengadores, Hulk y Dr. Strange y el Spider-Verse.

3.1 X-Men

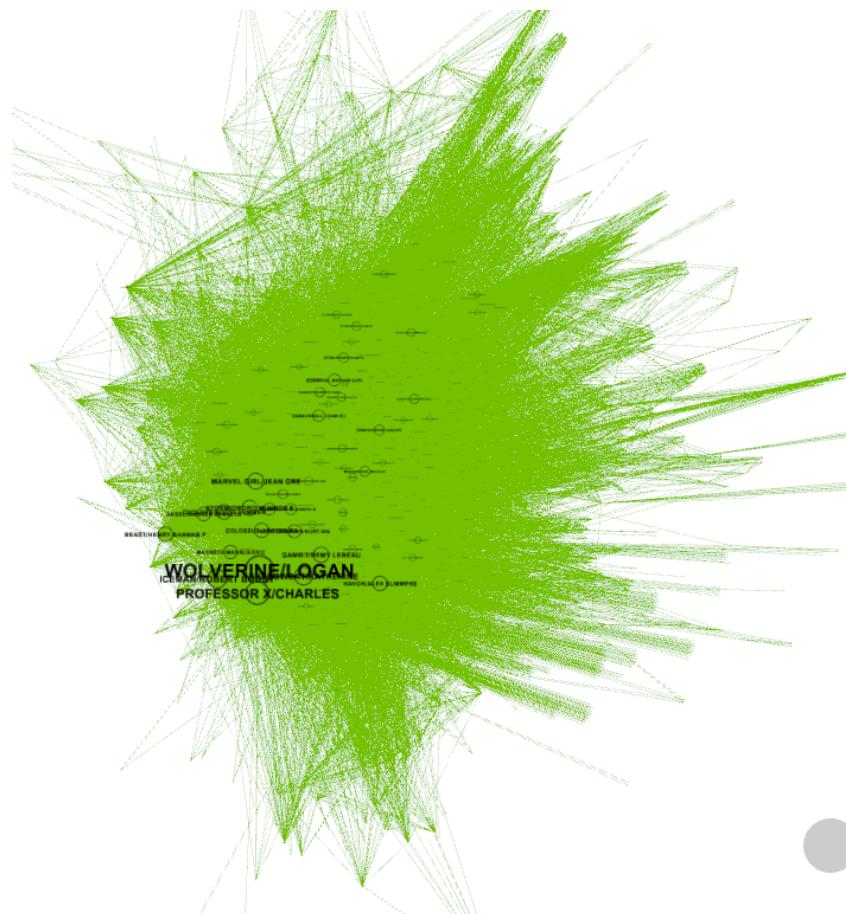


Fig 24

Nodos: 2239 (21,39% visible)

Aristas: 29750 (16,7% visible)

Grafo no dirigido

Fig 25

Como podemos ver, esta red se caracteriza por una infinidad de nodos muy pequeños. Después, sobresalen algunos principales como el de Wolverine o el del Profesor X, personajes muy destacados dentro de este universo y quienes sirven de puente entre los personajes de este mundo y también con los de otras comunidades.

```
[ ] #X-men
list5[0:10]

['WOLVERINE/LOGAN',
 'PROFESSOR X/CHARLES',
 'ICEMAN/ROBERT BOBBY',
 'BEAST/HENRY &HANK& P',
 'HAVOK/ALEX SUMMERS',
 'CYCLOPS/SCOTT SUMMER',
 'ANGEL/WARREN KENNETH',
 'STORM/ORORO MUNROE S',
 'COLOSSUS II/PETER RA',
 'SHADOWCAT/KATHERINE']
```

Fig 26: 10 nodos con mayor grado de la comunidad X-men

En NetworkX podemos visualizar los 10 nodos con mayor grado dentro de la comunidad. El primero, como hemos dicho, es Wolverine. Este es el personaje de esta saga con más cómics y películas; en la wiki de Marvel cuenta con 3186 apariciones (solo contamos con Wolverine de la Tierra 616, la realidad principal de este multiverso). Siguiéndole está el Profesor X, el fundador de los X-Men. A continuación podemos ver otros personajes, todos de ellos X-men. Iceman, Beast, Cíclope, por ejemplo, con miembros fundadores de X-men. Havok y Storm, por ejemplo, se asocian a menudo con los X-men (sus primera apariciones fueron en cómics de los X-men). Por todo esto vemos que esta comunidad está bien analizada. La red cuenta con 2239 nodos, de los cuales tan solo unos pocos tienen un grado elevado, al igual que el resto de comunidades que iremos viendo.

3.2 Los Vengadores

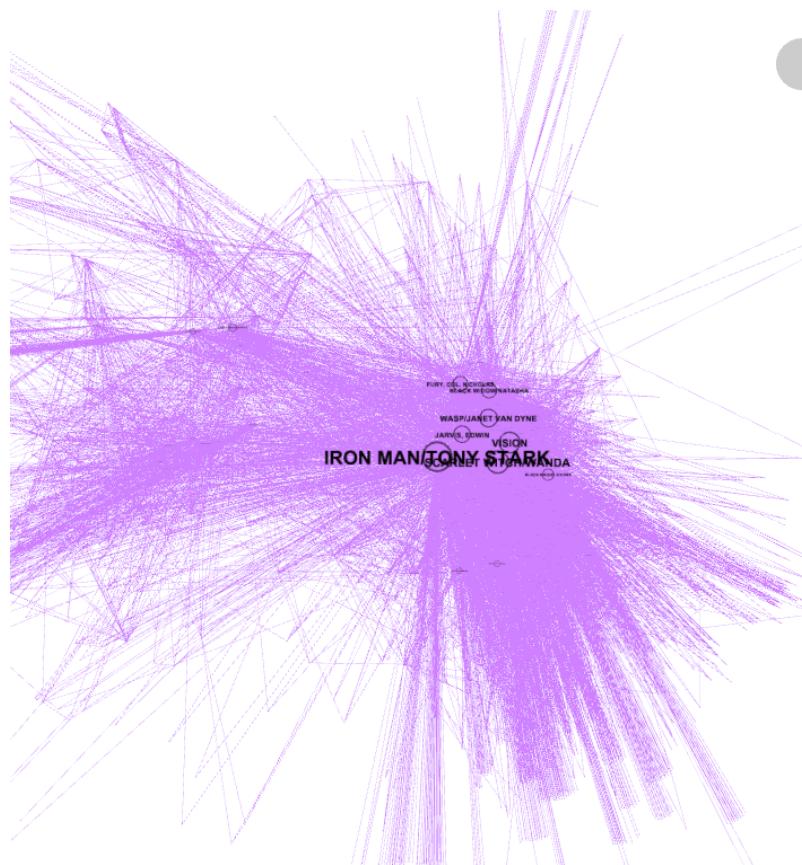


Fig 27

Nodos: 2505 (23,93% visible)
Aristas: 20662 (11,6% visible)
Grafo no dirigido

Fig 28

```
[ ] #Avengers
list3[0:10]

['IRON MAN/TONY STARK',
 'SCARLET WITCH/WANDA',
 'VISION',
 'INVISIBLE WOMAN/SUE',
 'WASP/JANET VAN DYNE',
 'JARVIS, EDWIN',
 'BLACK WIDOW/NATASHA',
 'BLACK KNIGHT V/DANE',
 'DR. DOOM/VICTOR VON',
 'BEETLE/ABNER RONALD']
```

Fig 29: 10 nodos con mayor grado de la comunidad Vengadores

Los Vengadores es también una comunidad bastante céntrica del universo de Marvel. En ella destacan muy pocos personajes a pesar de contar con dos mil quinientos nodos. Estos, a su vez, tienen un grado muy alto al tratarse de protagonistas dentro de este universo que sirven de puente entre comunidades.

Si pensamos únicamente en las películas de los Vengadores del MCU (Marvel Cinematic Universe), que son más conocidas que los cómics, ya vemos a personajes de los Vengadores que nos suenan: Iron Man, la Bruja Escarlata, Visión, la Viuda Negra, La Avispa...

Aparte de ellos, vemos que está Jarvis, la inteligencia artificial que controla el traje de Iron man, por lo que tiene sentido que esté en esta comunidad.

Black Knight apareció por primera vez en un cómic de los Vengadores.

Por otro lado, hay personajes como Dr. Doom y La mujer invisible que de primeras podemos pensar que solo estarían en la comunidad de los 4 Fantásticos (aparecen principalmente en esta saga), pero en la página de la wiki de Marvel de Dr. Doom, podemos leer sobre este encuentro en *Counter-earth*:

When [Franklin Richards](#) was kidnapped by [Onslaught](#), Doom (who saw this as an opportunity to steal the being's powers) joined the Fantastic Four, the [Avengers](#) and the [X-Men](#) to battle Onslaught in [Central Park](#). An enraged [Hulk](#) was able to crack open Onslaught's shell.

Fig 30

Vemos que eventos como este hacen que personajes importantes como los que se acaban de mencionar existan en esta comunidad.

3.3 Spider-Verse

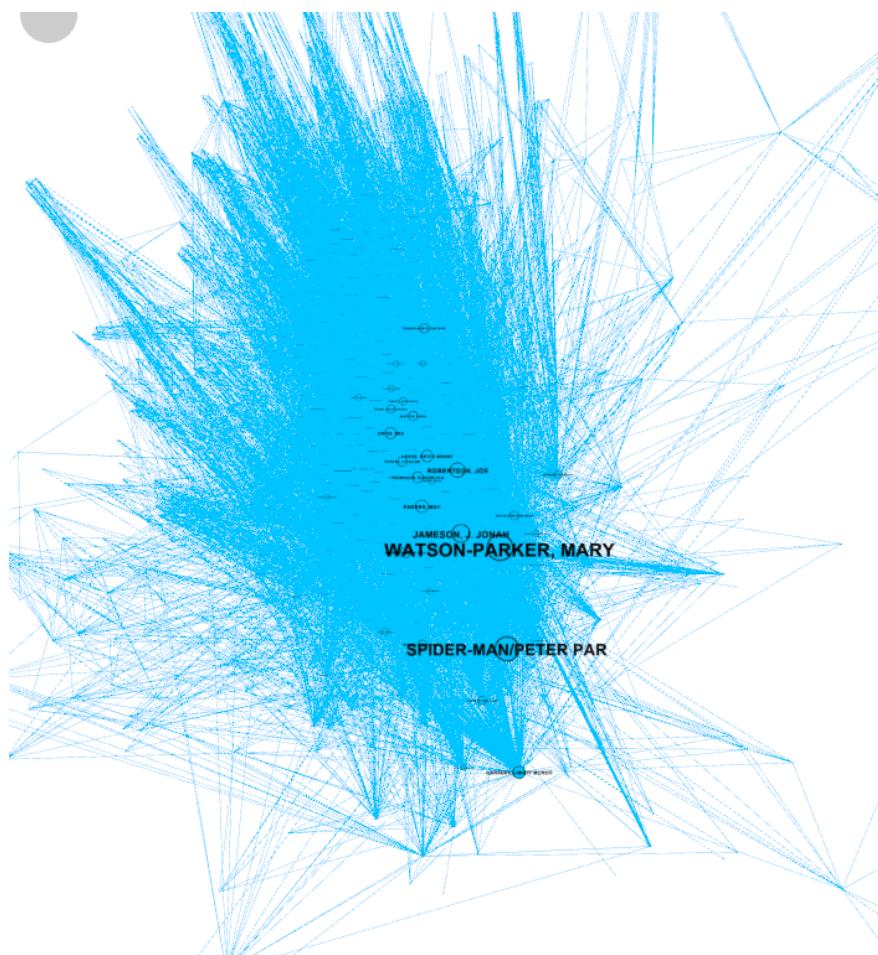


Fig 31

Nodos: 1326 (12,67% visible)

Aristas: 12691 (7,13% visible)

Grafo no dirigido

Fig 32

```
[39] #Spider-verse
list2[0:10]

['SPIDER-MAN/PETER PAR',
 'WATSON-PARKER, MARY',
 'DAREDEVIL/MATT MURDO',
 'JAMESON, J. JONAH',
 'ROBERTSON, JOE',
 'NELSON, FRANKLIN FOG',
 'PARKER, MAY',
 'PUNISHER II/FRANK CA',
 'KINGPIN/WILSON FISK',
 'JAMESON, COL. JOHN']
```

Fig 33: 10 nodos con mayor grado de la comunidad Spider-verse

Como podemos observar en la imagen, muy pocos nodos destacan en esta comunidad. El universo de Spiderman está muy centrado en su protagonista, Spider-man, por lo que la historia gira en torno a él y las apariciones de otros

personajes en sus aventuras. El segundo nodo más importante es Mary Jane Watson, el interés amoroso principal de Spiderman. Otro nodo notable es J.J. Jameson, el carismático periodista que está siempre atento a la vida de éste superhéroe, por lo que tiene sentido que esté en esta comunidad. También vemos a la tía de Peter (Spiderman), May, o villanos a los que Spiderman se ha enfrentado, como Kingpin, u otros superhéroes con los que ha colaborado, como Daredevil.

Becoming Daredevil

Though some of his early enemies were considered laughable like ([Matador](#),^[35] [Stilt Man](#),^[36] and [Leap Frog](#)),^[9] Daredevil gathered an impressive and dangerous rogues gallery. [Electro](#) was the first costumed villain Daredevil faced.^[37] Also, in one of his first encounters, he clashed with his long-time nemesis the [Owl](#).^[38] In his first team-up with another superhero, he assisted [Spider-Man](#) in defeating the [Circus of Crime](#).^[39] His next opponent, the [Purple](#)

Fig 34: Página de la wiki de Marvel de Daredevil

3.4 Dr. Strange y Hulk

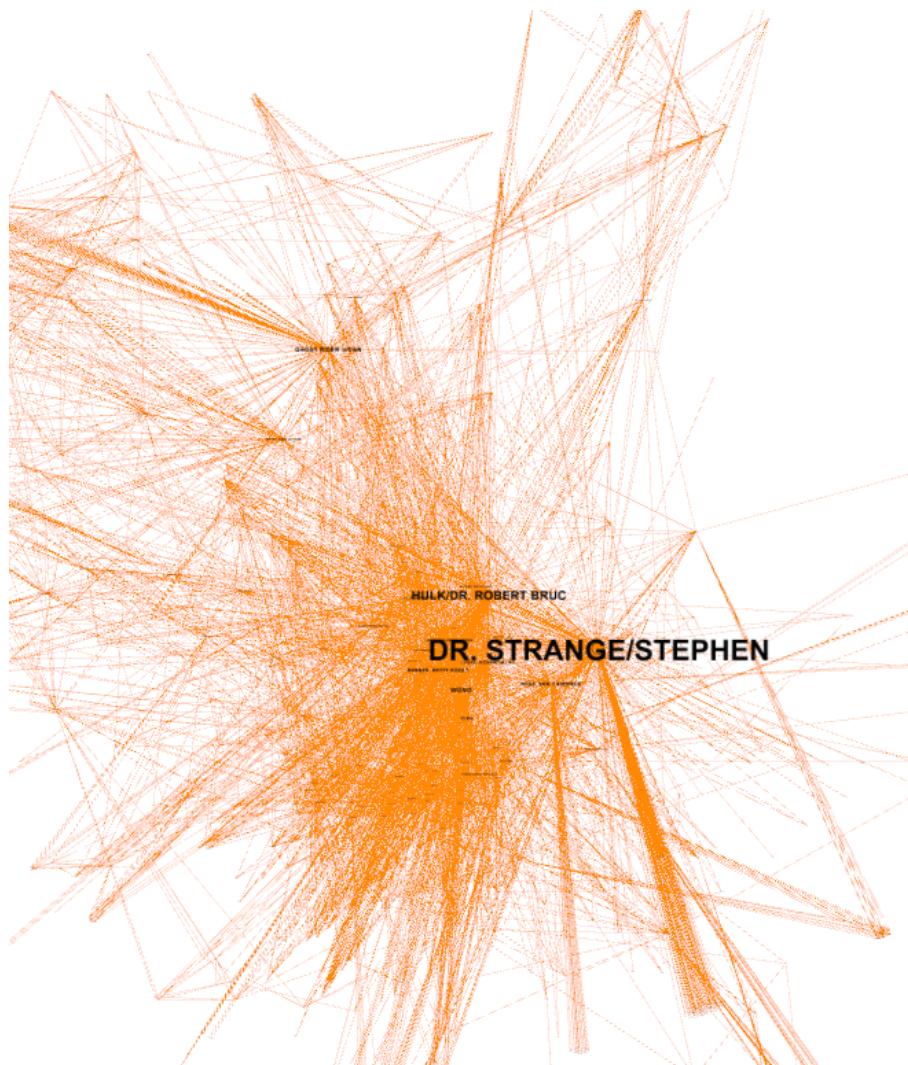


Fig 35

Nodos: 987 (9,43% visible)
Aristas: 4884 (2,74% visible)
Grafo no dirigido

Fig 36

```
[ ] #Hulk and Dr. Strange
list4[0:10]

['DR. STRANGE/STEPHEN',
 'HULK/DR. ROBERT BRUC',
 'JONES, RICHARD MILHO',
 'WONG',
 'ROSS, GEN. THADDEUS',
 'GHOST RIDER III/DAN',
 'NORRISS, SISTER BARB',
 'HELLCAT/PATSY WALKER',
 'BANNER, BETTY ROSS T',
 'DOC SAMSON/DR. LEONA']
```

Fig 37: 10 nodos con mayor grado de la comunidad Hulk y Dr. Strange

Con el Dr. Strange y Hulk se forma una nueva comunidad, la cual contiene casi mil nodos. Para nuestra sorpresa, estos dos personajes están bastante relacionados. El Dr. Strange y Hulk son protagonistas en el MCU, y comparten ciertas aventuras que nos llevan a nosotros a visualizar esta red, probablemente más por sus conocidos en común que por su propia relación. La importancia del resto de nodos es bastante despreciable al tratarse de personajes secundarios sin muchas interacciones ni protagonismo.

3.5 Fantastic Four

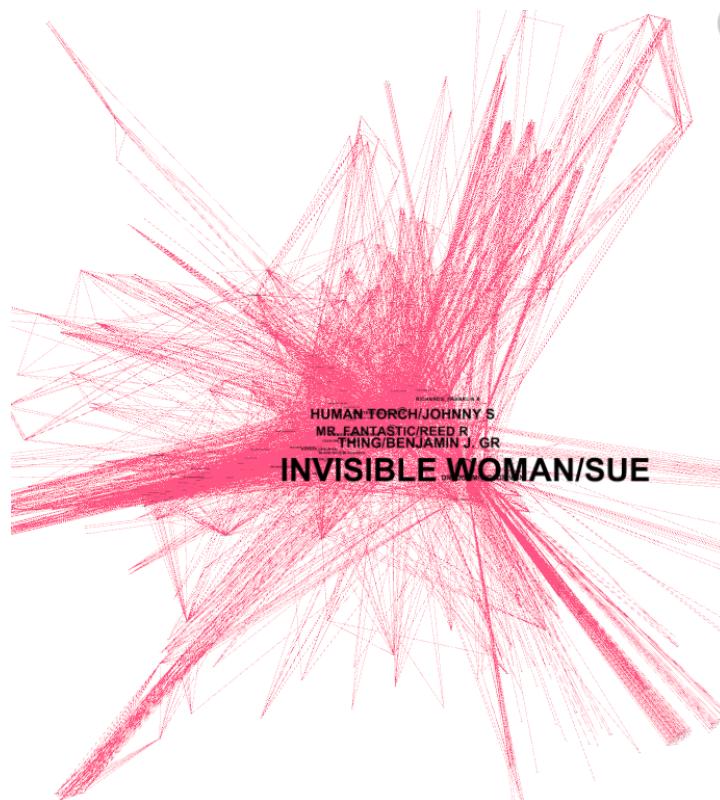


Fig 38

Nodos: 760 (7,26% visible)

Aristas: 4777 (2,68% visible)

Grafo no dirigido

Fig 39

```
[ ] #Fantastic 4
list6[0:10]

['THING/BENJAMIN J. GR',
 'MR. FANTASTIC/REED R',
 'HUMAN TORCH/JOHNNY S',
 'SILVER SURFER/NORRIN',
 'RICHARDS, FRANKLIN B',
 'UATU',
 'MEDUSA/MEDUSALITH AM',
 'BLACK BOLT/BLACKANTO',
 'GORGON [INHUMAN]',
 'KARNAK [INHUMAN]']
```

Fig 40: 10 nodos con mayor grado de la comunidad Fantastic 4

En la red de los Cuatro Fantásticos, como cabe esperar, destacan los cuatro personajes principales (La mujer invisible, La cosa, Mr. fantástico y la Antorcha humana).

Además de ellos, vemos personajes cuyas primeras apariciones han sido en cómics de esta saga, como Silver Surfer, Uatu o Black Bolt.

3.6 Comportamientos especiales

3.6.1 Mutantes X

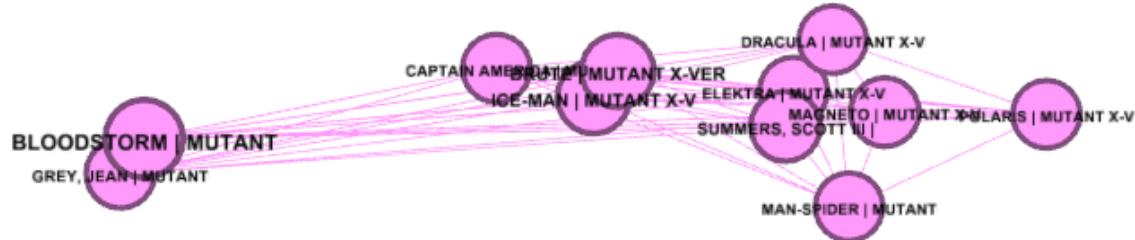


Fig 41: Una pequeña muestra de los nodos de la comunidad Mutant X

Cabe mencionar también una comunidad que, aunque no sea tan importante como las anteriormente mencionadas, sí que resulta bastante destacable. Es el caso de los Mutantes X, una red que se encuentra muy separada de todas las demás a excepción de unas pocas conexiones, sobre todo con la comunidad de los X-Men. En concreto, en la Figura 1 es la comunidad rosa que aparece apartada a la derecha.

Al tratarse de una comunidad pequeña y bastante aislada, podemos observar que el grado de los nodos es bastante similar al estar todos relacionados con todos.

3.6.2 Capitán América

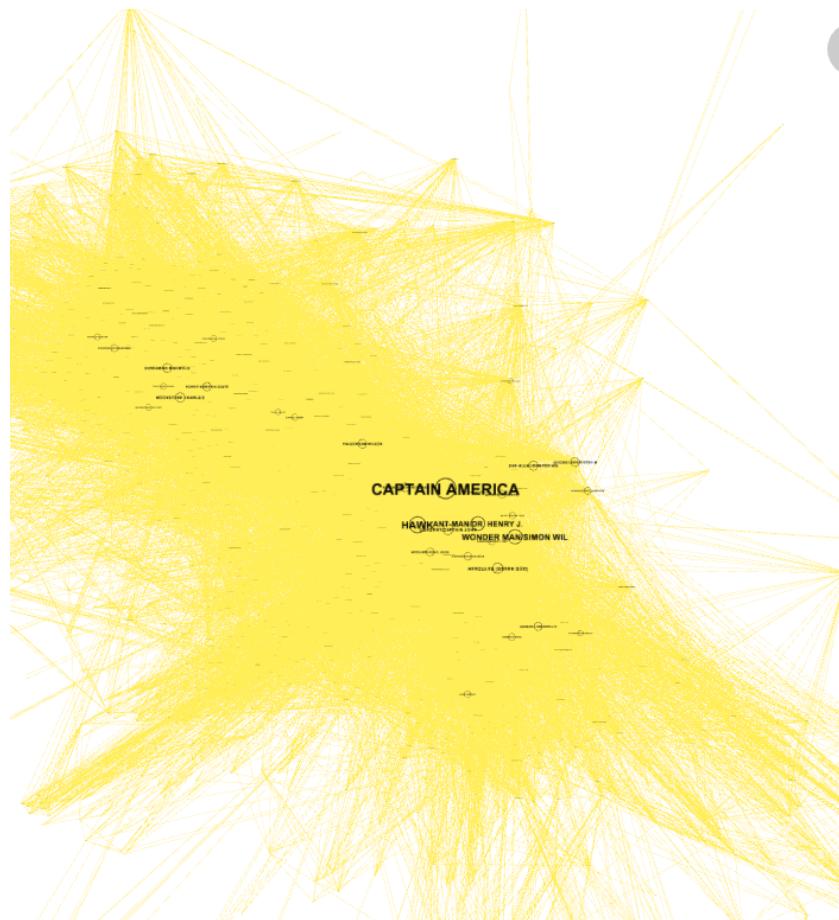


Fig 42: la comunidad “Capitán América”

```
#Capitan America
list1[0:10]

['CAPTAIN AMERICA',
 'HANK',
 'ANT-MAN/DR. HENRY J.',
 'SHE-HULK/JENNIFER WA',
 'WONDER MAN/SIMON WIL',
 'HERCULES [GREEK GOD]',
 'QUICKSILVER/PIETRO M',
 'QUASAR III/WENDELL V',
 "BLACK PANTHER/T'CHAL",
 'CRYSTAL [INHUMAN]']
```

Fig 43: 10 nodos con mayor grado de la comunidad “Capitán América”

Este personaje presenta un comportamiento en la red. Algunas de las imágenes anteriores donde veíamos los nodos más importantes según distintas métricas se hicieron en distintos momentos (pero no afecta las métricas, únicamente la división por comunidades), y por eso podemos ver, por ejemplo, que en la

imagen 7 y en la imagen 20 Capitán América parece estar en comunidades distintas; a veces junto a los Vengadores, a veces cuenta con su propia comunidad.

Esto es porque al tratarse de un nodo con un grado tan elevado, este puede pertenecer a unas comunidades o a otras dependiendo de la distribución de nodos en Gephi. Capitán América, de nuevo, es uno de los mayores protagonistas de Marvel, lo cual queda reflejado en su elevado número de conexiones con otros nodos de la red. Además, ocurre algo parecido con la comunidad que hemos llamado “Hulk & Dr. Strange”, como podemos ver, de los 10 nodos más importantes de esta comunidad, 7 de ellos eran conocidos por los miembros de este grupo (Capitán América, Hawkeye, Ant-man, She-hulk, Hercules, Quicksilver y Black Panther), algo que no nos ha ocurrido con ninguna otra comunidad; por tanto tiene sentido que esto, en realidad, sea una comunidad de personajes que son tan importantes, y tienen tantas conexiones, que forman su propia comunidad.

4. Bibliografía

Información de los personajes de Marvel:

<https://marvel.fandom.com/es/wiki/Categor%C3%ADa:Apariciones>

Red social del proyecto:

<https://github.com/gephi/gephi/wiki/Datasets>