



**Universidad Zaragoza**

TRABAJO FIN DE GRADO:

**CONTROL DE DIFUSIÓN DE RUMORES EN REDES  
COMPLEJAS**

Pablo Ibáñez Freire

---

Directores:  
Jesús Gómez Gardes  
Luis Mario Floría Peralta

CURSO 2014-2015



# Índice general

<b>1. Introducción.</b>	<b>1</b>
1.1. Redes complejas. . . . .	1
1.1.1. Redes complejas. Formalismo. . . . .	2
1.2. Modelos epidemiológicos. . . . .	4
1.2.1. Modelo SIR. . . . .	5
<b>2. Modelo para la propagación de rumores.</b>	<b>7</b>
2.1. Redes empleadas. . . . .	7
2.1.1. Red de Erdős-Rényi. . . . .	7
2.1.2. Red libre de escala. . . . .	9
2.2. Mecanismo de propagación. . . . .	12
2.3. Control de la difusión del rumor. . . . .	14
<b>3. Resolución del modelo.</b>	<b>15</b>
<b>4. Resultados.</b>	<b>19</b>
<b>5. Conclusiones.</b>	<b>23</b>



# Capítulo 1

## Introducción.

El objetivo de este trabajo es dar una aproximación a los procesos de “spreading” en concreto en el marco de la propagación de un rumor. Entendemos por proceso de “spreading” aquellos procesos que llevan asociados la propagación de un agente mediante la interacción entre varios individuos, como se mostrará, este agente podrá ser desde un virus hasta una idea. Como se aprecia a primera vista, dos son los elementos claves de estos procesos “la propagación de un agente” y “la interacción entre varios individuos”. Si nos disponemos a afrontar el problema de entender y posteriormente desarrollar un modelo matemático que recoja las ideas esenciales del proceso y que además nos permita hacer predicciones sobre el comportamiento del sistema. Es conveniente en primera instancia trabajar con estos dos elementos por separado y posteriormente llevarlos a un marco común. En resumidas cuentas, nuestro problema inicial, ahora queda dividido en dos problemas más sencillos, determinar cómo se lleva a cabo el proceso de propagación del agente y de qué forma se lleva a cabo la interacción entre los individuos donde este proceso sucede.

Siempre con la idea en mente de que en última instancia nos interesa cómo se comporta un rumor, elegiremos como punto de partida de nuestro estudio los modelos epidemiológicos, que nos ayudarán a entender el proceso de propagación y las redes complejas que nos proporcionan información de cómo están configuradas las estructuras que emergen fruto de la interacción entre individuos.

En las siguientes secciones expondremos brevemente estos elementos, explicando cuál ha sido la razón por la que han sido elegidos para nuestro empeño y dando una breve descripción de los mismos.

### 1.1. Redes complejas.

En los ámbitos en los que trabajamos, la propagación de un virus o bien de un rumor, los agentes que son susceptibles de ser infectados o de creer un rumor son seres humanos. Entonces necesitamos encontrar una estructura que encapsule las relaciones que se establecen entre las diferentes personas.

Las relaciones que se establecen entre los seres humanos son complicadas de estudiar. Esto se debe parcialmente a que estas relaciones se llevan a cabo entre un gran número de individuos pero existen además otras consideraciones que hacen difícil trabajar con este tipo de relaciones. Por ejemplo, si consideramos un gran número de individuos, un individuo elegido al azar no interactuará con el mismo número de personas que otro individuo elegido también al azar. Si a esto se le suma el hecho de que las personas no interactúan de forma local (pensemos por ejemplo tanto en los medios transporte, que permiten a las personas recorrer grandes distancias en cortos periodos de

tiempo, como también en los modernos métodos de comunicación, que hacen que personas que se encuentran en lugares alejados puedan estar en contacto de manera casi inmediata) tenemos como resultado la imposibilidad de utilizar métodos que se empleaban en otras disciplinas para describir sistemas con muchos grados de libertad, como las aproximaciones de campo medio o la hipótesis del “homogenous mixing” empleada en los modelos epidemiológicos.

Las estructuras que recogen esta información y nos permiten trabajar con ella son las denominadas “**redes complejas**”. Una red es un grafo (red), la palabra compleja es debida a que estos grafos presentan una estructura no trivial, a diferencia de lo que sería por ejemplo el grafo asociado con un retículo cristalino.

Las redes complejas son un importante instrumento en muchos campos, no sólo en los mencionados anteriormente. La razón de esto es que mediante las redes complejas podemos trabajar con sistemas en los que las relaciones entre sus elementos tienen tal complejidad que no se pueden tratar mediante los métodos tradicionales. Un campo donde las redes complejas han sido, y siguen siendo de gran utilidad, es el de la biología, debido a que en los procesos biológicos que constituyen la vida hay un gran número de elementos que interaccionan de manera intrincada, como los genes, las proteínas, las rutas metabólicas... [3][4]. Otro campo donde las redes están demostrando ser de gran utilidad es en el campo de la neurología [5][?], mediante, el uso de las redes neuronales se está logrando hacer grandes avances en nuestro entendimiento de cómo funciona el cerebro. También la teoría de redes se emplea ampliamente es en el campo de la economía [6], donde las complicadas relaciones que se establecen entre mercados pueden ser comprendidas mediante su uso. Estos son solo unos ejemplos, entre muchos otros, de los campos donde la teoría de redes se ha convertido en una parte indispensable que ha permitido hacer grandes avances.

### 1.1.1. Redes complejas. Formalismo.

Como se ha mencionado, una red compleja es un grafo que tiene ciertas propiedades. Por lo tanto, el formalismo que se emplea cuando se busca caracterizar y trabajar con este tipo de objetos es el formalismo de la teoría de grafos. De manera simplificada, podemos definir un grafo como un catálogo de componentes, llamados nodos, y las relaciones directas entre ellos, denominadas links. Hemos introducido entonces dos nuevos elementos, los nodos y los links, podemos también ahora introducir dos nuevos parámetros que nos ayudaran a dar una primera caracterización de la red:

- Número de nodos: lo denotaremos con la letra  $N$ . Frecuentemente  $N$  se conoce como el tamaño de la red. Los nodos se numeran con  $i=1,2,3,...,N$ .
- Número de links: lo denotaremos con la letra  $L$ , representa el número total de interacciones que se dan en la red. Para referirnos a un link lo haremos a través de los dos nodos que relaciona. Por ejemplo el link  $(3,5)$  conecta los nodos 3 y 5.

Los links pueden ser de dos tipos: no direccionados o direccionados. En un link dirigido (o direccionado) la relación que se establece entre los nodos que conecta no es bidireccional, en el sentido de que si, por ejemplo, 2 está conectado con 3, esto no significa que 3 esté conectado con 2. Un ejemplo de este tipo de redes sería el de la red WWW, donde que una página enlace con otra página, no significa que esta última tenga un enlace con la página precedente. En nuestro caso, como nos reduciremos al ámbito de las relaciones sociales, nos ocuparemos solo de redes no direccionadas. De forma sencilla se ve que cuando dos personas interaccionan entre ellas lo hacen de manera bidireccional.

Para trabajar con la red, necesitamos algún tipo de objeto que encapsule toda la información del grafo, a saber; el número de nodos y como estos están conectados. Por lo general, el elemento que se emplea es la denominada matriz de adyacencia que está constituida de la manera siguiente: se trata de una matriz  $N \times N$  en la que cada elemento  $(i,j)$  puede tomar dos valores:

- 0 si los nodos  $i$  y  $j$  no están conectados.
- 1 si los nodos  $i$  y  $j$  están conectados.

De esta definición se pueden extraer de forma inmediata ciertas características de la matriz de adyacencia: para el caso de una red no direccional, la matriz de adyacencia será simétrica y en el caso en el que no se considere que un nodo está conectado consigo mismo los elementos de la diagonal serán cero. Este tipo de matrices de adyacencia serán las empleadas en nuestro caso.

Introduzcamos ahora una serie de elementos asociados a cada red que nos permitirán obtener información valiosa sobre cómo dicha red está constituida. Definimos en primer lugar el grado  $k_i$  de un nodo:

$$k_i = \sum_j^N A_{ij}$$

donde  $A_{ij}$  es la matriz de adyacencia definida arriba. El grado de un nodo se corresponde con el número de links que éste tiene. En una red no direccional, el número total de links puede ser calculado mediante la expresión:

$$L = \frac{1}{2} \sum_{i=1}^N k_i.$$

Una importante propiedad de las redes es el grado medio o la conectividad media que para una red no direccional se define como:

$$\langle k \rangle = \frac{1}{N} \sum_{i=1}^N k_i = \frac{2L}{N}$$

Finalmente, podemos introducir la distribución de la conectividad, es decir, la probabilidad de que un nodo elegido al azar tenga conectividad  $k$ . Esta probabilidad viene dada por la expresión:

$$p_k = \frac{N_k}{N}$$

que como se observa esta sujeta a la relación de normalización:

$$\sum_{k=1}^{\infty} p_k = 1.$$

La distribución de la conectividad nos da valiosa información sobre cómo se organiza la red y sobre muchas de sus propiedades intrínsecas. Posteriormente esta herramienta será utilizada para discutir cuáles son las diferencias entre distintos tipos de redes.

## 1.2. Modelos epidemiológicos.

A pesar de que en un principio puede parecer que un rumor y agente infeccioso no tienen nada en común, si examinamos detenidamente su comportamiento llegaremos a la conclusión de que comparten ciertos aspectos, por lo tanto es conveniente examinar los modelos epidemiológicos, sobre los que se ha trabajado de manera extensa durante años y que nos servirán de punto de partida para desarrollar un modelo que represente el comportamiento de rumor.

Examinemos en primera instancia los elementos comunes que comparten los procesos de propagación de un virus y de un rumor. De manera sencilla, podríamos resumir el proceso de contagio de un virus de la manera siguiente; para que un individuo se infecte, es necesario que entre en contacto con otro individuo infectado y a su vez este nuevo individuo infectado se convierte en una nueva fuente de infección, que es capaz de infectar a otros individuos con cierta probabilidad. Examinado con un poco de detalle el comportamiento por el cual se propaga un rumor, llegamos a la conclusión de que se trata de un comportamiento muy similar a aquél por el cual un virus se contagia: un individuo que considera que el rumor es cierto lo compartirá, con cierta probabilidad, con los individuos con los que esté en contacto. Si alguno de ellos se cree este rumor tenderá a compartirlo también con otros individuos; es decir, el comportamiento de “contagio” en ambos casos es equivalente. Aunque, como se ha expuesto, el proceso de propagación es similar en ciertos aspectos, también tiene ciertas diferencias que no podemos obviar, sobre todo referidas al proceso por el cual un individuo deja de estar infectado, se recupera, o deja de considerar cierto un rumor. Por ejemplo, en el caso de una infección, con el paso del tiempo, por lo general, la propia acción del cuerpo hace que la infección desaparezca mientras que en principio, esto no sucede en el caso de un rumor; si se considera que cierta información es cierta y no hay nada que indique lo contrario, seguiremos considerando el rumor como cierto independiente del tiempo que transcurra. Por otro lado, para que un individuo deje de considerar cierto un rumor es necesario la mediación con otro individuo que sepa que no es cierto. Esto en el caso de un proceso infeccioso originado por un virus no sucede; el hecho de interaccionar con un individuo sano no da lugar a la recuperación de la infección. Por contra, otro elemento importante en común es que una vez que un individuo enfermo ha superado la infección, éste se vuelve inmune a la misma, del mismo modo que una persona que conoce que cierta información es falsa no volverá a considerarla como verdadera.

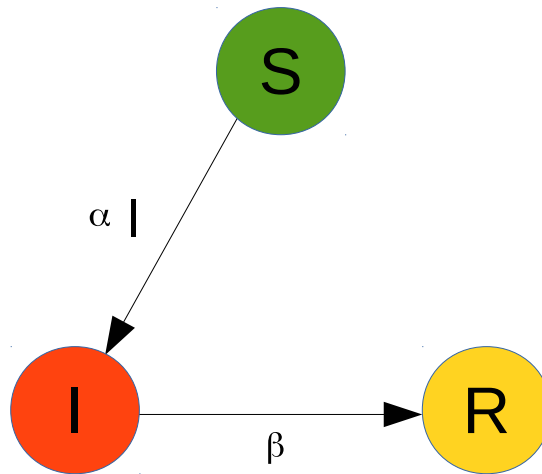
De estas consideraciones podemos concluir que en ambos casos un individuo puede encontrarse en tres estados diferentes mutuamente excluyentes; un estado inicial en el que es **susceptible**, otro estado en el que individuo se encuentra **infectado** y finalmente un estado final en el que éste se **recupera**.

El modelo epidemiológico que considera estos tres casos, se conoce como modelo **SIR** (por sus siglas en inglés “susceptible”, “infected”, “recovered”). Puesto que nos servirá de punto de partida para desarrollar nuestro modelo para la propagación de un rumor, lo expondremos brevemente a continuación.

### 1.2.1. Modelo SIR.

El modelo SIR fue desarrollado en 1927 por W.O.Kermack y A.G.McKendrick[7]. Dentro del marco de los modelos epidemiológicos, pertenece al grupo de los denominados modelos compartimentales, caracterización que es debida a que se divide la población en tres compartimentos; los susceptibles, los infectados, y los recuperados. Es un modelo que se corresponde con el comportamiento de enfermedades como la varicela o las paperas, procesos infecciosos de los que se obtiene inmunidad una vez que estos han sido superados. Junto con la hipótesis compartimental, la otra hipótesis sobre la que está fundamentado la versión clásica de este modelo es la denominada hipótesis de “homegenous mixing”. Lo que se supone con esta hipótesis es que cada individuo tiene la misma probabilidad de entrar en contacto con un individuo infectado. Esta hipótesis es la que elimina la necesidad de conocer todas las relaciones que se establecen entre los individuos, es decir es la que elimina la necesidad de trabajar con la red.

La descripción del modelo es la siguiente, como ya se ha mencionado, tenemos tres posibles estados, susceptible, infectado y recuperado. La evolución temporal está regida por las siguientes reglas:



- La tasa de contagio es  $\alpha$ .
- La tasa de recuperación es  $\beta$ .
- Una vez que un individuo se encuentra recuperado, permanece recuperado.

Con lo que la evolución temporal de los individuos que se encuentran en cada uno de los estados posibles viene dada por el siguiente sistema de ecuaciones diferenciales:

$$\begin{aligned}\frac{dS}{dt} &= -\alpha SI \\ \frac{dI}{dt} &= \alpha SI - \gamma I \\ \frac{dR}{dt} &= \gamma I\end{aligned}$$

Una de las características de estas ecuaciones es que podemos observar que hay un parámetro crítico para que se pueda producir una epidemia. Dicha condición se sigue directamente del sistema de ecuaciones anterior, para que haya posibilidad de que se produzca una epidemia se debe cumplir que  $\frac{dI}{dt} > 0$  de donde se sigue la siguiente condición:  $S > \frac{\gamma}{\alpha}$ . Es decir, para un  $\alpha$  y un  $\gamma$  fijados hay un número crítico de susceptibles el cual da lugar a una epidemia. Si denotamos con  $S_0$  el número inicial de susceptibles, la condición para que se de lugar a una epidemia se escribe como:

$$S_0 > \frac{\gamma}{\alpha}$$

Cabe señalar que la aplicación de la teoría de redes a este modelo modifica esta condición, uno de los resultados más importantes es que, en determinadas redes como por ejemplo las que se dan en las relaciones humanas o bien en la WWW, esta condición prácticamente desaparece, haciendo que para un gran abanico de valores de los parámetros  $\alpha$  y  $\gamma$  se produzca una epidemia[8]. Esto da una explicación a un gran número de fenómenos (como son por ejemplo la gran facilidad con la que se propaga un virus informático en la WWW, o bien las epidemias estacionales de gripe que se producen con relativa frecuencia) que la teoría clásica, en la que no se tiene en cuenta la estructura de red, se ve incapaz de explicar. Mediante este tipo de modelos se lleva a una conclusión sobre las redes antes mencionadas y es su gran vulnerabilidad frente a este tipo de procesos.

## Capítulo 2

# Modelo para la propagación de rumores.

En este capítulo desarrollaremos nuestro modelo para la propagación de rumores. Primero expondremos las redes sobre las que se trabajará y las razones para su elección, posteriormente se indicará el modelo desarrollado para la propagación del rumor explicando cuales han sido las consideraciones tenidas en cuenta.

### 2.1. Redes empleadas.

Mediante el estudio de las relaciones que se establecen entre individuos se ha llegado a la conclusión de que las redes que se establecen son las denominadas redes **libres de escala**; este tipo de redes serán las elegidas para nuestro estudio. También emplearemos otro tipo de redes, las redes aleatorias o de “**Erdős-Rényi**”, con el fin de estudiar cuales son los cambios en la dinámica cuando se modifica la estructura de la red.

#### 2.1.1. Red de Erdős-Rényi.

En primera instancia, las redes aleatorias o redes de Erdős-Rényi[9], en honor a los matemáticos Pál Erdős y Alfréd Rényi que durante los años sesenta mediante una serie de artículos desarrollaron toda una teoría en torno a este tipo de redes. Se tratan de los primeros modelos propuestos para entender como estaban constituidas las redes reales.

Un red aleatoria se define como sigue:

Una **red aleatoria** esta constituida de  $N$  nodos, donde cada pareja de nodos está conectada con un probabilidad  $p$ .

Basándonos en la definición anterior, podemos desarrollar un sencillo procedimiento para construir una red aleatoria:

- (1) Comenzamos con  $N$  nodos aislados
- (2) Seleccionamos una pareja de nodos, y generamos un numero aleatorio entre 0 y 1. Si este número es mayor que  $p$  conectamos ambos nodos, de otra forma los dejamos desconectados.
- (3) Repetimos el paso (2) para cada uno de los  $N(N - 1)/2$  pares de nodos.

Para obtener la distribución de grado, tenemos que tener en cuenta que la probabilidad de que un nodo tenga exactamente  $k$  conexiones viene dada por el producto de tres términos:

- La probabilidad de que tenga  $k$  conexiones,  $p^k$ .
- La probabilidad de que no este conectado con los nodos restantes  $(N - k - 1)$ , es decir,  $(1 - p)^{N-k-1}$ .
- El número de formas en las que podemos seleccionar  $k$  conexiones de las  $N - 1$  potenciales conexiones que podemos tener, esto es:

$$\binom{N-1}{k}$$

Con lo que finalmente la distribución de grado obtenida resulta ser:

$$p_k = \binom{N-1}{k} p^k (1-p)^{N-1-k}$$

Es decir se trata de una distribución binomial. En el límite cuando  $\langle k \rangle \ll N$ , la distribución binomial puede ser correctamente aproximada a una distribución de Poisson:

$$p_k = e^{-\langle k \rangle} \frac{\langle k \rangle^k}{k!}$$

Es importante mencionar que las redes aleatorias no se corresponden en la gran mayoría de los casos con las redes observadas en la realidad, éste hecho tiene importantes implicaciones. Significa que en la mayor parte de los casos los individuos que constituyen un sistema no interaccionan de manera aleatoria. Las relaciones que se establecen entre ellos siguen unas determinadas reglas, y son estas reglas que rigen a escala microscópica las que dan lugar las diferentes estructuras que presentan las redes.

Como se ha mencionado con anterioridad, aun cuando las redes aleatorias no se corresponden con redes reales, las presentaremos y las utilizaremos en las simulaciones, para comprobar cómo cambios en la estructura de la red modifican los resultados.

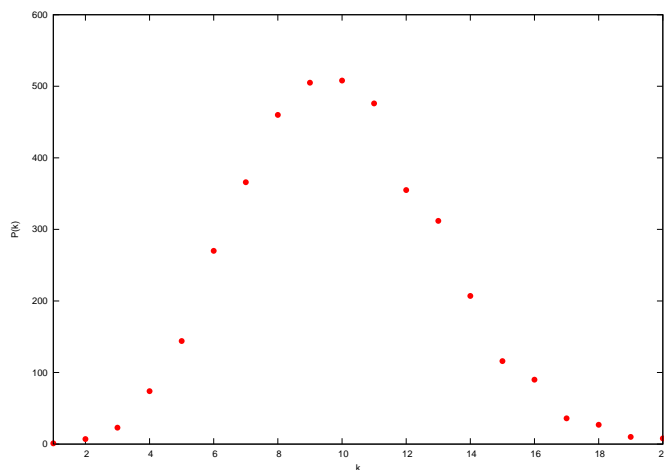


Figura 2.1: Distribución de la conectividad para una red de Erdős-Rényi.  $N = 4000$  ;  $p = 0,0025$ .

### 2.1.2. Red libre de escala.

El principal objeto con el que trabajaremos serán las redes libres de escala. Estas redes son, de forma siempre aproximada, las que surgen cuando tratamos con las relaciones que se establecen entre seres humanos. Aunque han demostrado también tener una aplicación mucho mayor, dado que en un gran número de sistemas las relaciones que se establecen entre sus componentes dan lugar a redes similares.

Como una primera definición, diremos que las redes libre de escala son aquellas redes cuya distribución de grado sigue una ley de potencias de la siguiente forma:

$$p_k \sim k^{-\gamma}$$

donde  $\gamma$  es un parámetro que caracteriza la red.

La razón de que este tipo de redes reciban el nombre de libres de escala es la siguiente. Definamos en primer lugar los momentos de la distribución, el momento enésimo de la distribución viene dado por la siguiente expresión:

$$\langle k^n \rangle = \sum_{k_{min}}^{\infty} k^n p_k = \int_{k_{min}}^{\infty} k^n p(k) dk$$

Los momentos de orden menor tienen importantes interpretaciones:

- $n=1$ : el primer momento es el grado medio,  $\langle k \rangle$ .
- $n=2$ : el segundo momento nos permite calcular la varianza  $\sigma^2 = \langle k^2 \rangle - \langle k \rangle^2$ , cuya raíz cuadrada, nos da la desviación estándar que es una medida de cómo están distribuidos los valores en torno a la media.

En el caso de una red libre de escala tenemos que:

$$\langle k^n \rangle = C \int_{k_{min}}^{k_{max}} k^n k^{-\gamma} dk = C \frac{k_{max}^{n-\gamma+1} - k_{min}^{n-\gamma+1}}{n - \gamma + 1}$$

Mientras que el valor de  $k_{min}$  está fijado, el valor de  $k_{max}$  depende en general del tamaño del sistema, con lo que para entender el comportamiento de  $\langle k^n \rangle$  es conveniente considerar el límite asintótico  $k_{max} \rightarrow \infty$ . En este límite observamos que el valor de  $\langle k^n \rangle$  depende de los valores de  $n$  y  $\gamma$ :

- Si  $n - \gamma + 1 \leq 0$  entonces el término  $k_{max}^{n-\gamma+1}$  se hace cero conforme  $k_{max}$  aumenta, lo que significa que los momentos de orden  $n \leq \gamma - 1$ , serán finitos.
- Si  $n - \gamma + 1 > 0$  entonces  $k_{max}^{n-\gamma+1}$  se hace infinito cuando  $k_{max} \rightarrow \infty$  con lo que los momentos de orden  $n > \gamma - 1$  divergen.

Este hecho tiene importantes consecuencias, para la mayoría de las redes libres de escala reales,  $\gamma$  toma valores entre 2 y 3 [1, cap. 4, p. 53], lo que conlleva que el primer momento  $\langle k \rangle$  se finito. Sin embargo, los momentos de orden superior,  $\langle k^2 \rangle, \langle k^3 \rangle \dots$  divergen, es decir, se hacen infinitos. En esta divergencia está la razón de que estas redes se denominen libres de escala. Para entender mejor este hecho analicemos paralelamente qué sucede con las redes aleatorias. En el caso de una red general, si elegimos un nodo al azar, obtendremos un valor para la conectividad que se encontrará típicamente en el rango dado por:

$$k = \langle k \rangle \pm \sigma_k$$

Veamos que valores toma  $\sigma$  para cada tipo de red:

- En el caso de una **red aleatoria**, que sigue una distribución de Poisson, el valor para  $\sigma_k = \langle k \rangle^{1/2}$  será siempre menor que  $\langle k \rangle$  con lo que los nodos tienen una conectividad en el rango dado por  $k = \langle k \rangle \pm \langle k \rangle^{1/2}$ . Es decir, los nodos de la red tienen una conectividad comparable y el valor de  $\langle k \rangle$  nos da una **escala** para la conectividad típica de sus elementos.
- Para una red **libre de escala** si se da el caso de que  $\gamma < 3$ , el primer momento es finito, mientras que el segundo momento es infinito. Entonces cuando  $N$ , el número de nodos, es elevado, las fluctuaciones en torno al valor medio  $\langle k \rangle$  pueden ser arbitrariamente grandes. Esto significa que, cuando elegimos un nodo de forma aleatoria, no sabemos que esperar; la conectividad del nodo podría ser arbitrariamente grande o pequeña. Con lo que podemos concluir que las redes con  $\gamma < 3$  no tienen una escala característica: son **libres de escala**.

Otro rasgo importante y característico de las redes libres de escala es la existencia de ciertos nodos denominados “**hubs**”. Un “hub” es un nodo que posee una gran conectividad en comparación con la conectividad promedio de la red. Este tipo de nodos juegan un papel vital en los procesos dinámicos que se llevan a cabo en una red, por ejemplo, en el caso de un proceso epidemiológico, si un “hub” resulta infectado, la infección se propagará rápidamente por el resto de la red debido a la gran cantidad de nodos que ahora son susceptibles de ser contagiados.

### Modelo de Barabási-Albert

Una vez expuestas las diferencias entre los dos tipos de redes empleadas, podemos preguntarnos cuál es la razón de que las interacciones entre varios sujetos den lugar a una u otra. Se ha expuesto cuál es el método por el que se origina una red aleatoria; dado una serie de nodos estos se conectan de forma aleatoria dada una determinada probabilidad. Como se ha mencionado, las redes aleatorias no se encuentran por lo general en la realidad mientras que las redes libres de escala se encuentran mucho más frecuentemente. Entonces, podemos preguntarnos cuál es la razón de que esto suceda, ¿por qué las redes aleatorias no se dan en la realidad?, ¿cuál es el proceso que da lugar a las redes libres de escala y que las dota de una mayor universalidad?. En primer lugar cabe señalar que el método por el cual se genera una red aleatoria parte de una premisa que es falsa y que no se da en la realidad, i.e., el modelo de las redes aleatorias asume que la red tiene un número fijo de nodos,  $N$ . Pero en la realidad esto no sucede; en las redes reales el número de nodos crece continuamente gracias a la adición de nuevos nodos. Como ejemplos podemos pensar en la WWW; cada vez que se crea una nueva página, ésta se enlaza con otras páginas existentes que a su vez forman parte de una estructura anterior. Otro ejemplo quizás menos evidente es el de las interacciones entre las proteínas que en un principio puede parecer estáticas, dado que el número de genes para cada persona es fijo, pero si lo examinamos detenidamente esto no es cierto nuestro número de genes ha crecido desde unos pocos hasta los cerca de 20000 que contienen una célula humana. Este proceso de evolución se ha llevado a cabo durante millones de años. Por lo tanto si buscamos un modelo que se corresponda con las redes reales debemos tener en cuenta que el número de nodos no es fijo.

El modelo de la red aleatoria también asume que las parejas de que se elijen para unirse lo hacen de modo aleatorio pero en la realidad esto no sucede. De hecho, en las redes reales los nodos que se incorporan a la red no eligen con qué nodos conectarse de forma aleatoria sino que prefieren unirse con los nodos que tienen una mayor conectividad. Este proceso se conoce como “preferential attachment”.

En resumen tenemos dos elementos en el modelo de las redes aleatorias que no son correctos:

- Crecimiento: Las redes reales son el resultado de un proceso que continuamente incrementa el número de nodos.
- “Preferential attachment”: En las redes reales, los nuevos nodos se incorporan a la red, eligiendo como nodos con los cuales conectarse aquellos que tienen una mayor conectividad.

Basándose principalmente en estos dos principios, Barabási-Albert [10] propusieron un modelo que permite generar redes libres de escala:

Comenzamos con  $m_0$  nodos; las conexiones entre ellos se eligen de manera aleatoria. Éste será el núcleo sobre el cual comenzamos a construir nuestra red. El resto de la red se desarrolla siguiendo los dos siguientes pasos:

- A cada instante de tiempo añadimos un nuevo nodo con  $m$  ( $\leq m_0$ ) conexiones que conectan el nuevo nodo a  $m$  nodos ya existentes en la red.
- La probabilidad  $p(k)$  de que un nuevo nodo se conecte al nodo  $i$  depende de la conectividad  $k_i$  de la forma:

$$p(k_i) = \frac{k_i}{\sum_j k_j}$$

Mediante este algoritmo generamos una red libre de escala con  $\gamma = 3$ . Este algoritmo será el empleado en las simulaciones cuando necesitemos trabajar con una red libre de escala.

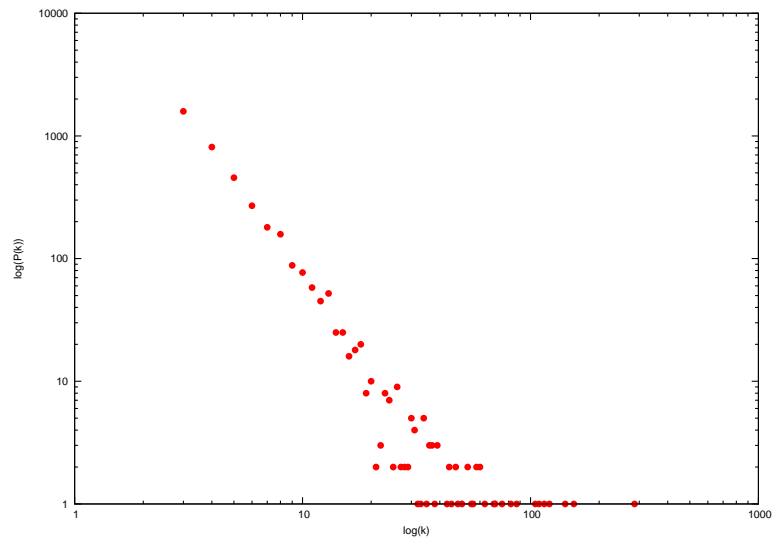


Figura 2.2: Distribución de la conectividad para una red libre de escala, generada mediante el modelo de Barabási-Albert.  $N = 4000$  ;  $m = 3$ .

## 2.2. Mecanismo de propagación.

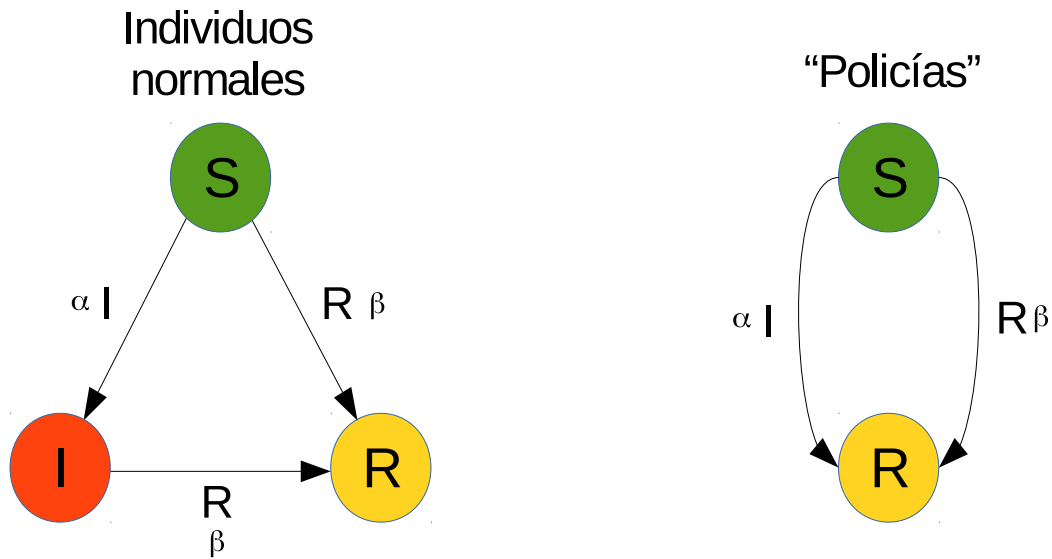
Basándonos en los modelos epidemiológicos desarrollamos también un modelo compartimental. En este caso los individuos podrán encontrarse en tres estados diferentes, cuando son susceptibles de creerse el rumor (S), una vez que han sido “infectados” (I), es decir que se han creído el rumor y lo comparten con los individuos cercanos y finalmente pueden encontrarse “recuperados” (R), una vez que conocen que el rumor no es cierto.

La dinámica propuesta es la siguiente:

- Mediante la mediación con un individuo infectado, un individuo sano puede ser contagiado con una probabilidad  $\alpha$ , es decir el individuo se cree el rumor. Por otro lado, otra opción que debemos tener en cuenta es la situación en la que mediante la interacción con un individuo que conoce que el rumor es falso el individuo sano puede pasar directamente a recuperado, el individuo recuperado le informa de la existencia del rumor y de que éste no es cierto, es decir un individuo sano también puede transitar al estado recuperado mediante la interacción con un individuo recuperado con probabilidad  $\beta$ .
- Un individuo que se encuentra infectado únicamente puede recuperarse cuando otro individuo recuperado interacciona con él, comunicándole que el rumor no es cierto. Es decir, un individuo infectado puede pasar a recuperado mediante la interacción con un individuo recuperado con probabilidad  $\beta$ .
- Finalmente un individuo que se encuentra recuperado, permanecerá en este estado. Su función en la dinámica del sistema será la de favorecer que otros individuos susceptibles o infectados, pasen a estar recuperados.

Ésta sería la dinámica seguida por los individuos “normales”, pero es necesario introducir otro tipo de individuos con el fin de modelizar correctamente la propagación del rumor. Primero de todo, cabe señalar que en el caso en que nos restringiéramos a utilizar solo los individuos indicados anteriormente, el rumor se extendería por toda la red, no habría ningún mecanismo por el cual algún individuo pudiese conocer que el rumor es falso. Para solucionar este hecho, introducimos un nuevo tipo de individuos, los “policías”. Mediante la adición de los mismos, buscamos introducir un mecanismo mediante el cual se pueda hacer frente al rumor. Estos nuevos individuos se corresponderían con determinadas personas que tienen cierto tipo de información que les permite discernir si el rumor es cierto o falso por sus propios medios, sin necesidad de la mediación con otros individuos. Por lo tanto un “policía” únicamente se puede encontrar en dos estados, un estado en el que susceptible, es decir todavía no conoce la existencia del rumor, y un otro estado en el que esta recuperado, conoce la existencia del rumor e interacciona con otros individuos, comunicándoles que el rumor es falso. Su dinámica es por lo tanto diferente:

- Mediante la mediación con un individuo infectado, un “policía” puede transitar directamente al estado recuperado con una probabilidad  $\alpha$ , de la misma forma mediante la interacción con otro individuo recuperado este también pasará a encontrarse recuperado con una probabilidad  $\beta$ .
- Cuando el “policía” se encuentre recuperado, la dinámica es la misma que para un individuo normal que se encuentre recuperado. Es decir su contribución será la de mediar para que el resto de individuos sean capaces de pasar al estado recuperado.



En resumen, con la adición de este nuevo tipo de individuos tenemos el proceso que nos faltaba anteriormente, es decir, un proceso mediante el cual podemos tener personas que conozcan que el rumor es falso, recuperadas, sin la necesidad de tener con anterioridad individuos de este tipo.

Desde el punto de vista puramente de la evolución del sistema, las personas "normales" y los "policías" son tratadas de la misma forma cuando interaccionan entre ellos, en el sentido de que si un individuo interacciona con otro y este es susceptible o se encuentra recuperado es irrelevante del tipo que sea.

Dado que este nuevo tipo de individuos son los encargados de comenzar con el proceso mediante el cual podemos controlar el rumor y evitar que este se extienda, se puede intuir desde un principio, y se comprobará posteriormente mediante las simulaciones, como su número y colocación son críticos. De hecho, como se mostrará, es suficiente un número muy reducido de estos agentes, siempre y cuando sean situados en puntos estratégicos de la red, para conseguir que el rumor apenas tenga efecto sobre la misma.

### 2.3. Control de la difusión del rumor.

Si lo que queremos es encontrar una serie de mecanismos que nos permitan detener la extensión del rumor, es evidente que estos estarán relacionados con los “policías”, dado que son estos individuos los que favorecen los mecanismos del sistema para frenar y hacer desaparecer el rumor. En primer lugar, puede parecer que la opción mas sencilla sería aumentar el número de estos sujetos. Sin embargo debido a la estructura de red que rige las relaciones entre los individuos, se mostrará que es mucho mas eficiente (en el sentido de que con un número menor de policías lograremos que su efecto en la expansión del rumor sea mucho mayor) desarrollar ciertas estrategias que nos permitan determinar cuáles son los sitios de la red en los que debemos situar a los “policías” para que su efecto se vea potenciado.

Dado que el ámbito en el que se lleva a cabo la propagación del rumor es el de las relaciones entre personas y que como hemos visto este tipo de relaciones dan lugar a redes libres de escala, serán estas redes las que debemos analizar detenidamente para determinar dónde debemos situar a los “policías”. Como se mencionó, una de las características de este tipo de redes es la existencia de ciertos nodos, los denominados “hubs”, cuya conectividad es muy superior a la media. Se llega inmediatamente a la conclusión de que en el caso de que en estos nodos se encuentre un “policía”, su acción sobre la red se verá notablemente potenciada. Basándonos en estas consideraciones podemos desarrollar una serie de estrategias para la distribución de los policías:

- Mayor Grado: En primer lugar debido a la existencia de los “hubs”, podemos recorrer la red y calcular la conectividad de cada nodo para posteriormente distribuir los “policías” entre los nodos de mayor a menor conectividad. De esta forma nos aseguramos que en los “hubs” haya un “policía”.
- Aleatorio-Aleatorio: Otra opción para la colocación de los policías consiste en elegir un nodo al azar y de entre sus vecinos elegir otro al azar en el que situaremos al policía. La razón para esto es la siguiente: si bien cuando elegimos un nodo al azar, la probabilidad de que éste sea un “hub” es muy pequeña, la probabilidad de que eligiendo uno de sus vecinos al azar tengamos un “hub” es mucho mayor. Esto es debido a la alta conectividad de los “hubs”, que hace que la probabilidad de que un nodo esté conectado con uno de ellos es mucho mayor que para el resto.
- Aleatorio-Mayor Grado: De forma similar al caso anterior, elegimos en primera instancia un nodo de forma aleatoria y posteriormente, de entre sus vecinos, ahora elegimos al que tenga una mayor conectividad. De esta forma la probabilidad de que este nodo sea un “hub” es mayor que en el caso anterior.
- Mínima distancia máxima: Podemos calcular la distancia máxima de cada nodo, es decir, para cada nodo, determinamos cuál es la distancia al nodo más lejano. Entonces, si distribuimos los “policías” en los nodos cuya distancia máxima sea mínima, los procesos iniciados en estos nodos se propagaran rápidamente por toda la red (en este caso la información de que el rumor no es cierto).

Cabe mencionar que aun a pesar de que puede parecer que no tiene mucho sentido proponer ciertas estrategias diferentes a la de situar los “policías” en los “hubs” de la red, si la examinamos con detenimiento, observamos que esta estrategia tiene un inconveniente: para llevarla a cabo es necesario conocer por completo la estructura de la red. Por otro lado, en las estrategias de “Aleatorio-Aleatorio” y “Aleatorio-Mayor Grado”, pese a no ser tan eficientes no es necesario conocer toda la red para implementarlas de forma correcta.

## Capítulo 3

# Resolución del modelo.

En esta sección intentaremos obtener una expresión que nos permita determinar cómo será la evolución temporal del sistema. En concreto, esto se llevará a cabo entendiendo el proceso como una cadena de Markov. Este tipo de proceso se expone brevemente a continuación.

**Proceso de Markov.** Los procesos de Markov forman parte de un tipo de procesos más generales conocidos como procesos estocásticos. Un proceso estocástico es una sucesión de variables aleatorias  $X_1, X_2, \dots$  donde el subíndice indica el tiempo. Un proceso de Markov se dice de Markov si satisface la llamada propiedad de Markov, es decir, la probabilidad condicionada:

$$P(X_n = x_n | X_{n-1} = x_{n-1}, \dots, X_1 = x_1) = P(X_n = x_n | X_{n-1} = x_{n-1})$$

o dicho con brevedad, un proceso de Markov sólo tiene memoria de lo acontecido en el instante previo. Llamamos al segundo miembro probabilidad de transición.

Para intentar dar una expresión de la evolución temporal de nuestro modelo, entendiéndolo en el marco de los procesos de Markov, debemos hacer algunas consideraciones previas. En primer lugar estamos empleando un enfoque probabilístico, con lo que ahora para un determinado tiempo  $t$  un individuo  $i$  no se encontrará en un estado determinado. Sino que se encuentra:

- En el estado  $S$ (susceptible) con probabilidad  $S_i(t)$ .
- En el estado  $I$ (infectado) con probabilidad  $I_i(t)$ .
- En el estado  $R$ (recuperado) con probabilidad  $R_i(t)$ .

Sobre estas tres probabilidades se exige la condición de normalización, de forma que tenemos que:  $S_i(t) + I_i(t) + R_i(t) = 1 \forall t$ .

Antes de comenzar con el cálculo de las probabilidades para cada nodo en función del tiempo, es conveniente explicar de manera un poco más detenida cómo se realiza el cálculo de las probabilidades de transición, puesto que estas probabilidades serán empleadas posteriormente en nuestros cálculos.

Para explicar de qué forma calculamos las probabilidades de transición nos apoyamos en la imagen adjunta. En esta imagen se representa una situación genérica en la que el nodo central es susceptible y se encuentra en contacto con otros nodos en diferentes estados; infectados, recuperados y otros también susceptibles. Supongamos por sencillez que el nodo central es un individuo normal (no un “policía”). Calculemos, por ejemplo, la probabilidad de que este nodo se encuentre infectado al tiempo  $t+1$ . Para que esto suceda, es suficiente con que alguno de los nodos vecinos que se encuentren infectados lo contagien. Para calcular la probabilidad de que esto suceda en un principio podríamos pensar en calcular la probabilidad de que alguno de los nodos lo infecte. Sin embargo, la dificultad del cálculo de esta probabilidad se incrementa con el número de nodos que pueden contagiar al nodo central. Por este motivo, calcularemos la probabilidad del suceso complementario, i.e., que ningún nodo infectado infecte al nodo central, y luego tomaremos de nuevo la probabilidad complementaria. Habremos calculado así la probabilidad de que al menos un nodo contagie al nodo central. La probabilidad de que un nodo infectado contagie a otro nodo con el que está en contacto, viene dada por el parámetro  $\alpha$ , con lo que la probabilidad complementaria, es decir, que el nodo no lo contagie será  $1 - \alpha$ ; finalmente fijándonos en la figura, tendremos que la probabilidad de que ningún nodo infecte al nodo central será:

$$(1 - \alpha) \cdot (1 - \alpha) \cdot (1 - \alpha) = (1 - \alpha)^3$$

Con lo que podemos concluir que la probabilidad de que, al tiempo  $t+1$ , el nodo central se encuentre infectado viene dada por:

$$1 - (1 - \alpha)^3$$

De la misma forma podemos escribir la probabilidad de que el nodo central pase a estar recuperado, debido a la mediación con alguno de los nodos recuperados:

$$1 - (1 - \beta)^2$$

Como se ha mencionado, cuando entendemos la dinámica como un proceso de Markov, los nodos no se encuentran en un estado definido para cada tiempo, sino que a cada uno de sus posibles estados se le asocia una probabilidad, con lo que por ejemplo, ahora la probabilidad de que cierto nodo  $i$  infecte a otro nodo con el cual está conectado, es  $\alpha I_i(t)$ . Entonces para la situación descrita en la figura, la probabilidad de contagio para un determinado instante de tiempo  $t$  será (asumiendo la independencia de las probabilidades  $I_1, I_2$  e  $I_3$ ):

$$1 - (1 - \alpha I_1(t)) \cdot (1 - \alpha I_2(t)) \cdot (1 - \alpha I_3(t))$$

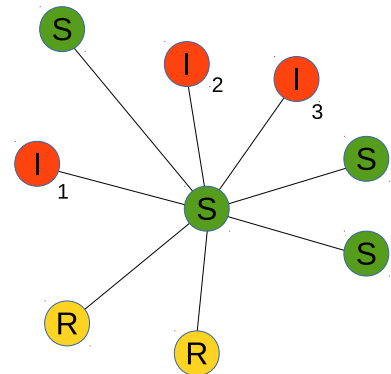


Figura 3.1: Ejemplo cálculo de probabilidades.

Para el caso mas general podemos escribir la probabilidad de transición ( $S \rightarrow I$ ), de que un nodo susceptible,  $i$ , se infecte a un tiempo  $t$  como

$$P_i(t)_{S \rightarrow I} = 1 - \prod_{j=1}^N (1 - \alpha A_{ij} I_j(t)),$$

donde  $N$  es el número total de nodos (el producto recorre toda la red),  $A_{ij}$  se corresponde con la matriz de adyacencia de la red;  $A_{ij} = 1$  si los nodos  $i$  y  $j$  están conectados o bien  $A_{ij} = 0$  si los nodos están separados o por otro lado si  $i = j$ .

Escribamos finalmente el conjunto de probabilidades de transición para los individuos normales:

- $P_i(t)_{S \rightarrow I} = 1 - \prod_{j=1}^N (1 - \alpha A_{ij} I_j(t))$
- $P_i(t)_{S \rightarrow R} = 1 - \prod_{j=1}^N (1 - \beta A_{ij} R_j(t))$
- $P_i(t)_{I \rightarrow R} = 1 - \prod_{j=1}^N (1 - \beta A_{ij} R_j(t))$

En el caso de los policías, para calcular la probabilidad de transición de susceptible a recuperado, debemos tener en cuenta que la transición se puede llevar a cabo mediante la mediación con una persona infectada o bien mediante la mediación con una persona recuperada. Por lo tanto la probabilidad de que no pase al estado recuperado viene dada por la probabilidad de que ambos sucesos sucedan simultáneamente, es decir, que no transite debido a la mediación con un infectado y que no transite mediante la mediación con un recuperado, esta probabilidad se corresponde con el producto de las dos probabilidades, con lo que la probabilidad buscada será su complementaria:

- $P_i(t)_{S \rightarrow R} = 1 - [\prod_{j=1}^N (1 - \alpha A_{ij} I_j(t))] [\prod_{j=1}^N (1 - \beta A_{ij} R_j(t))]$

Una vez llegados a este punto estamos preparados para escribir las ecuaciones para la evolución temporal. Como se ha mencionado, nos encontramos ante un proceso de Markov, por lo que el estado a tiempo  $t+1$  depende de cuál era el estado a tiempo  $t$ . Para cada nodo tendremos un conjunto de ecuaciones, una para cada estado posible. Escribamos por ejemplo las ecuaciones para la evolución del estado recuperado en el caso de un nodo normal:

$$R_i(t+1) = R_i(t) + I_i(t) \cdot P_i(t)_{I \rightarrow R} + S_i(t) \cdot (1 - P_i(t)_{S \rightarrow I}) P_i(t)_{S \rightarrow R} \quad (3.1)$$

Examinemos cada término:

- $R_i(t)$  Dado que una vez que el individuo se encuentra recuperado no puede transitar a otro estado, si se encuentra recuperado a tiempo  $t$  lo estará a  $t+1$ .
- $I_i(t) \cdot P_i(t)_{I \rightarrow R}$  Si el individuo se encuentra infectado a tiempo  $t$ , es posible que mediante la interacción con un individuo recuperado, este pase a estar recuperado en  $t+1$ .
- $S_i(t) \cdot (1 - P_i(t)_{S \rightarrow I}) P_i(t)_{S \rightarrow R}$  Si el individuo es susceptible a tiempo  $t$ , es posible que mediante la interacción con un individuo recuperado se encuentre recuperado a  $t+1$ . Con lo que la probabilidad de que esta transición suceda será la probabilidad de que el individuo esté susceptible a tiempo  $t$  por la probabilidad de que pueda pasar a recuperado multiplicado por la probabilidad de que no se contagie por la acción de un infectado.

Siguiendo un planteamiento similar y utilizando la condición de normalización, que nos permite expresar una ecuación en función de las otras, podemos escribir el resto de las ecuaciones para cada tipo de individuo:

■ Individuos normales:

- $R_i(t+1) = R_i(t) + I_i(t) \cdot P_i(t)_{I \rightarrow R} + S_i(t) \cdot (1 - P_i(t)_{S \rightarrow I})P_i(t)_{S \rightarrow R}$
- $I_i(t+1) = I_i(t) \cdot (1 - P_i(t)_{I \rightarrow R}) + S_i(t) \cdot P_i(t)_{S \rightarrow I}(1 - P_i(t)_{S \rightarrow R})$
- $S_i(t+1) = 1 - R_i(t+1) - I_i(t+1)$

■ Individuos “policías”:

- $R_i(t+1) = R_i(t) + S_i(t) \cdot P_i(t)_{S \rightarrow R}$
- $S_i(t+1) = 1 - R_i(t+1)$

Como aquello que nos interesa conocer es el impacto del rumor en la población, debemos encontrar algún tipo de marcador que dé cuenta de cómo éste ha sido y que, a su vez, nos permita comparar cómo varía dicho impacto cuando modificamos ciertos parámetros del modelo. Estos parámetros serán los valores de  $\alpha$  y  $\beta$ , las posiciones de la red en donde se sitúan los “policías” (sin embargo, el número inicial de infectados será siempre pequeño, puesto que son un número reducido de personas las que originan un rumor). También comprobaremos como varían los resultados para diferentes tipos de redes. Una vez fijados estos parámetros iniciales, procedemos con la iteración de las ecuaciones anteriores. Éste proceso nos permite obtener resultados de una forma mucho más eficiente que mediante la simulación directa del sistema.

Como cantidad que cuantifica el impacto del rumor, hemos elegido el número de individuos que son “infectados”, es decir, el número de individuos que en algún momento han llegado a considerar el rumor como cierto. La expresión que nos aporta esta información se obtiene teniendo en cuenta que, para un instante de tiempo, el número promedio de personas que son infectadas viene dado por la expresión  $\sum_{i=1}^N S_i(t) \cdot P_i(t)_{S \rightarrow I} \cdot (1 - P_i(t)_{S \rightarrow R})$ , con lo que el número total de sujetos infectados durante el proceso se puede calcular mediante la expresión:

$$\#(S \rightarrow I) = \sum_{t=0}^{t=\infty} \left[ \sum_{i=1}^N S_i(t) \cdot P_i(t)_{S \rightarrow I} \cdot (1 - P_i(t)_{S \rightarrow R}) \right]$$

## Capítulo 4

# Resultados.

En esta sección expondremos los resultados obtenidos mediante los cálculos realizados con las herramientas expuestas en la sección anterior. Comprobaremos cómo se ajustan las ecuaciones obtenidas anteriormente a los datos obtenidos mediante la simulación directa del proceso.

Con el fin de comprobar la validez de las herramientas desarrolladas, realizaremos una simulación directa del proceso mediante el método de Monte Carlo y comprobaremos cómo las expresiones anteriores se ajustan a los resultados obtenidos mediante dichas simulaciones.

Para realizar la simulación de Monte Carlo procedemos de la manera siguiente; a cada instante de tiempo, recorremos los nodos y, para cada nodo inspeccionamos sus vecinos, entonces en función del estado del vecino examinado comprobamos si este modifica el estado del nodo, por ejemplo en el caso de que el nodo fuese susceptible y el vecino examinado estuviese infectado, generamos un número aleatorio,  $n$ , entre 0 y 1; en el caso en el que  $n < \alpha$ , el nodo se infectará, en caso contrario examinaremos otro vecino. Si al final del proceso el nodo no ha modificado su estado, permanecerá en el estado inicial. Para cada conjunto de valores de los parámetros realizaremos varias simulaciones, consiguiendo así un conjunto de datos sobre los que podremos hacer estadística y obtener de esta forma medidas más precisas de los valores que queremos conocer.

En primer lugar, comparamos cómo evoluciona en el tiempo el sistema para diferentes valores de los parámetros.

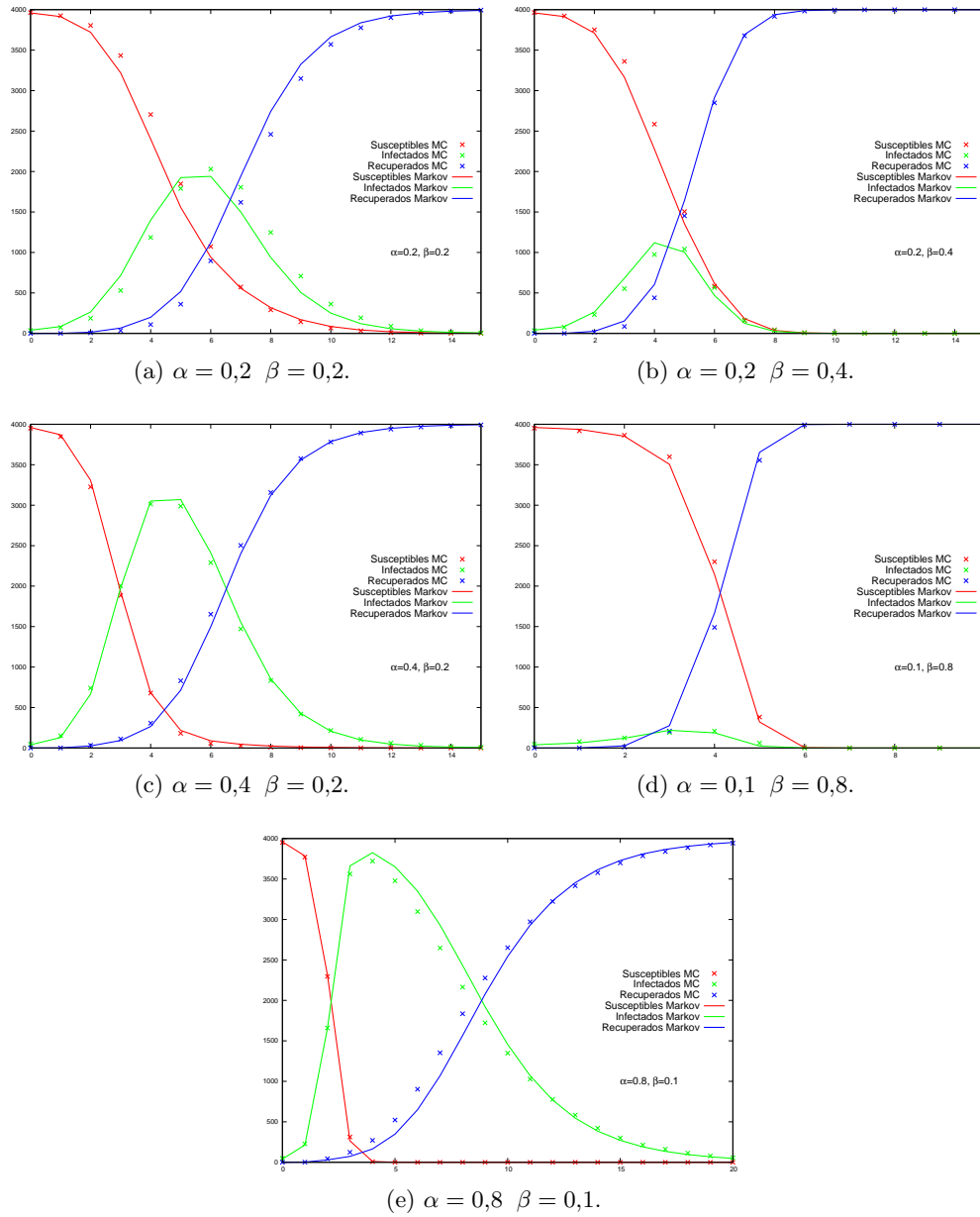


Figura 4.1: Evolución temporal del número de individuos susceptibles, infectados y recuperados para diferentes valores de los parámetros. En todos los casos, el porcentaje de policías es del 1 % y el porcentaje inicial de infectados es también del 1 %. La red empleada es una red libre de escala, generada mediante el Modelo de Barabási-Albert con  $N = 4000$  ;  $m = 3$  ;  $\langle k \rangle = 6$

Como se aprecia en las imágenes, salvo pequeñas desviaciones, hay un buen acuerdo entre los datos aportados por la simulación de Monte Carlo y lo predicho por las ecuaciones de Markov.

En segundo lugar conviene comprobar la validez de la ecuación:

$$\#(S \rightarrow I) = \sum_{t=0}^{t=\infty} \left[ \sum_{i=1}^N S_i(t) \cdot P_i(t)_{S \rightarrow I} \cdot (1 - P_i(t)_{S \rightarrow R}) \right]$$

mediante la cual determinamos el impacto del rumor en la población.

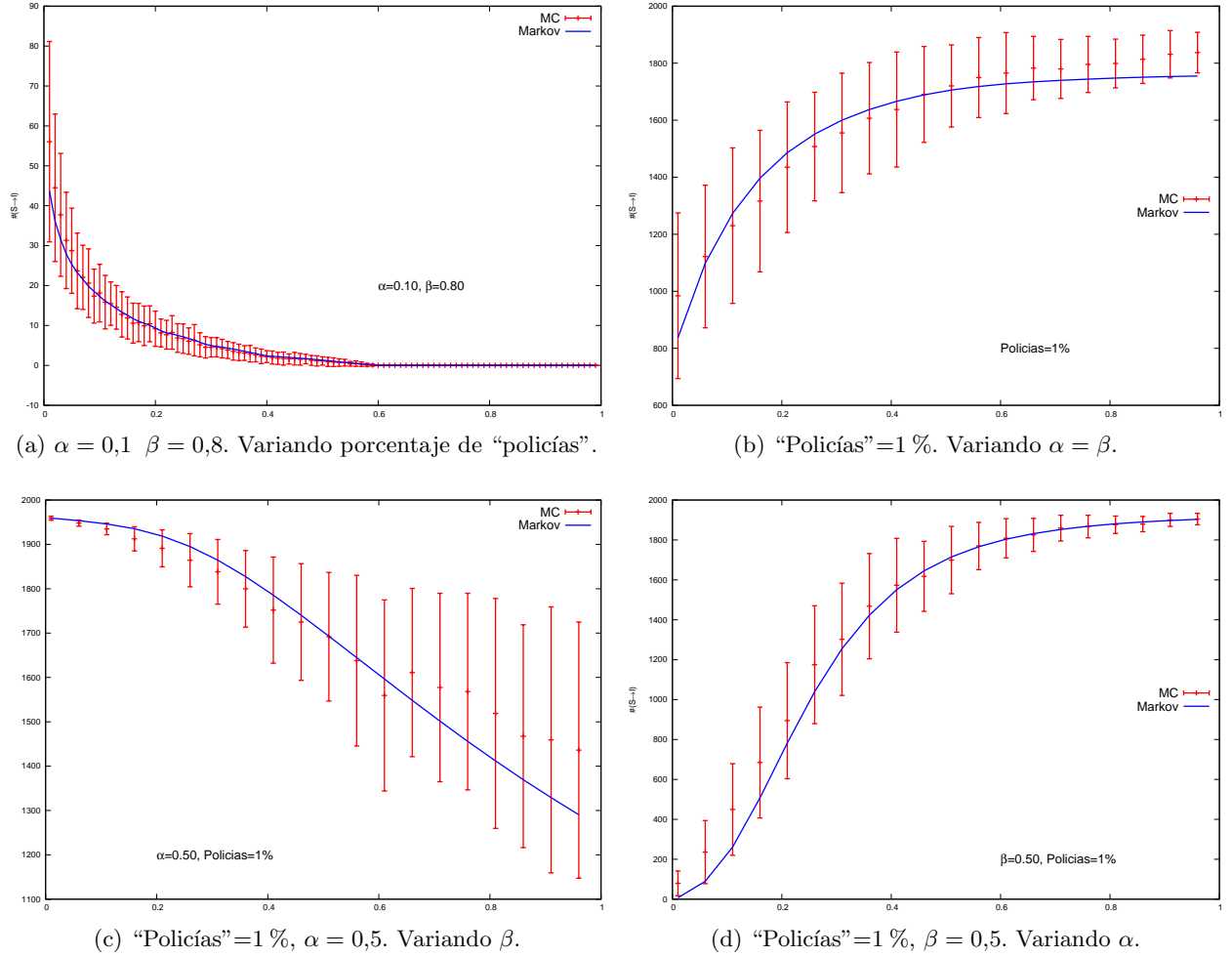


Figura 4.2: Comparación del numero de individuos que alguna vez han estado infectados, calculado mediante la expresión dada por el proceso de Markov y determinados mediante la simulación directa. En todos los casos, el porcentaje de policías es del 1 % y el porcentaje inicial de infectados es también del 1 %. La red empleada es una red libre de escala, generada mediante el Modelo de Barabási-Albert con  $N = 2000$  ;  $m = 3$  ;  $\langle k \rangle = 6$ .

En este caso también se observa que el ajuste entre los datos dados por la simulación directa y los estimados mediante las ecuaciones es correcto.

Finalmente, mostramos los resultados obtenidos para la implementación de las diferentes estrategias de control de la difusión del rumor. Mostramos el numero de individuos que en algún momento han sido infectados, frente al porcentaje de “policías” en la red.

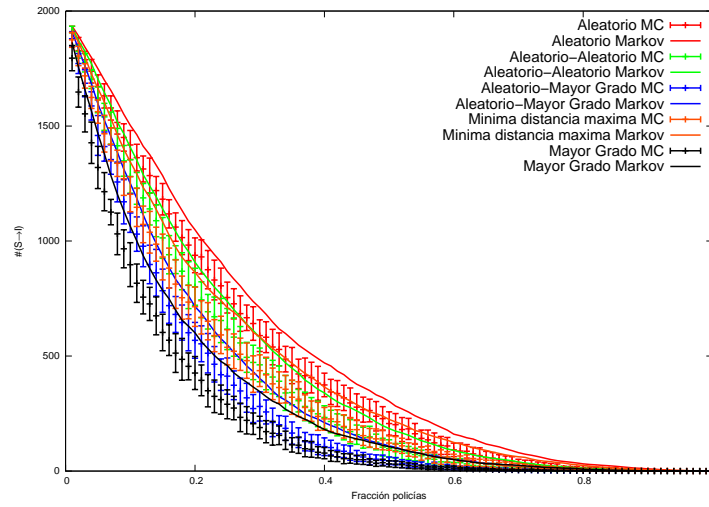


Figura 4.3: Estrategias para el control del rumor. La red empleada es una red de Erdős-Rényi con  $N = 2000$  ;  $p = 0,003$  ;  $\langle k \rangle = 6$ .  $\alpha = 0,4$  ;  $\beta = 0,3$ .

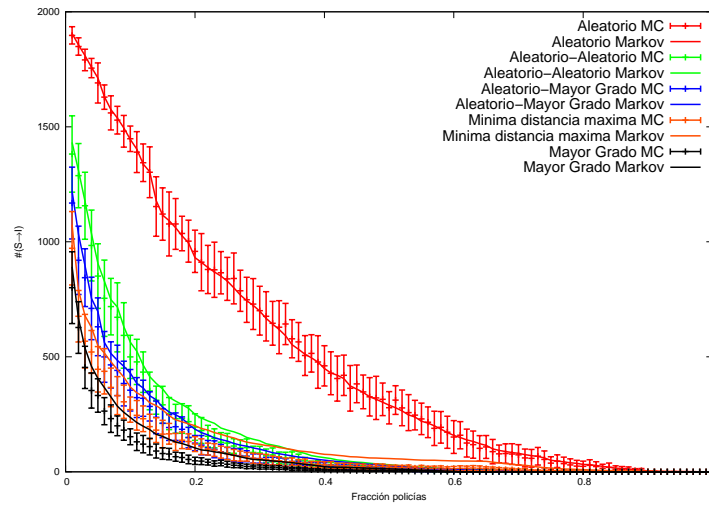


Figura 4.4: Estrategias para el control del rumor. La red empleada es una red libre de escala, generada mediante el Modelo de Barabási-Albert con  $N = 2000$  ;  $m = 3$  ;  $\langle k \rangle = 6$ .  $\alpha = 0,4$  ;  $\beta = 0,3$ .

## Capítulo 5

# Conclusiones.

Basándonos en los modelos epidemiológicos clásicos desarrollados con el fin de modelizar el comportamiento de un agente infeccioso, hemos construido un modelo que pretende explicar como se comporta un rumor cuando éste se propaga entre una serie de personas. El elemento central que hemos implementado y que es la principal diferencia entre el comportamiento de un virus y el de un rumor, es la adición de cierto mecanismo de “inmunización activa”; a diferencia del comportamiento propio de un proceso infeccioso, ahora los individuos pueden “recuperarse” mediante la interacción con otros individuos. Parte esencial de este nuevo mecanismo de “inmunización” es la existencia de un nuevo tipo de individuos, los “policías”, que actúan como centros de propagación para el proceso de “recuperación”.

Paralelamente, otro importante elemento del modelo es el estudio detallado de la estructura que emerge fruto de las interacciones entre personas. Dicho estudio se lleva a cabo mediante el empleo de la teoría de redes. La motivación para el uso de estos elementos ha sido el fuerte impacto que han tenido en un gran número de campos (en concreto en el ámbito de los procesos epidemiológicos donde la consideración de este tipo de estructura ha dado lugar a la explicación de ciertos fenómenos, como por ejemplo el bajo umbral de epidemia) que no podían ser explicados mediante el uso de otras técnicas, en particular mediante el empleo de aproximaciones de campo medio. En nuestro caso, hemos optado por trabajar con la red entera, sin realizar otro tipo de aproximaciones como podrían ser la aproximaciones de campo medio heterogéneo[11], que si bien tienen en cuenta cierta estructura en la relación entre los individuos, no dan lugar a una precisión en los resultados tan alta como la obtenida mediante el uso de la red.

Hemos desarrollado nuestro modelo mediante el uso de los elementos antes mencionados. Además hemos propuesto una serie de estrategias para la colocación de los “policías” que nos permiten minimizar el impacto del rumor en la red de modo que éste sea lo más pequeño posible. Cabe mencionar en este punto como estas estrategias también guardan cierta semejanzas con las estrategias de “inmunización” que se estudian en el ámbito de la propagación de las enfermedades infecciosas. Sin embargo, hay una importante diferencia: como se ha mencionado con anterioridad, el proceso mediante el cual los individuos se “recuperan” difiere del de un proceso infeccioso originado por un virus, pues aquí los individuos “recuperados” se convierten en fuentes de un nuevo proceso (la propagación de la recuperación). Con lo que aparece una competición entre los dos procesos, el proceso de “infección” y el proceso de “recuperación”. Esta diferencia también hace que las estrategias para frenar el avance, o bien del rumor o bien del virus, difieran. En el caso de un virus, se busca inmunizar a los individuos que se encuentren en posiciones de la red que permitan frenar el avance de la infección [12][13], mientras que en el caso de la propagación del rumor, lo que se busca es que la posición de los “policías” sea tal, que la acción del proceso de “recuperación” sea máxima.

Para mostrar de manera más clara como la estructura de la red condiciona la dinámica de los procesos que se llevan a cabo en la misma, elaboramos la siguiente tabla. En la primera columna se indica el porcentaje de individuos que en algún momento han llegado a considerar el rumor cierto, es decir el impacto del rumor en la red. Mientras que en la segunda y tercera columna se muestra la fracción de “policías” que hay que introducir en la red para lograr este objetivo, colocándolos siguiendo la estrategia de “Mayor Grado”, pero tratándose de una red de “Erdős-Rényi” en la segunda columna y de una red libre de escala en la tercera.

% Impacto	Erdős-Rényi	Libre de escala
40 %	0.11	0.01
20 %	0.21	0.04
5 %	0.36	0.14
1 %	0.54	0.3

Se aprecia claramente cómo en el caso de una red libre de escala la estrategia para frenar el avance del rumor es mucho más eficiente y simplemente hace falta añadir un número muy reducido de “policías” para que el impacto del rumor se vea notablemente reducido. Esto es debido a las diferencias en la estructura de la red, en concreto a la existencia de “hubs” en las redes libres de escala.

Una contribución importante del trabajo presentado es la modelización de la dinámica de propagación de rumores en términos de un proceso markoviano. El buen ajuste de los resultados de la dinámica markoviana a los datos de las simulaciones numéricas, parece indicar que genéricamente, la hipótesis de independencia de las probabilidades asociadas a nodos del entorno, subyacente al cálculo de las probabilidades de transición, no distorsiona de forma importante los resultados, constituyendo así una buena aproximación. Ello permite un procedimiento enormemente eficiente de obtención de resultados.

En resumen, hemos mostrado que haciendo uso de los modelos epidemiológicos clásicos e introduciendo ciertas modificaciones (principalmente el procesos de “inmunización activa” y empleando la teoría de redes, la cual nos permite encapsular y trabajar con las complicadas relaciones que se establecen entre las personas) somos capaces de desarrollar un modelo que nos permite entender en cierta medida cuál es el comportamiento de un rumor y mediante el que podemos desarrollar estrategias para controlar su propagación.

# Bibliografía

- [1] Albert-László Barabási, *Network Science*. Cambridge University Press, 2015.
- [2] Sergey Dorogovtsev, *Evolution of Networks: From Biological Nets to the Internet and WWW*. Oxford University Press, 2003.
- [3] O.Sporns, G.Tononi, and R.Kötter., *The Human Connectome: A Structural Description of the Human Brain*. ,PLoS Computational Biology,1:4,2005.
- [4] Z.N. Oltvai and A.-L.Barabási., *Network Biology: Understanding the cell's functional organization*. , Nature Reviews Genetics,5:101,2004.
- [5] J.J.Hopfield., *Neuronal networks and physical systems with emergent collective computational abilities*. , PNAS,79:2554,1982
- [6] Wataru Soumaa, Yoshi Fujiwarab, Hideaki Aoyamac., *Complex networks and economics*. ,Physica A: Statistical Mechanics and its Applications, 2003, vol. 324, no 1, p. 396-401.
- [7] Kermack, William O.; McKendrick, Anderson G., *A contribution to the mathematical theory of epidemics*. ,En Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences. The Royal Society, 1927. p. 700-721.
- [8] Gómez, S., Arenas, A., Borge-Holthoefer, J., Meloni, S. and Moreno, Y. , *Discrete-time Markov chain approach to contact-based disease spreading in complex networks*. , EPL (Europhysics Letters), 2010, vol. 89, no 3, p. 38009.
- [9] P.Erdős and A.Rényi, *On random graphs*. ,Publicationes Mathematicae, 6:290,1959.
- [10] A.-L.Barabási and R.Albert., *Emergence of scaling in random networks*. ,Science,286:509-512,1999.
- [11] Vespignani, A., *Modelling dynamical processes in complex socio-technical systems*. ,Nature Physics, 8(1), 32-39,2012.
- [12] Pastor-Satorras, R., and Vespignani, A., *Immunization of complex networks*. ,Physical Review E, 65(3), 036104,2002.
- [13] Kitsak, M., Gallos, L. K., Havlin, S., Liljeros, F., Muchnik, L., Stanley, H. E., and Makse, H. A., *Identification of influential spreaders in complex networks*. ,Nature Physics, 6(11), 888-893,2010.