

# Modelos metapoblacionales para la difusión de epidemias

Trabajo de fin de grado  
Alberto Aleta Casas

Director  
Yamir Moreno Vega

Departamento de Física Teórica  
Universidad de Zaragoza  
Zaragoza, Junio 2014



*According to the evidence put forward in the preceding pages the space-time events in the body of a living being which correspond to the activity of its mind, to its selfconscious or any other actions, are (considering also their complex structure and the accepted statistical explanation of physicochemistry) if not strictly deterministic at any rate statistico-deterministic.*

*What is life?*

Erwin Schrödinger, 1944



# Índice

<b>1</b>	<b>Introducción</b>	<b>6</b>
<b>2</b>	<b>Fundamentos de la teoría de redes</b>	<b>7</b>
2.1	La matriz de adyacencia . . . . .	7
2.2	Grado de un vértice y distribución de grado . . . . .	7
2.3	Distribución <i>power-law</i> . . . . .	8
2.4	Redes <i>scale-free</i> . . . . .	9
<b>3</b>	<b>Difusión de epidemias</b>	<b>13</b>
3.1	Modelos compartimentales . . . . .	13
3.2	<i>Homogeneous mixing</i> . . . . .	14
3.3	Difusión en redes <i>scale-free</i> . . . . .	17
<b>4</b>	<b>Modelos metapoblacionales</b>	<b>21</b>
4.1	La difusión en metapoblaciones: procesos reacción-difusión . . . . .	21
4.2	<i>Global invasion threshold</i> . . . . .	23
4.3	Aplicaciones: la pandemia de gripe A/H1N1 . . . . .	25
<b>5</b>	<b>Estudio de la persistencia en metapoblaciones</b>	<b>29</b>
5.1	Modelo . . . . .	29
5.2	Resultados . . . . .	30
<b>6</b>	<b>Conclusiones</b>	<b>33</b>



# 1 Introducción

Un sistema complejo es aquel que muestra comportamientos que son consecuencia de la interacción entre los elementos que lo constituyen y que no pueden explicarse del análisis de dichos elementos por separado. A estos comportamientos se les denomina comportamientos emergentes o colectivos[1].

Encontramos sistemas complejos en multitud de disciplinas como la física, la biología o la economía. Sin embargo, detrás de cada uno de ellos siempre existe una red que define las interacciones entre sus componentes, la cual es fundamental para comprender el funcionamiento de estos sistemas.

Un proceso que resulta de particular interés en el estudio de los sistemas complejos es el fenómeno de la difusión en redes complejas, capaz de explicar desde la propagación de una enfermedad hasta la de un virus de ordenador o la de un rumor.

La introducción de la física estadística, la teoría de transiciones de fase y la de fenómenos críticos ha resultado de gran utilidad para describir el comportamiento macroscópico de los brotes epidémicos. En particular, la aproximación de campo medio ha resultado un gran éxito al conseguir reducir enormemente los grados de libertad del sistema[2].

En los últimos años el interés por el estudio de las epidemias se ha incrementado enormemente. Vivimos en una sociedad cada día más grande y conectada, lo que es el caldo de cultivo perfecto para una epidemia. Así, mientras que la peste negra tardó 4 años en pasar de Francia a Suecia en el siglo XIV, en el año 2009 la gripe A se propagó de México a España en apenas un mes.

Los modelos clásicos de difusión de epidemias se centran en estudiar la propagación dentro de poblaciones aisladas. Sin embargo, hoy en día se hace necesario extender estos modelos para abarcar múltiples poblaciones separadas espacialmente que no son independientes entre sí. Es en este contexto en el que surgen los modelos metapoblacionales.

El presente trabajo tiene como objetivo último desarrollar un modelo computacional que nos permita estudiar la persistencia de una epidemia en una metapoblación. Para conseguirlo será necesario analizar y comprender cómo se puede modelar la difusión de una epidemia en una población y el papel que juega la estructura de la red subyacente en el proceso.

Por ello, el texto se organiza como sigue:

- i) Comenzaremos con una breve introducción a la teoría de redes en la que definiremos el concepto de grado y su influencia en la estructura de una red.
- ii) A continuación estableceremos los principios básicos que gobiernan la difusión de las epidemias. Formularemos las hipótesis de “homogeneous mixing” y de campo medio que formarán una parte fundamental de los modelos metapoblacionales.
- iii) Integraremos todos estos conceptos en la construcción de un modelo metapoblacional y determinaremos algunas de sus propiedades más importantes.
- iv) Finalizaremos aplicando todos los conocimientos adquiridos en el desarrollo de un modelo computacional. Estudiaremos la persistencia de una epidemia en una metapoblación en función del tiempo promedio de inmunización de los individuos, de su movilidad y de la estructura de la red.

Los resultados que obtengamos podrán servir como punto de partida para futuros desarrollos teóricos.

## 2 Fundamentos de la teoría de redes

Un *grafo* es un conjunto de *vértices* unidos por enlaces (*edges*). Esta denominación se emplea al hablar de la representación matemática, mientras que para sistemas reales se prefiere la palabra *red*[3]. Así, en física, tenemos *redes* formadas por *sites* y *bonds*; en sociología *redes* de *actores* y *ties*; y en los sistemas complejos *redes* de *nodos* unidos por *links*.

A pesar de la aparente simplicidad de este esquema está claro que existen numerosas diferencias entre una red cristalográfica, una red de transporte o una red metabólica. A continuación veremos algunas de las propiedades más importantes de las redes, en especial aquellas relacionadas con el tipo de redes que vamos a emplear para estudiar la difusión de epidemias.

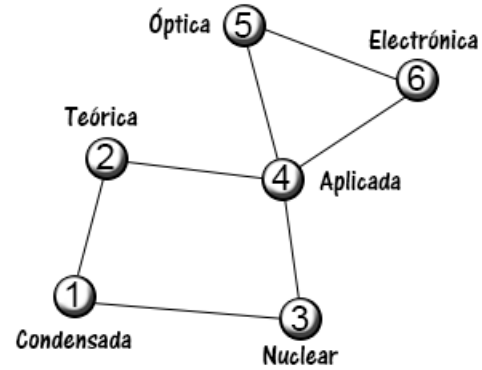
### 2.1 La matriz de adyacencia

Existen muchas formas de representar una red matemáticamente, pero la más cómoda para realizar cálculos es emplear una *matriz de adyacencia*.

Se define como matriz de adyacencia a la matriz  $A$  cuyos elementos  $A_{ij}$  valen 1 si existe un enlace entre los vértices  $i$  y  $j$  y 0 en cualquier otro caso.

Por ejemplo, la matriz de adyacencia correspondiente a la red mostrada en la figura 1 será

$$A = \begin{pmatrix} 0 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 0 \end{pmatrix}$$



**Figura 1:** Red de áreas y departamentos de física en la Universidad de Zaragoza (distribución espacial).

En ocasiones es útil asociar a los enlaces un número real en lugar de un 0 o un 1. A estas redes se les denomina *redes pesadas*. Un ejemplo lo podemos encontrar en el modelo *tigh-binding* de la física del estado sólido, donde los enlaces entre sites se representan por un parámetro  $t$  cuyo valor depende de la integral de solapamiento entre los orbitales de sites vecinos[4, p.143].

### 2.2 Grado de un vértice y distribución de grado

El *grado* de un vértice es igual al número de enlaces que posee. Así, en una red de  $N$  vértices el grado  $k_i$  de un nodo  $i$  se puede expresar en términos de la matriz de adyacencia como

$$k_i = \sum_{j=1}^N A_{ij}$$

Sea  $L$  el número total de enlaces. Como cada enlace está presente en dos vértices tendremos que

$$2L = \sum_{i=1}^N k_i \quad \Rightarrow \quad L = \frac{1}{2} \sum_{ij} A_{ij}$$

Con estas relaciones el grado medio  $\langle k \rangle$  de un vértice se puede expresar como

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



Un tipo especial de redes son aquellas cuyos vértices tienen todos el mismo grado. A estas redes se les conoce como *redes regulares* y son las que nos solemos encontrar al estudiar el modelo tight-binding mencionado anteriormente (*periodic lattice*).

Se conoce como *distribución de grado* a la distribución de probabilidad de un grado en la red. Definimos como  $P(k)$  a la fracción de vértices de una red que poseen grado  $k$ . Así,  $P(k)$  representa la probabilidad de que un vértice aleatorio de la red sea de grado  $k$ .

Aunque la distribución de grado no determina completamente una red, sí que es una característica fundamental de su estructura y nos permite conocer muchas de sus propiedades.

### 2.3 Distribución *power-law*

El modelo más sencillo en el que podemos pensar para generar una red es el de Erdős-Renyi (ER)[5]. En este modelo cada par de nodos se une con la misma probabilidad  $p$ , de forma que se obtiene una *red aleatoria*.

La distribución de grado de una red aleatoria finita sigue una distribución binomial

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

De forma que cuando  $N \rightarrow \infty$  la distribución tiende a una distribución de Poisson

$$P(k) = e^{-\langle k \rangle} \frac{\langle k \rangle^k}{k!}$$

Veamos qué implica este resultado. El número medio de amigos que un usuario estadounidense tiene en Facebook es de 350[6] siendo el número total de usuarios de ese país  $\sim 2 \cdot 10^8$ . Así, si las relaciones sociales siguiesen el patrón de una red aleatoria, la probabilidad de que un usuario tuviese 100 amigos sería de  $\sim 5 \cdot 10^{-48}$  y la de tuviese 600 amigos  $\sim 4 \cdot 10^{-26}$ . En otras palabras, en una red aleatoria la mayoría de los vértices tienen aproximadamente el mismo grado. Esto contradice claramente nuestra experiencia, pues es comprensible suponer que una persona famosa tendrá bastantes más amigos que la media.

Ante este resultado surge una pregunta, ¿es este modelo relevante para sistemas reales? La respuesta es no[7]. A pesar de ello se mantuvo vigente hasta 1999 debido a la carencia de los datos necesarios para comprobar experimentalmente su validez. No obstante, a día de hoy se sigue empleando a nivel académico debido a su sencillez.

El punto de inflexión se produjo en 1999. En este año Barabási et al. realizaron un mapa de Internet donde los nodos eran las páginas web y los links los enlaces entre unas y otras. El resultado fue que el 80% de los nodos tenían 4 links o menos, mientras que un 0.01% de ellos enlazaban con más de 1000 páginas, todo lo contrario a lo que uno encuentra en una red aleatoria[8]. Tras repetir el análisis con otras redes llegaron a la conclusión de que la probabilidad de que un vértice de una red real interaccione con otros  $k$  vértices (i.e. que tenga grado  $k$ ) decae siguiendo una ley de potencias o *power-law*[9],

$$P(k) \sim k^{-\alpha}$$

A pesar de su simplicidad, este tipo de distribución tiene algunas propiedades muy interesantes. Comencemos estudiando el valor de la constante de normalización. Sabemos que toda distribución de probabilidad tiene que cumplir que

$$\sum_{k=0}^{\infty} P(k) = 1$$

Podemos descartar el valor de  $k = 0$  ya que si un vértice no está unido a ningún otro podemos considerar que no forma parte de la red. Así, tendremos que

$$\sum_{k=1}^{\infty} C k^{-\alpha} = 1 \quad \Rightarrow \quad C = \frac{1}{\sum_{k=1}^{\infty} k^{-\alpha}} = \frac{1}{\zeta(\alpha)}$$

donde  $\zeta(\alpha)$  es la función zeta de Riemann. Con esto la distribución queda

$$P(k) = \frac{k^{-\alpha}}{\zeta(\alpha)}$$

En ocasiones nos va a interesar analizar la distribución a partir de un cierto valor  $k_{\min}$ , de forma que

$$P(k) = \frac{k^{-\alpha}}{\sum_{k=k_{\min}}^{\infty} k^{-\alpha}} = \frac{k^{-\alpha}}{\zeta(\alpha, k_{\min})}$$

donde  $\zeta(\alpha, k_{\min})$  es la función zeta generalizada.

Resulta razonable aproximar las sumas anteriores por integrales, ya que la power-law varía lentamente para grandes valores de  $k$ . Así, suponiendo que  $\alpha > 1$  para que la integral sea convergente, tendremos que

$$C \approx \frac{1}{\int_{k_{\min}}^{\infty} k^{-\alpha} dk} = (\alpha - 1) k_{\min}^{\alpha-1} \quad \Rightarrow \quad P(k) \approx \frac{\alpha - 1}{k_{\min}^{1-\alpha}} k^{-\alpha}$$

Una propiedad muy importante de esta distribución aparece al estudiar los momentos. El momento  $m$  de la distribución vendrá dado por

$$\langle k^m \rangle = \sum_{k=0}^{\infty} k^m P(k) \approx C \int_{k_{\min}}^{\infty} k^{m-\alpha} dk = \frac{(\alpha - 1) k_{\min}^{\alpha-1}}{m - \alpha + 1} [k^{m-\alpha+1}]_{k_{\min}}^{\infty}$$

Para evitar que la integral diverja será necesario que  $m - \alpha + 1 > 0 \Rightarrow \alpha > m + 1$ . Así, el primer momento será finito si y solo si  $\alpha \geq 2$ ; mientras que el segundo lo será si y solo si  $\alpha \geq 3$ . Como veremos más adelante, la importancia de este resultado está en que la mayoría de las redes reales que siguen una power-law tienen un exponente en el rango  $2 < \alpha \leq 3$ . Esto significa que en el límite en el que  $N \rightarrow \infty$  el grado medio de estas redes carece de sentido, pues las fluctuaciones serán demasiado grandes (la varianza es infinita).

## 2.4 Redes *scale-free*

Las redes cuya distribución de grado sigue una power-law se conocen como redes libres de escala o *scale-free*. El origen de este nombre se encuentra en el hecho de que la estructura de una red de este tipo no cambia aunque escalemos el tamaño, es decir, posee invariancia de escala.

Una función  $f(x)$  se dice invariante de escala si cumple que

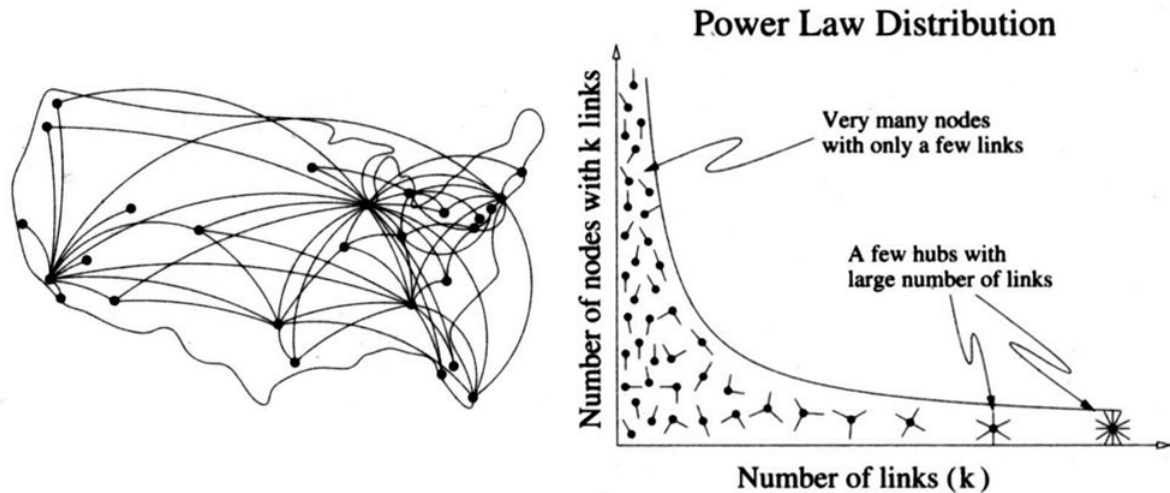
$$f(\lambda x) = \lambda^{\Delta} f(x)$$

para cierto valor de  $\Delta$  y para todo factor de escala  $\lambda$ . Está claro que la distribución power-law,  $P(k) \sim k^{-\alpha}$ , es invariante de escala con  $\Delta = -\alpha$ .

Esta propiedad aparece en numerosos aspectos de la física. Por ejemplo, en la transición de fase de un sistema magnético la longitud de correlación, el calor específico, la susceptibilidad y la magnetización son

invariantes de escala [10, p.223]. Es más, juega un papel fundamental en el funcionamiento de los discos duros[10, p.539].

Una de las principales características de las redes scale-free es la presencia de unos pocos nodos muy conectados (*hubs*) y de una gran cantidad muy poco conectados (figura 2), al contrario de lo que sucedía en las redes aleatorias donde todos tenían aproximadamente el mismo grado. Si analizamos la forma en la que se generan las redes no resulta sorprendente que la mayoría de ellas sigan este esquema.



**Figura 2:** Esquema de la red de aeropuertos de los Estados Unidos[7].

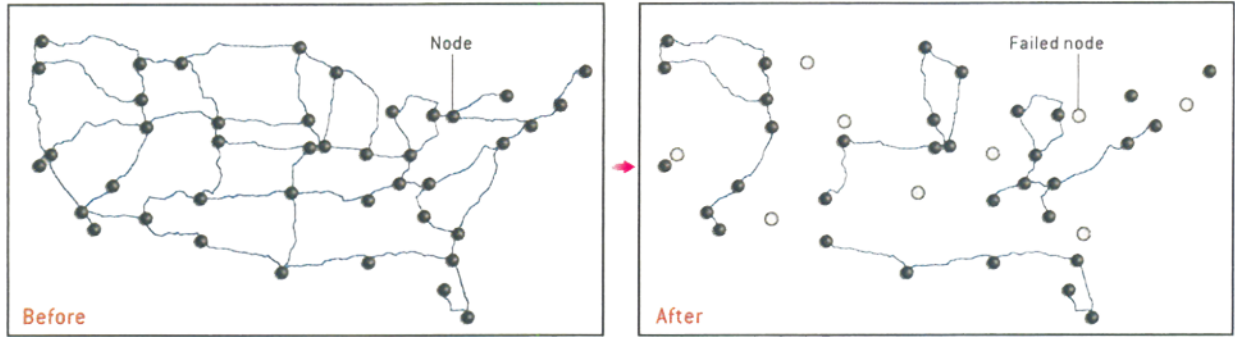
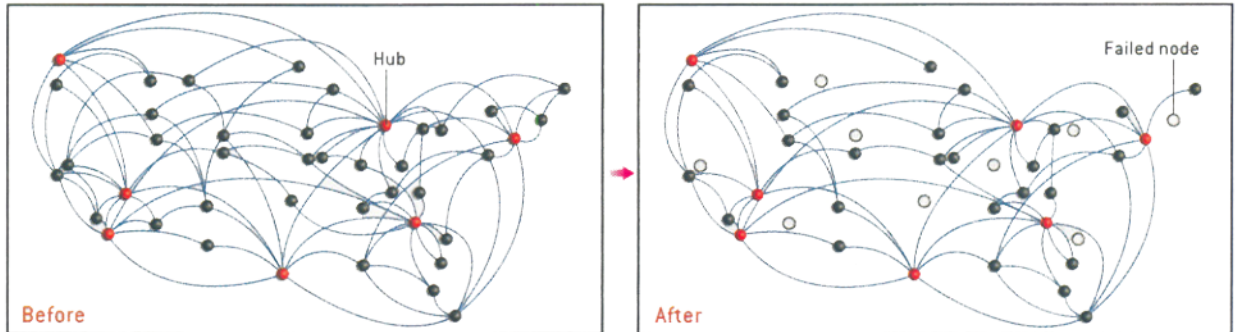
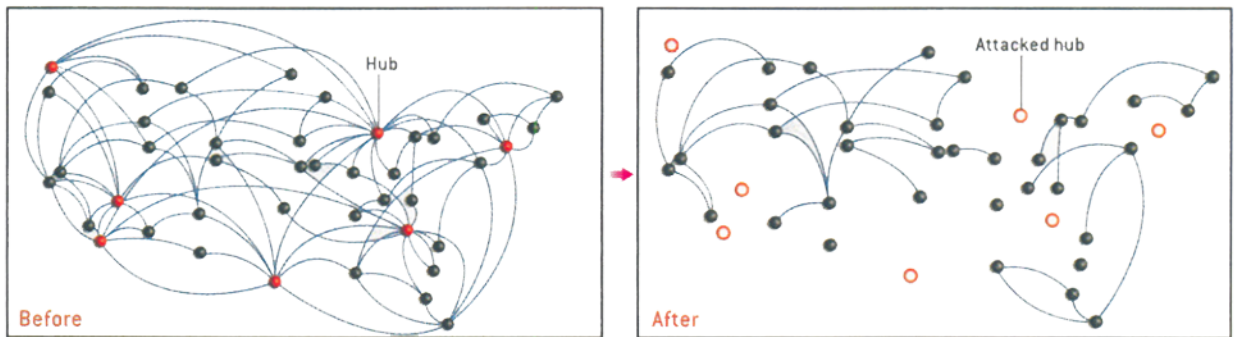
Pensemos por simplicidad en la red de páginas web. En el modelo ER se partía de una situación en la que los  $N$  nodos estaban presentes y se unían entre sí con una probabilidad constante  $p$ . Sin embargo, mientras que en 1990 existía una página web, a día de hoy hay varios millones[11]. Además, no todas ellas son iguales. Está claro que lo más probable es que una nueva página enlace con alguna muy conocida como Google o Youtube, de forma que el enlace es *preferencial*.

Así, estos dos mecanismos, el crecimiento y el enlace preferencial, nos permiten explicar la existencia de hubs de una manera sencilla: “*los ricos se vuelven cada vez más ricos*” [12]. Pero, ¿cuál es la importancia de los hubs?

Supongamos que tenemos una red aleatoria y que eliminamos de ella una fracción de nodos elegidos de manera aleatoria. Es lo que sucede, por ejemplo, si en la red de routers se producen averías que inutilizan una pequeña cantidad de ellos. Debido a la distribución uniforme de los grados, es bastante probable que la red se vea muy afectada incluso puede que se divida en regiones no conectadas entre sí como se refleja en la figura 3(a).

Si, en cambio, la red es scale-free, lo más probable es que los nodos eliminados sean de muy poco grado (ya que son los más abundantes) y que los hubs se mantengan intactos. Al tener un grado tan alto, su presencia permitirá que los nodos restantes de la red sigan estando todos conectados entre sí, figura 3(b). Se dice que las redes scale-free son muy robustas ante accidentes. Sin embargo, por este mismo motivo está claro que un ataque dirigido a los hubs dañará enormemente a la red, figura 3(c).

Por tanto, los hubs forman una parte fundamental de la red porque cualquier cambio en ellos se transmitirá a un gran número de nodos.

**Random Network, Accidental Node Failure (a)****Scale-Free Network, Accidental Node Failure (b)****Scale-Free Network, Attack on Hubs (c)****Figura 3:** Efecto de la eliminación de una fracción de los nodos en una red aleatoria y en una scale-free [12].

Para concluir, en la tabla 1 se presentan algunos ejemplos de redes scale-free. Pese a tener orígenes muy diferentes, todas ellas tienen un exponente entre 2 y 3.

Red	$N$	$\langle k \rangle$	$\alpha$
Internet (routers)	3700 - 10500	3,6 - 4,1	2,1
WWW	$2 \cdot 10^8$	15	2,1
Actores de cine	225,226	61	2,3
Proteínas de la levadura	1870	2,40	2,4
Metabólica	765	4,82	2,2
Co-ocurrencia de palabras	460902	34,93	2,7

**Tabla 1:** Número de nodos, grado medio y exponente de la distribución de varias redes[13, 1].

En concreto, en la figura 4 se representa la red de routers a escala internacional. Se observan claramente grandes hubs cerca de las grandes capitales mundiales y un mayor número de nodos poco conectados.



**Figura 4:** Mapa de las conexiones de Internet entre continentes[14].

### 3 Difusión de epidemias

Se dice que se ha producido una *epidemia* cuando el número de individuos afectados por una enfermedad en una población es superior al que se espera en condiciones normales. En ocasiones se emplea el término *outbreak* o *brote* debido a que tiene una connotación menos negativa y no causa tanta alarma en la población, aunque desde el punto de vista epidemiológico *epidemia* y *outbreak* significan exactamente lo mismo [15].

La difusión de una enfermedad depende tanto de las propiedades del organismo patógeno<sup>a</sup> como de las redes de contactos entre individuos. En una red de contactos hay un nodo por cada persona y se establece un link si dos personas entran en contacto de una forma que permita la propagación de la infección. Esta puntualización es importante ya que mientras que la gripe o la tuberculosis se transmiten por el aire, los parásitos lo hacen por contacto físico y el VIH lo hace durante las relaciones sexuales. Así, será crucial conocer las propiedades de la red subyacente para entender la propagación de la enfermedad[16].

Aunque el problema parece complicado de abordar debido a la gran cantidad de factores diferentes involucrados, vamos a ver cómo estableciendo ciertas simplificaciones vamos a ser capaces de desarrollar modelos muy sencillos pero de gran poder predictivo.

#### 3.1 Modelos compartimentales

En primer lugar suponemos que la población puede dividirse en diferentes clases o *compartimentos* en función de la etapa de la enfermedad en la que se encuentran. Este tipo de modelos se denominan *compartimentales*[17].

El caso más sencillo es considerar dos clases: *susceptibles* ( $S$ ) e *infectados* ( $I$ ). Este es el caso típico de las enfermedades bacterianas, o está presente la bacteria o no lo está. Si pensamos en virus, en cambio, debemos considerar tres clases: susceptibles ( $S$ ), infectados ( $I$ ) y *recuperados* o inmunes<sup>b</sup> ( $R$ ). Esto se debe a que al superar una infección vírica los anticuerpos permanecen en el organismo durante un tiempo que va desde uno o dos años en el caso de la gripe hasta toda la vida en el de la varicela. Durante este periodo el individuo no puede volver a contraer la misma infección, por lo que se le considera inmune.

El siguiente paso es modelar el comportamiento del patógeno. Para ello emplearemos tres parámetros:

- $\beta$ : probabilidad por unidad de tiempo de que una persona infectada contagie a una susceptible con la que tiene contacto.
- $\mu$ : probabilidad por unidad de tiempo de que una persona infectada se recupere.
- $\gamma$ : probabilidad por unidad de tiempo de que una persona inmune se vuelva susceptible.

Estos parámetros se determinan de una forma completamente experimental, lo que tiene consecuencias positivas y negativas. Por una parte es un modelo muy versátil ya que no es necesario conocer completamente el patógeno, basta con saber la variación del número de susceptibles, infectados y recuperados en función del tiempo. Además, permite agregar de forma sencilla efectos como puede ser la estacionalidad (hay enfermedades que son más contagiosas en invierno que en verano). Por otra parte, esto implica que los valores numéricos solo se pueden obtener una vez que la epidemia ha terminado. Como veremos esto es un gran problema a la hora de tratar de reproducir la evolución de una nueva epidemia a tiempo real.

Con estos ingredientes ya es posible construir los modelos compartimentales más sencillos:

---

<sup>a</sup>Agente capaz de producir una enfermedad en un ser vivo.

<sup>b</sup>Realmente esta clase denota aquellos individuos que ya no pueden enfermar, por lo que puede incluir también a las personas que hayan fallecido y no solo a los que se han inmunizado. Por este motivo en inglés se prefiere emplear el término *removed* en lugar de *recovered* aunque la mayoría de las veces se usan indistintamente.



- Modelo *SIS*: en este modelo los individuos susceptibles pasan a infectados con una probabilidad  $\beta$  los cuales se recuperan con una probabilidad  $\mu$  volviendo al estado susceptible (figura 5a).

$$S \xrightarrow{\beta} I \xrightarrow{\mu} S$$

- Modelo *SIR*: en este caso los individuos que se recuperan se vuelven inmunes para siempre a la enfermedad (figura 5b).

$$S \xrightarrow{\beta} I \xrightarrow{\mu} R$$

- Modelo *SIRS*: si la inmunidad no dura toda la vida los individuos pueden volver al estado susceptible con una probabilidad  $\gamma$ .

$$S \xrightarrow{\beta} I \xrightarrow{\mu} R \xrightarrow{\gamma} S$$



**Figura 5:** En los modelos compartimentales la población se divide en compartimentos en función del estado de la enfermedad en el que se encuentran. Ilustraciones adaptadas de [18] y [19].

El gran potencial de los modelos compartimentales es que se pueden ampliar muy fácilmente. Para abarcar más enfermedades sería necesario añadir nuevos compartimentos como *latente* ( $E$ ), donde el individuo está infectado pero todavía no es contagioso; o *asintomático* ( $A$ ), estado muy peligroso ya que es contagioso pero el individuo no es consciente. Se pueden incluir nacimientos y defunciones en el modelo suponiendo la población variable, aunque en la mayoría de los casos prácticos la difusión es tan rápida que no se ve afectada si se desprecian estos efectos. Cabe destacar, además, que hemos supuesto que la población de cada compartimento es homogénea. Para introducir el factor de la edad, por ejemplo, bastaría con hacer que los parámetros de la infección dependiesen de la edad del individuo.

No obstante, normalmente se emplean los modelos básicos como una primera aproximación y luego se amplían en función de las características de la enfermedad que queramos estudiar. Por ello, en adelante nos centraremos en el modelo *SIRS* para obtener una visión general sobre la difusión.

### 3.2 *Homogeneous mixing*

Los modelos desarrollados hasta ahora tienen en cuenta las propiedades de los individuos y de la enfermedad pero no hemos hablado en ningún momento de un elemento crucial, la red de contactos.

Inicialmente vamos a suponer *homogeneous mixing*, es decir, que todos los individuos pueden tener contacto entre sí. Al hacer esto realmente estamos eliminando el papel de la red, cosa que va en contra de los resultados experimentales. Por ejemplo, la difusión de la mayoría de las enfermedades de transmisión sexual no puede entenderse sin tener en cuenta la distribución espacial y la heterogeneidad de la actividad sexual en una población[20]. No obstante, nos servirá como punto de partida y nos permitirá comprobar las diferencias que supone aplicar el modelo a una red.

Sea  $S(t)$  el número de individuos que son susceptibles a tiempo  $t$ ,  $I(t)$  el número de infectados y  $R(t)$  el número de inmunes. Técnicamente, como el proceso de difusión de una enfermedad es aleatorio, estos

números no están determinados unívocamente. Si la enfermedad volviese a extenderse entre la población, aunque partiésemos de las mismas condiciones iniciales, los resultados serían ligeramente diferentes. Sin embargo, por ahora supondremos que estamos ante un proceso determinista<sup>c</sup>.

Supongamos que en cada paso de tiempo cada infectado entra en contacto con una persona. La probabilidad de que ésta pertenezca al estado susceptible será  $S(t)/N$  donde  $N$  es el número total de personas. Así, por unidad de tiempo, la probabilidad de que un infectado contagie a uno sano será  $\beta S(t)/N$ . Como el número total de enfermos es  $I(t)$ , en cada paso de tiempo se estarán generando  $\beta I(t)S(t)/N$  nuevas infecciones. De forma similar,  $\mu I(t)$  personas pasarán al estado inmune y  $\gamma R(t)$  volverán a ser susceptibles. Si escribimos esto en forma de ecuación,

$$S(t + \delta t) = S(t) - \beta \frac{S(t)I(t)}{N} \delta t + \gamma R(t) \delta t \Rightarrow \frac{dS(t)}{dt} = \frac{S(t + \delta t) - S(t)}{\delta t} = -\beta \frac{S(t)I(t)}{N} + \gamma R(t)$$

Procediendo de forma análoga con los otros dos estados y normalizando ( $s \equiv S/N$ ,  $i \equiv I/N$  y  $r \equiv R/N$ ) se obtiene el sistema de ecuaciones diferenciales que determina la evolución de una epidemia en el modelo SIRS bajo la hipótesis de homogeneous mixing,

$$\frac{ds(t)}{dt} = -\beta s(t)i(t) + \gamma r(t) \quad (3.1)$$

$$\frac{di(t)}{dt} = \beta s(t)i(t) - \mu i(t) \quad (3.2)$$

$$\frac{dr(t)}{dt} = \mu i(t) - \gamma r(t) \quad (3.3)$$

Realmente una de las ecuaciones es redundante ya que  $S + I + R = N \Rightarrow s + i + r = 1$ , por lo que el sistema se puede resumir en dos ecuaciones diferenciales. A pesar de ello, el modelo SIRS no se puede resolver analíticamente en el caso general.

La predicción más significativa de estas ecuaciones es la existencia de un *epidemic threshold* o umbral epidémico[21]. Es decir, existe una cierta combinación de las condiciones iniciales que determina si es posible o no que se produzca una epidemia. Veámoslo.

Tomemos como condiciones iniciales  $s(0) = s_0 > 0$ ,  $i(0) = i_0 > 0$  y  $r(0) = 0$ . Así, en  $t = 0$  la ecuación (3.2) queda

$$\left. \frac{di}{dt} \right|_{t=0} = i_0(\beta s_0 - \mu)$$

Para que se pueda producir una epidemia será necesario que, como mínimo, inicialmente el número de infectados se incremente, es decir

$$\begin{aligned} \left. \frac{di}{dt} \right|_{t=0} > 0 &\Rightarrow \beta s_0 - \mu > 0 \\ &\Rightarrow s_0 > \frac{\mu}{\beta} \end{aligned} \quad (3.4)$$

Pero  $s \leq 1 \Rightarrow s_0 \leq 1$ . Teniendo en cuenta esta condición,  $1 \geq s_0 > \frac{\mu}{\beta}$ , por lo que

$$\lambda \equiv \frac{\beta}{\mu} > 1 \quad (3.5)$$

Esto nos permite definir el umbral epidémico como  $\lambda_c \equiv 1$  de forma que si  $\lambda > \lambda_c$  la propagación de la epidemia es posible, mientras que si  $\lambda < \lambda_c$  el brote inicial desaparece antes de infectar a una cantidad no despreciable de la población.

<sup>c</sup>Podemos considerar que los resultados que vamos a obtener son los que conseguiríamos si reptiésemos el proceso muchas veces con las mismas condiciones iniciales y después hicieramos el promedio[1].



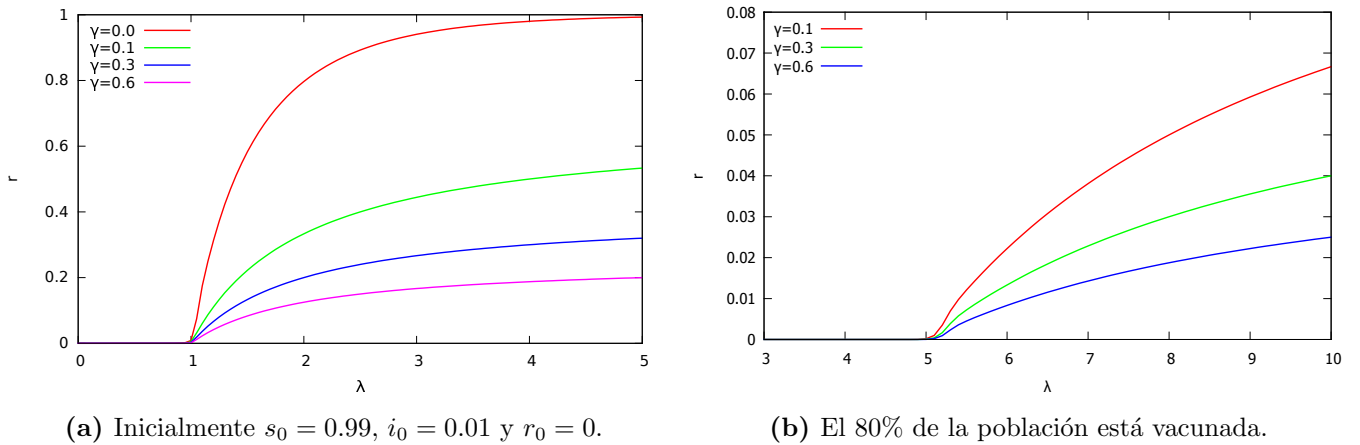
Volviendo a la ecuación (3.4) definimos

$$R_0 \equiv \frac{s_0 \beta}{\mu} = s_0 \lambda \quad (3.6)$$

de forma que si  $R_0 > 1$  la epidemia es posible y en caso contrario desaparece. En epidemiología clásica a esta cantidad se le denomina *basic reproductive number* y da cuenta del número de infecciones secundarias que se producen cuando se introduce un individuo infectado en una población donde todos son susceptibles. Esta condición es más restrictiva que la del umbral epidémico, aunque como normalmente  $s_0 \approx 1$  podemos considerar que  $R_0 \approx \lambda$  [22].

Detengámonos un momento en esta expresión. Los parámetros  $\beta$  y  $\mu$  dependen principalmente de la enfermedad, por lo que será difícil modificarlos. Sin embargo, si conseguimos reducir el valor de  $s_0$  podemos llegar a evitar que se produzca la epidemia, cosa que podemos realizar mediante la vacunación.

Veamos un par de ejemplos reales. La viruela tiene aproximadamente un umbral epidémico  $\lambda = 5$ . Si queremos eliminar la enfermedad, necesitaremos que  $R_0 < 1 \Rightarrow s_0 < 1/\lambda = 0.2$ . Así, si conseguimos vacunar al 80% de la población seremos capaces de eliminar la enfermedad como, de hecho, se logró en los años 70. La rubeola, en cambio, tiene un umbral epidémico de  $\lambda = 18$ , lo que implica que necesitamos vacunar al  $\sim 94\%$  de la población, algo difícil pero a simple vista no imposible. El problema de ésta y otras enfermedades con valores tan altos del umbral epidémico es que en toda población hay una serie de individuos que no pueden recibir la vacuna: alérgicos, enfermos, personas con el sistema inmune debilitado, ancianos... Y en ocasiones son los suficientes como para que no se pueda reducir el valor de  $s_0$  hasta los niveles que previenen la epidemia [18].



**Figura 6:** Solución numérica de las ecuaciones (3.1), (3.2) y (3.3). Se representa la fracción de población recuperada de la enfermedad en estado estacionario para varios valores de  $\lambda$  y  $\gamma$  manteniendo  $\mu = 0.2$ . La condición necesaria para que  $r \neq 0$  es que  $i(t \rightarrow \infty) \neq 0$  si  $\gamma \neq 0$ , mientras que si  $\gamma = 0$  basta con que se haya producido un brote epidémico.

Para finalizar con el análisis de la hipótesis de homogeneous mixing podemos comprobar la validez de estas predicciones resolviendo numéricamente el modelo, figura 6.

La expresión del umbral epidémico que hemos obtenido depende únicamente de  $\beta$  y  $\mu$ , por lo que deberíamos obtener  $\lambda_c = 1$  independientemente del valor de  $\gamma$  siempre que  $s_0 \approx 1$ . En la figura 6a se aprecia claramente la validez de este resultado. Cabe resaltar que en el caso en el que  $\gamma = 0$  se recupera el modelo SIR, de ahí que exista tanta diferencia entre dicha curva y el resto, aunque vemos que el umbral es el mismo que en el modelo SIRS. Si considerásemos el modelo SIS también obtendríamos el mismo resultado, ya que el análisis se ha efectuado sobre la ecuación (3.2) y ésta no depende de la presencia de la fase inmune.

En la figura 6b se puede comprobar cómo cambian los resultados al partir de una situación en la que solo el  $\sim 20\%$  de la población es susceptible de contraer la enfermedad. Bajo estas condiciones ya no es

válida la aproximación  $R_0 \approx \lambda$ , por lo que  $\lambda_c$  deberá ser distinto de uno. Según la ecuación (3.6) para que  $R_0 > 1$  será necesario que  $\lambda > 5$ , es decir,  $\lambda_c = 5$ . Vemos claramente como los resultados numéricos están de acuerdo con esta condición.

### 3.3 Difusión en redes *scale-free*

La suposición de que es posible entrar en contacto con cualquier persona de una población contradice claramente nuestra experiencia diaria. En el mundo real, la mayoría de las personas tienen un círculo de conocidos con el que se relaciona normalmente: familia, amigos, compañeros de trabajo... Matemáticamente podemos representar al conjunto de contactos potenciales de una persona mediante una red. Como veremos a continuación, la estructura de dicha red puede tener un gran efecto sobre la forma en la que se difunde la epidemia.

La diferencia principal respecto al caso de homogeneous mixing es que un nodo solo podrá infectarse si uno de sus vecinos está infectado. Debido a la heterogeneidad de la red esto implica que hay que considerar la evolución de cada nodo por separado. Para poder analizar este sistema analíticamente vamos a plantear una hipótesis de campo medio: los nodos que poseen el mismo grado se comportan de la misma manera[23][2].

Por esta hipótesis, si el estado de un nodo viene dado por  $\epsilon_i$  tendremos que  $\epsilon_i = \epsilon_j$  si el grado de los nodos  $i$  y  $j$  es el mismo. Denotando por  $D_k$  al conjunto de nodos de grado  $k$ , esto es equivalente a decir que  $i, j \in D_k$ . En otras palabras, si la probabilidad de que el nodo  $i$  pertenezca al estado susceptible es  $s_i(t)$ , la probabilidad de que los nodos de grado  $k$  pertenezcan al estado susceptible será

$$s_k(t) = \frac{1}{N_k} \sum_{i \in D_k} s_i(t) = s_i(t)$$

por lo que en vez de plantear las ecuaciones de evolución de cada nodo, basta con plantear una por grado.

Denotemos por  $\Theta(t)$  a la probabilidad de que un link de un nodo apunte a un nodo infectado. Un nodo de grado  $k$  tendrá  $k$  links, de forma que la probabilidad de que esté en contacto con uno infectado será  $k\Theta(t)$ . Así, tendremos que

$$\frac{ds_k(t)}{dt} = -\beta s_k(t)k\Theta(t) + \gamma r_k(t) \quad (3.7)$$

$$\frac{di_k(t)}{dt} = \beta s_k(t)k\Theta(t) - \mu i_k(t) \quad (3.8)$$

$$\frac{dr_k(t)}{dt} = \mu i_k(t) - \gamma r_k(t) \quad (3.9)$$

donde  $s_k + i_k + r_k = 1$ .

Falta determinar una expresión para  $\Theta(t)$ . La probabilidad de que un link apunte a un nodo con  $k$  links será proporcional a  $kP(k)$ . Por tanto, la probabilidad de que un link esté unido a un nodo infectado será

$$\Theta(t) = \frac{\sum_k kP(k)\rho_k(t)}{\sum_s sP(s)} = \frac{\sum_k kP(k)\rho_k(t)}{\langle k \rangle} \quad (3.10)$$

En esta aproximación estamos considerando que la probabilidad de que un link se una a un nodo infectado es independiente del nodo de partida. Existen modelos más detallados en los que se tiene en cuenta la probabilidad condicional  $P(k|k')$  de que un nodo de grado  $k'$  se una con uno de grado  $k$ . Sin embargo, esta aproximación es suficiente para obtener las condiciones que determinan si se va a producir o no una epidemia[24].

Para determinar el punto crítico que diferencia el estado epidémico del no epidémico vamos a suponer que  $r \approx 0$ . Esto es cierto en el modelo SIS para todo tiempo  $t$ , mientras que en el SIR y en el SIRS lo es solo al inicio del proceso. No obstante, los resultados serán válidos para todos ellos ya que si no se produce un brote al inicio del proceso es imposible que se produzca después.

Bajo estas suposiciones la condición necesaria para que se pueda producir un brote es

$$\left. \frac{di_k(t)}{dt} \right|_{t=0} > 0 \quad \Rightarrow \quad \beta(1 - i_k(t))k\Theta(t) > \mu i_k(t)$$

para algún valor de  $k$ . Despejando  $i_k$ ,

$$i_k < \frac{\beta k \Theta}{\mu + \beta k \Theta}$$

Y sustituyendo esta ecuación en la definición de  $\Theta$  (3.10),

$$\Theta > \frac{1}{\langle k \rangle} \sum_k \frac{k^2 P(k) \beta \Theta}{\mu + \beta k \Theta} \equiv f(\Theta)$$

A la vista de la figura 7 para que exista algún valor de  $\Theta$  que cumpla la condición es necesario que

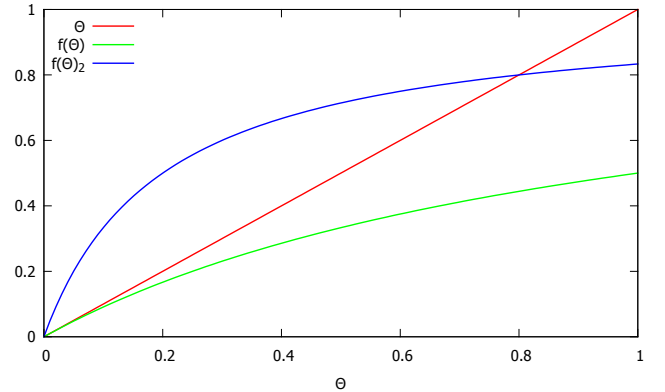
$$\left. \frac{df(\Theta)}{d\Theta} \right|_{\Theta=0} > 1$$

Es decir,

$$\frac{1}{\langle k \rangle} \sum_k k^2 P(k) \cdot \frac{\beta}{\mu} > 1$$

Recordando que  $\lambda = \beta/\mu$  llegamos a que

$$\lambda > \frac{\langle k \rangle}{\langle k^2 \rangle}$$



**Figura 7:** Representación gráfica de  $y = \Theta$  e  $y = f(\Theta)$  con diferentes valores de  $k\beta$ .

Por lo que el umbral epidémico para la difusión en una red es

$$\lambda_c = \frac{\langle k \rangle}{\langle k^2 \rangle} \quad (3.11)$$

Además, de forma similar al caso anterior, podemos definir el basic reproductive number como

$$R_0 = \frac{\beta}{\mu} \frac{\langle k^2 \rangle}{\langle k \rangle} = \lambda \frac{\langle k^2 \rangle}{\langle k \rangle} \quad (3.12)$$

Ambas expresiones son muy similares a las obtenidas suponiendo homogeneous mixing, (3.5) y (3.6). Sin embargo, en esta ocasión tenemos términos adicionales relacionados con la estructura de la red, como era de esperar. Veamos qué implican estos resultados.

En el caso de una red aleatoria se puede demostrar[7] que

$$\langle k^2 \rangle = \langle k \rangle \left( 1 + \frac{\langle k \rangle}{N} \right)$$

En el límite  $N \rightarrow \infty$ , por tanto,  $\langle k^2 \rangle = \langle k \rangle$  por lo que el punto crítico queda

$$\lambda_{\text{rand}} = \frac{1}{\langle k \rangle} \quad (3.13)$$

De manera que cuanto más conectada esté la red menor será el punto crítico y más fácil será que se produzca una epidemia. No obstante, el resultado realmente interesante surge al considerar una red scale-free que, además, son las redes que nos encontramos en el mundo real.

Recordemos que en la sección 2.3 habíamos determinado que en una red scale-free  $\langle k^2 \rangle = \infty$  cuando  $2 < \alpha \leq 3$ . Además, en la 2.4 veíamos como la mayoría de las redes reales cumplen esta condición. Por ejemplo, la red de routers tiene un exponente  $\alpha = 2,1^d$ .

Aplicando este resultado a la ecuación (3.11) obtenemos algo sorprendente,

$$\lambda_c = \frac{\langle k \rangle}{\langle k^2 \rangle} = 0$$

Es decir, en una red scale-free con un exponente tal que  $2 < \alpha \leq 3$  no existe umbral epidémico. Por tanto, existe una probabilidad no nula de que se produzca una epidemia para cualquier valor de  $\beta$  y  $\mu$  (siempre que  $\beta > 0$ ). Otra manera de verlo es mediante el basic reproductive number. Por la ecuación (3.12) si  $\langle k^2 \rangle = \infty$  entonces  $R_0 > 1$ .

Pensemos de nuevo en la vacunación. En este caso no nos basta con reducir la población inicial susceptible ya que el que el umbral sea cero depende de la estructura de la red.

Es en este punto donde cobra sentido la discusión sobre la importancia de los hubs de la página 10. Veíamos que, aunque una red scale-free era muy resistente a la pérdida (inmunización) de nodos aleatorios, era muy débil frente a ataques dirigidos (vacunación).

De esta forma, lo que era una desventaja podemos convertirlo en algo positivo. Si conseguimos localizar a los grandes hubs de la red y los inmunizamos, seremos capaces de hacer que la red deje de ser conexas y detendremos la difusión de la epidemia. Además, el eliminar a los nodos de mayor grado hará que la diferencia entre los de menor y mayor grado no sea tan grande, incluso es posible que  $\langle k^2 \rangle$  tome un valor finito. En consecuencia, se reestablecerá un umbral finito y será posible detener la epidemia aunque la red siga siendo conexas.

Para finalizar, vamos a comprobar si estas predicciones se ajustan a las simulaciones. En este caso vamos a emplear simulaciones de Monte Carlo, lo que nos va a permitir introducir el factor de aleatoriedad inherente a este tipo de procesos (recordemos que los parámetros  $\beta$ ,  $\mu$  y  $\gamma$  son probabilidades). Estos efectos son especialmente importantes al inicio de la epidemia, cuando el número de infectados es todavía muy pequeño[26].

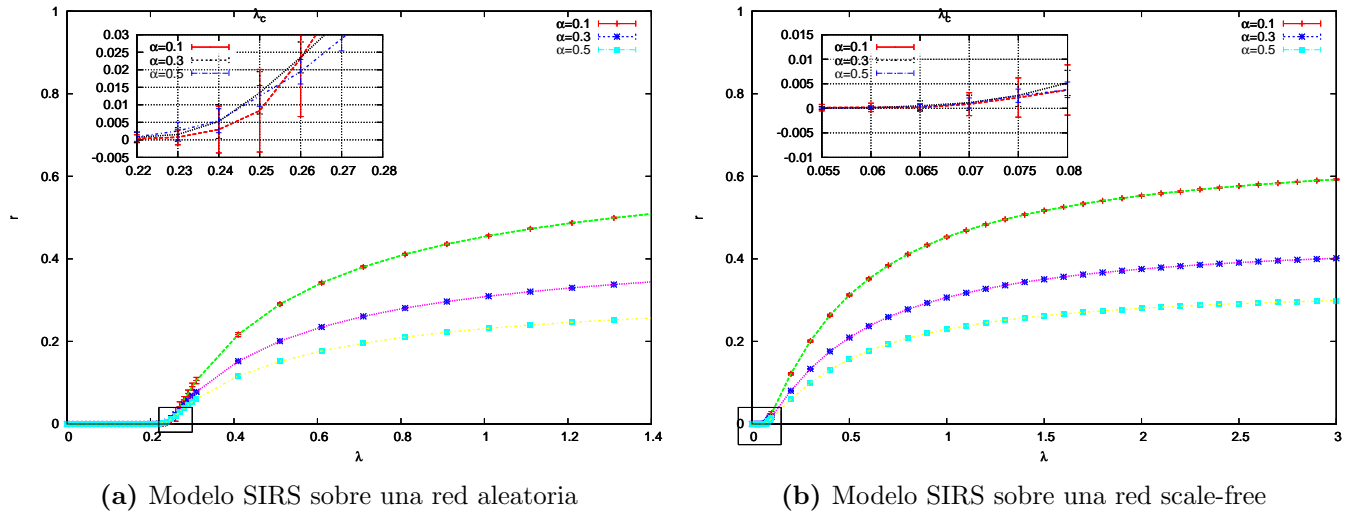
En primer lugar empleamos una red aleatoria con  $\langle k \rangle = 4.04$ . Si la predicción de la ecuación (3.13) es correcta, deberíamos obtener  $\lambda_c \approx 0.25$ . Como se puede comprobar en la figura 8a el punto crítico es precisamente  $\sim 0.25^e$ .

En la red scale-free, a pesar de tener un grado medio similar, los resultados son bastante diferentes. De acuerdo a la ecuación (3.11) idealmente deberíamos obtener el punto crítico en el 0, pero al usar una red finita obtenemos un resultado ligeramente diferente.

La red empleada posee  $\langle k^2 \rangle = 66.44$ , lejos de un valor infinito. No obstante, el valor es lo suficientemente grande como para introducir una gran diferencia respecto al caso anterior. Así, teniendo en cuenta este

<sup>d</sup>Cabe destacar que en la práctica la difusión de virus entre ordenadores y la de epidemias entre personas se puede tratar con los mismos modelos. La primera vez que se propuso aplicar un modelo epidémico a una red scale-free fue al tratar de explicar la persistencia de los virus en los ordenadores[25].

<sup>e</sup>El que la transición entre un estado y otro no sea abrupta se denomina efecto de tamaño finito y es una consecuencia del tamaño finito de la red[1].



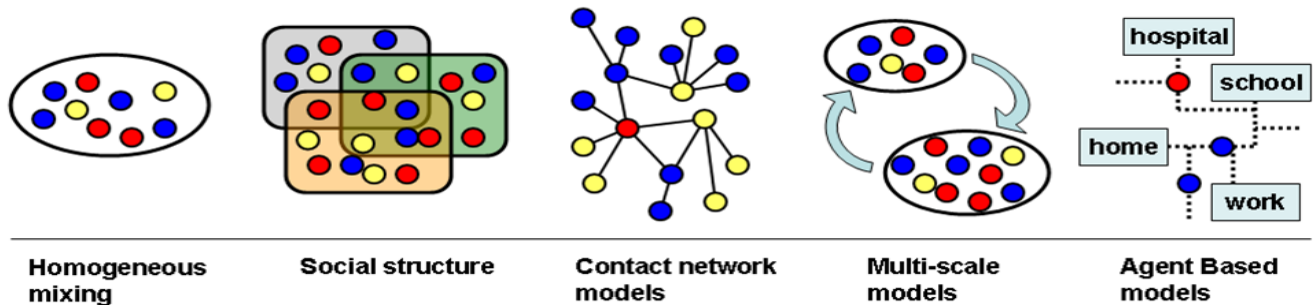
**Figura 8:** Simulación MC de la difusión de una epidemia según el modelo SIRS sobre redes de  $\sim 10.000$  nodos. La red aleatoria posee  $\langle k \rangle = 4.04$ , mientras que la scale-free tiene  $\langle k \rangle = 4.59$  y  $\langle k^2 \rangle = 66.44$

parámetro deberíamos obtener  $\lambda_c = 0.69$ . Podemos ver en la figura 8b como claramente este resultado concuerda con las simulaciones.

A pesar de la simplicidad del modelo, estos resultados concuerdan cualitativamente con la epidemiología de una gran cantidad de patógenos. No obstante, para realizar predicciones cuantitativas que puedan emplearse para evaluar qué políticas aplicar ante un brote y analizar el coste-beneficio de ellas es necesario refinar este modelo. Es en este contexto en el que surgen los *modelos metapoblacionales*[26].

## 4 Modelos metapoblacionales

Como sucede en muchos otros campos, a la hora de proponer un modelo debemos tener en cuenta que cuanto más sencillo sea, peor precisión tendrán las predicciones y viceversa. Por otra parte, el agregar más parámetros al modelo, además de dificultar su manejo, posee un problema inherente: será necesario contar con muchos más datos para poder determinarlos. Si no podemos estimar estos parámetros con gran precisión no nos servirá de nada complicar el modelo ya que sus predicciones pueden no tener menos incertidumbre que la que nos darían modelos más sencillos pero con parámetros bien determinados.



**Figura 9:** Diferentes estructuras empleadas en el estudio de la difusión de epidemias. Los colores representan el estado de cada nodo. De izquierda a derecha: el modelo más básico es considerar homogeneous mixing; también podemos tener en cuenta la estructura social (edad, género...); para modelar la heterogeneidad de las relaciones sociales se emplean redes de contacto; en los modelos multi-escala (entre los que se incluyen los metapoblacionales) se tiene en cuenta la interacción entre diferentes subpoblaciones (como podrían ser las ciudades) pero dentro de ellas se emplea homogeneous mixing; en los modelos agent-based se recrean los movimientos e interacciones de cada una de las personas individualmente con gran precisión[17].

En la figura 9 se representan esquemáticamente los diferentes modelos que podemos emplear. Hasta ahora nos hemos limitado a los tres primeros ya que, como vamos a ver a continuación, son ingredientes clave para el desarrollo de los siguientes, los modelos metapoblacionales.

### 4.1 La difusión en metapoblaciones: procesos reacción-difusión

Una metapoblación es una población subdividida en comunidades discretas unidas por conexiones que representan los posibles movimientos de los individuos. Dentro de cada comunidad la dinámica de transmisión de las enfermedades se describe mediante los modelos compartimentales estándar[27].

Matemáticamente, la difusión en redes metapoblacionales se estudia como un proceso de reacción-difusión (RD) sobre una red compleja. Así, consideramos que los nodos son ciudades y que las personas se pueden mover (difundir) de una a otra a través de los links. En la sección 2 hemos presentado todos los conceptos de la teoría de redes que vamos a necesitar, pero no hemos hablado todavía sobre los procesos RD.

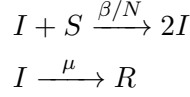
Los modelos RD se desarrollaron para explicar las reacciones químicas, aunque hoy en día tienen aplicaciones en muchos otros campos. En física, por ejemplo, podemos encontrar este tipo de procesos al estudiar la dinámica no-lineal del transporte en semiconductores[28].

A nivel microscópico, los procesos RD consisten en un conjunto de partículas que se difunde en el espacio y que están sujetas a una serie de reacciones que dependen del problema que se esté tratando (figura 10). Mientras que en los modelos RD fermiónicos se considera que hay un límite en el número de partículas que puede haber en cada nodo (por el principio de exclusión), en los modelos bosónicos no existe tal restricción[29].

Para modelar la difusión de una epidemia vamos a considerar un proceso RD bosónico en el cual las partículas serán las personas. Éstas se dividirán en tres categorías (susceptible, infectado o inmune) pudi-

endo pasar de una a otra mediante una “reacción”. Cabe destacar que el modelo que vamos a desarrollar es determinista, por lo que no es del todo válido para estudiar el inicio de la epidemia (en dicho punto las fluctuaciones estocásticas no son despreciables)[30].

Consideremos, por simplicidad, que podemos describir la epidemia mediante el modelo SIR. Las ecuaciones de las reacciones serán:



Reacciones que se producirán en cada población  $j$ . Bajo la hipótesis de homogeneous mixing (sección 3.2) la tasa de infección en cada nodo  $j$  la podemos denotar por  $\beta\Gamma_j$  donde

$$\Gamma_j = \frac{I_j S_j}{N_j}$$

Bajo la aproximación de campo medio (sección 3.3) tendremos que

$$I_k = \frac{1}{V_k} \sum I_j, \quad S_k = \frac{1}{V_k} \sum S_j$$

siendo  $V_k$  el número de nodos de grado  $k$ . En esta aproximación la tasa de infección será  $\beta\Gamma_k = \beta I_k S_k / N_k$ .

Podemos escribir la variación del número de infectados en cada subpoblación de grado  $k$  mediante una ecuación maestra,

$$I_k(t + \Delta t) - I_k(t) = W_k^+ - W_k^-$$

donde  $W_k^\pm$  representa el número de infectados que entran/abandonan los nodos de grado  $k$  tanto por contagio como por difusión. Si suponemos que el ritmo de difusión depende del grado del nodo,  $p_k$ , tendremos que

$$W_k^- = p_k I_k + (1 - p_k) \mu I_k$$

El primer término representa el número de individuos que abandonan el nodo y el segundo el número que se cura de los que se quedan. De forma similar,

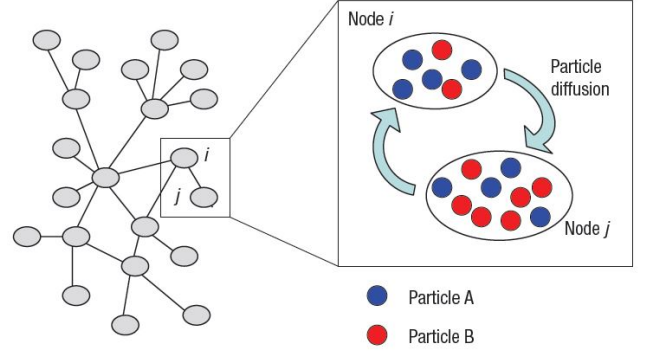
$$W_k^+ = (1 - p_k) \beta \Gamma_k + k \sum_{k'} P(k'|k) d_{k'k} [(1 - \mu) I_{k'} + \beta \Gamma_{k'}]$$

En este caso el primer término da cuenta del número de gente que se queda en su nodo y se infecta y el segundo el número de infectados que llegan de otros nodos. En efecto,  $d_{k'k}$  es la fracción de infectados que van de un nodo de grado  $k'$  a uno de grado  $k$ . El número de infectados del nodo de grado  $k'$  será  $[(1 - \mu) I_{k'} + \beta \Gamma_{k'}]$  (los que no se recuperan más los nuevos infectados). El otro término del sumatorio,  $P(k'|k)$ , es la probabilidad de que haya un link entre un nodo de grado  $k$  y otro de grado  $k'$ . Así, al sumar sobre todos los grados, obtendremos el número total de infectados que pueden llegar a través de un link al nodo de grado  $k$ . Por último, multiplicamos esta cantidad por el número de links del nodo,  $k$ .

Introduciendo estas definiciones llegamos a la ecuación diferencial que determina la evolución del número de infectados presentes en los nodos de grado  $k$ ,

$$\partial_t I_k = -p_k I_k + (1 - p_k) [-\mu I_k + \beta \Gamma_k] + k \sum_{k'} P(k'|k) d_{k'k} [(1 - \mu) I_{k'} + \beta \Gamma_{k'}]$$

De forma análoga se pueden obtener las expresiones para la evolución de  $S_k$  y  $R_k$ . Con estas ecuaciones quedaría determinada la evolución del sistema dadas las condiciones iniciales adecuadas.



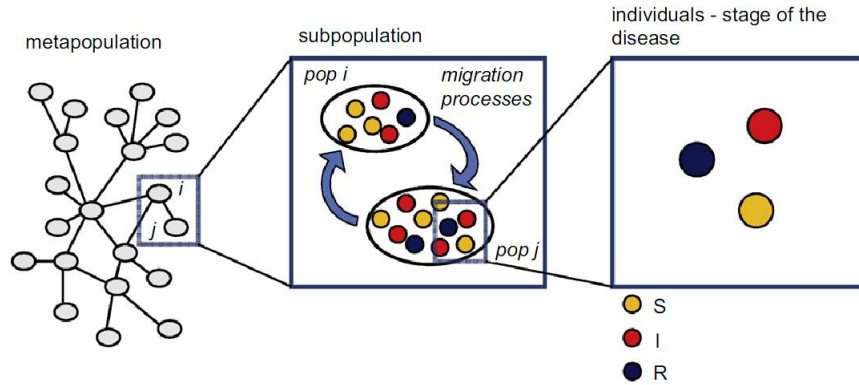
**Figura 10:** Esquema de un proceso RD aplicado en una red.

## 4.2 Global invasion threshold

En los modelos metapoblacionales, al igual que en los estudiados anteriormente, uno de los principales objetivos es hallar qué condiciones determinan que se pueda producir una epidemia. Es posible expresar estas condiciones mediante un único parámetro,  $R_*$ , que da cuenta del número medio de subpoblaciones a las que una subpoblación infectada puede transmitir la enfermedad mediante la movilidad de individuos infectados durante la duración de la epidemia. Este parámetro se conoce como *global invasion threshold*[31].

Así pues, si  $R_* < 1$  la epidemia permanece en la subpoblación inicial hasta que desaparece, mientras que si  $R_* > 1$  se puede propagar espacialmente en el sistema y alcanzar una dimensión global. Como vemos, este parámetro es una extensión del basic reproductive number de las secciones 3.2 y 3.3.

Las ecuaciones que acabamos de deducir no nos van a servir para calcular  $R_*$  ya que no tienen en cuenta ni las fluctuaciones inherentes al proceso de difusión ni la probabilidad de extinción del brote epidémico[30]. Por esta razón, vamos a tener que desarrollar otro modelo en el que se tenga en cuenta la aleatoriedad presente en el proceso.



**Figura 11:** Esquema del modelo SIR sobre una metapoblación[30].

Siguiendo con el modelo SIR, supongamos que la epidemia empieza en una subpoblación de grado  $k$  y  $N_k$  personas en la que  $R_0 > 1$ , de forma que localmente es posible que se produzca un brote. Denotemos por  $\alpha N_k$  al número total de individuos infectados que se generan en esta subpoblación. Cada persona infectada permanecerá de media en dicho estado durante un tiempo  $1/\mu$ , durante el cual podrá viajar a una subpoblación vecina de grado  $k'$  con una probabilidad  $d_{kk'}$ . Así, en primera aproximación, podemos considerar que el número de nuevos infectados que llegan a una subpoblación de grado  $k'$  provenientes de la subpoblación original es

$$\lambda_{kk'} = d_{kk'} \frac{\alpha N_k}{\mu}$$

Sea  $D_k^0$  el número de subpoblaciones “infectadas” de grado  $k$  al principio del proceso (generación 0). Cada una de ellas podrá comunicar la infección a sus vecinas, determinando el valor de  $D_k^1$ . Es decir, el número de subpoblaciones infectadas en la siguiente generación.

Al inicio del proceso el número de poblaciones en las que hay un brote será lo suficientemente pequeño como para poder considerar un proceso tipo árbol de forma que  $D_k^n$  podremos relacionarlo con  $D_k^{n-1}$ . Así,

$$D_k^n = \sum_{k'} D_{k'}^{n-1} (k' - 1) \lambda_{k'k} (R_0 - 1) P(k|k') \left( 1 - \frac{D_k^{n-1}}{V_k} \right) \quad (4.1)$$

Analicemos detenidamente esta expresión,

- Cada subpoblación de grado  $k'$  podrá infectar a  $k' - 1$  subpoblaciones. El factor  $-1$  da cuenta del hecho de que no se puede traspasar la epidemia a la subpoblación que le ha infectado.



- Como hay  $D_{k'}^{n-1}$  subpoblaciones de grado  $k'$  infectadas, el número total de subpoblaciones a las que puede pasar la epidemia desde una de grado  $k'$  será  $D_{k'}^{n-1}(k' - 1)$ .
- A continuación multiplicamos por la probabilidad de tengan un link con un vecino de grado  $k$ ,  $P(k|k')$
- Suponiendo que  $R_0 - 1 \ll 1$  podemos aproximar la probabilidad de que se produzca un brote en una subpoblación de grado  $k$  por  $\lambda_{kk'}(R_0 - 1)$
- Por último, debemos comprobar que la subpoblación de grado  $k$  no está infectada todavía,  $(1 - D_k^{n-1}/V_k)$

Sumando las contribuciones de todas las subpoblaciones obtenemos el número de subpoblaciones de grado  $k$  infectadas en la siguiente generación. Al estar al comienzo de la epidemia la probabilidad de que haya subpoblaciones infectadas es muy pequeña y podremos considerar  $(1 - D_k^{n-1}/V_k) \approx 1$ .

El siguiente paso es definir cómo se produce la difusión. Supongamos que en cada generación la fracción de individuos que abandona cada subpoblación es  $p$ . Lógicamente, la fracción de personas que viaja por cada link será  $p/k$  donde  $k$  es el número de links del nodo (difusión homogénea). Así,

$$\lambda_{k'k} = \frac{p}{k'} \frac{\alpha N_{k'}}{\mu}$$

Si la red no posee correlaciones entonces  $P(k|k') = kP(k)/\langle k \rangle[1]$ . Además, en difusión homogénea se cumple que  $N_{k'} = \bar{N}k'/\langle k \rangle$ , donde  $\bar{N}$  es el tamaño medio de las subpoblaciones. Introduciendo estas expresiones en (4.1),

$$D_k^n = (R_0 - 1) \frac{p\alpha\bar{N}kP(k)}{\mu\langle k \rangle^2} \sum_{k'} D_{k'}^{n-1} k'(k' - 1)$$

Definiendo  $\Theta^n \equiv \sum_{k'} D_{k'}^n k'(k' - 1)$  podemos reescribir esta expresión como

$$\Theta^n = (R_0 - 1) \frac{\langle k^2 \rangle - \langle k \rangle}{\langle k \rangle^2} \frac{p\bar{N}\alpha}{\mu} \Theta^{n-1}$$

Para que el número de subpoblaciones infectadas se incremente será necesario que

$$R_* \equiv (R_0 - 1) \frac{\langle k^2 \rangle - \langle k \rangle}{\langle k \rangle^2} \frac{p\bar{N}\alpha}{\mu} > 1 \quad (4.2)$$

Cualitativamente este resultado es muy similar al obtenido al estudiar la difusión en redes scale-free (3.12). Cuanto mayor sea la heterogeneidad de la red (mayor  $\langle k^2 \rangle$ ) más fácil será que se cumpla la condición  $R_* > 1$ . Es más, en el caso de una red scale-free ideal  $R_* \rightarrow \infty$ , el mismo resultado que en la difusión en una población.

Resulta interesante escribir la condición en función de  $p$ . Si  $R_0 - 1 \ll 1$  la constante  $\alpha$  se puede aproximar en el modelo SIR por[21]

$$\alpha \approx \frac{2(R_0 - 1)}{R_0^2}$$

De forma que

$$p\bar{N} \geq \frac{\langle k \rangle^2}{\langle k^2 \rangle - \langle k \rangle} \frac{\mu R_0^2}{2(R_0 - 1)^2} \quad (4.3)$$

La interpretación de este resultado es inmediata. Si existe una gran movilidad ( $p$  muy elevado) será fácil que se cumpla la condición y se producirá una epidemia a escala global. Análogamente, un valor

muy pequeño permitirá detener la difusión de la epidemia. Esto nos da una idea de cómo prevenir que se produzca una epidemia a escala mundial, si reducimos la movilidad de las personas (por ejemplo, cerrando los aeropuertos) deberíamos poder detener el proceso.

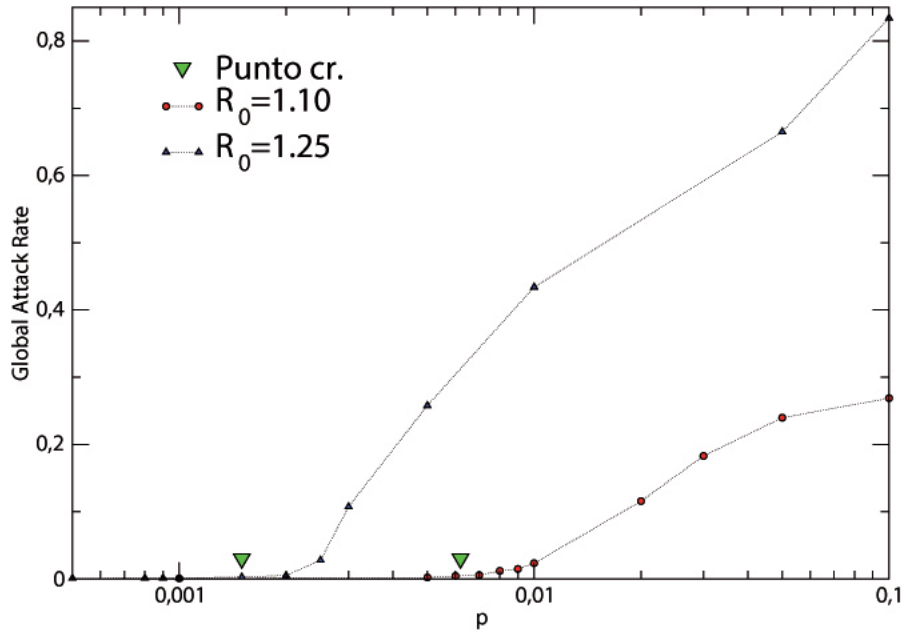
Sin embargo, en el caso de una red scale-free  $\langle k^2 \rangle$  será un número muy elevado, haciendo que el término de la derecha de la ecuación (4.3) tienda a cero. Por tanto, si queremos detener la epidemia restringiendo la movilidad deberemos reducirla prácticamente a 0, cosa bastante difícil.

Los datos de los que se dispone muestran que, efectivamente, las medidas de restricción del tráfico aéreo tienen un efecto muy limitado. Además, las simulaciones basadas en datos reales muestran que para que fuesen efectivas estas medidas la restricción debería ser prácticamente total[32][33].

Por último, veamos si estas predicciones se ajustan a los resultados de las simulaciones basadas en procesos de Monte Carlo.

Se define el *Global Attack Rate* (GAR) como la fracción de subpoblaciones en las que, una vez alcanzado el equilibrio, se ha producido un brote epidémico. Así, este parámetro nos permite medir el grado de propagación espacial de la epidemia. Representando el GAR frente a la movilidad  $p$  será posible comprobar la validez de la ecuación (4.3).

Para la simulación se ha empleado una red scale-free de 1000 nodos y exponente 2.1. El grado medio de la red es  $\langle k \rangle = 5.09$  y el segundo momento es  $\langle k^2 \rangle = 55.67$ . Además, en cada subpoblación hay, en promedio, 1000 personas y establecemos  $\mu = 0.2$ . Los resultados se muestran en la figura 12.



**Figura 12:** Fracción de subpoblaciones que han sufrido un outbreak (GAR) en función de la movilidad,  $p$ .

Como vemos el resultado es relativamente bueno. Hay que tener en cuenta que la expresión (4.3) está calculada suponiendo que  $R_0 - 1 \ll 1$ , por lo que no es de extrañar la diferencia observada. Una prueba de ello es que la concordancia es mejor en el caso de  $R_0 = 1.10$  que en el de  $R_0 = 1.25$ .

### 4.3 Aplicaciones: la pandemia de gripe A/H1N1

Para demostrar la validez y utilidad de los modelos metapoblacionales desarrollados hasta ahora vamos a analizar el estudio de Tizzoni et al.[34] sobre la pandemia de gripe A/H1N1 que se produjo en el 2009.

En este estudio emplean un modelo epidémico estocástico a escala global basado en metapoblaciones en el que integran datos demográficos y de movilidad reales. Este modelo se conoce como Global Epidemic

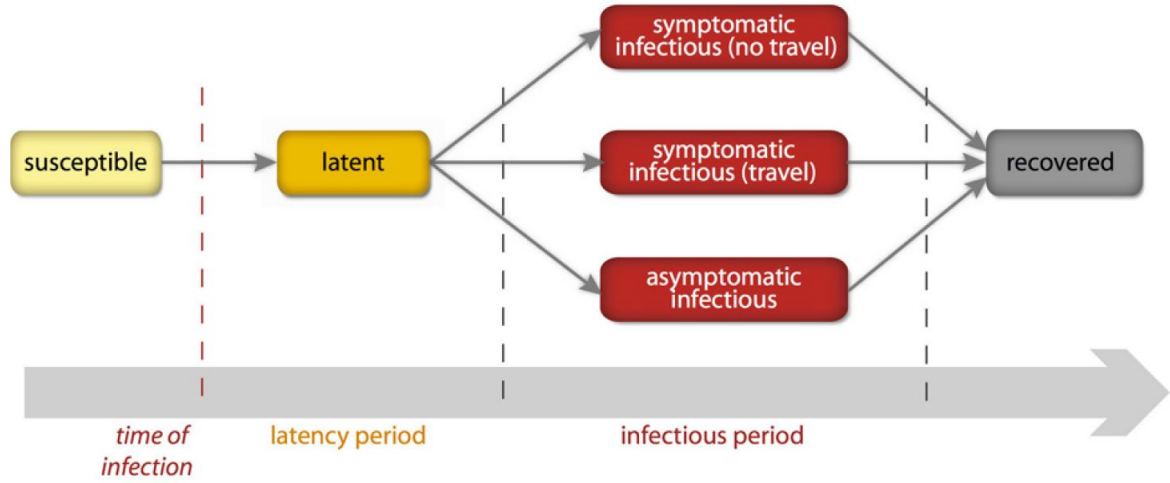
and Mobility (GLEAM) computational model[35].

El modelo GLEAM se puede dividir en tres “capas”: (i) la capa de poblaciones, (ii) la capa de transporte y (iii) la capa epidémica.

La capa de poblaciones se construye empleando una base de datos de gran resolución de la población mundial[36]. En ella se divide el planeta en una serie de celdas y se estima la población de cada una de ellas. Las subpoblaciones de la estructura metapoblacional se construyen alrededor de los grandes hubs de transporte, los aeropuertos. Las celdas se asocian al aeropuerto más cercano, obteniendo al final 3362 áreas repartidas entre 220 países.

De forma similar, la capa de movilidad se construye integrando los datos disponibles sobre el transporte aéreo a escala global con la movilidad entre subpoblaciones adyacentes. Así, el modelo es capaz de simular el número de pasajeros que viajan diariamente de una subpoblación a otra.

Por último, la capa epidémica hace uso de los modelos compartimentales adaptados a esta enfermedad. Cada individuo puede estar en el estado: susceptible, latente, sintomático infeccioso capaz de viajar, sintomático infeccioso incapaz de viajar, asintomático o recuperado permanente (figura 13). En cada subpoblación se considera homogeneous mixing.



**Figura 13:** Modelo compartimental de la gripe H1N1.

La transmisión de la enfermedad por los individuos sintomáticos viene dada por  $\beta$  y la de los asintomáticos se reescala mediante un factor  $r_\beta = 50\%$ . El periodo de incubación es  $\epsilon^{-1}$  tras el cual la probabilidad de pasar al estado asintomático es  $\mathcal{P}_a = 33\%$ . Para tener en cuenta a los individuos que dejarán de viajar por estar enfermos se aplica una penalización de un 50% a la probabilidad de desplazamiento de los individuos infectados. Finalmente, las personas infectadas que se recuperarán tras un tiempo  $\mu^{-1}$ .

El basic reproductive number que se obtiene con este modelo es

$$R_0 = \frac{\beta}{\mu} [r_\beta \mathcal{P}_a + (1 - \mathcal{P}_a)]$$

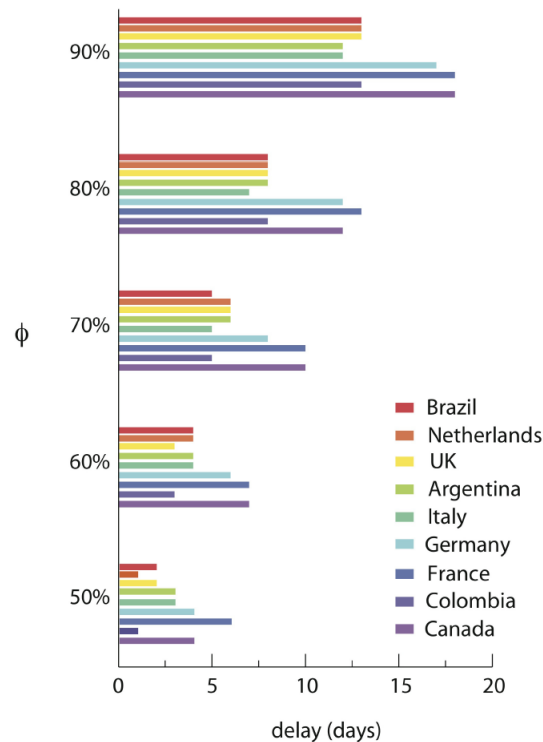
Para tener en cuenta los efectos estacionales se introduce en  $R_0$  un reescalado sinusoidal que depende de la fecha y de la región considerada. Así, se divide el mundo en tres regiones: hemisferio norte, hemisferio sur y región tropical. El parámetro, que se denota por  $\alpha(t)$ , varía entre  $\alpha_{\max}$  en invierno y  $\alpha_{\min}$  en verano,

$$\alpha(t) = \frac{1}{2} \left[ (\alpha_{\max} - \alpha_{\min}) \sin \left( \frac{2\pi}{365} (t - t_{\max}) + \frac{\pi}{2} \right) + \alpha_{\max} + \alpha_{\min} \right]$$

donde  $t_{\max}$  se define de forma que el mínimo sea el 15 de Julio y el máximo el 15 de Enero en el hemisferio norte y al contrario en el hemisferio sur. En la región tropical no se considera variación estacional.

Definido el modelo es posible estudiar el efecto de limitar el transporte aéreo, el punto en el que se debe producir el pico de la epidemia en cada región, el efecto de la vacunación... A continuación nos centraremos en los dos primeros aspectos.

La pandemia se originó en México y se extendió al resto del mundo por la red de transporte aéreo. Definiendo el tiempo de llegada a un país como la fecha en la que el primer caso sintomático fue detectado, es posible cuantificar el retraso en la aparición de la epidemia producido por la reducción del tráfico aéreo país por país. Así, la reducción de un 40% en el flujo de viajeros que se produjo en México se estima que retrasó de media la llegada de la epidemia a otros países unos 3 días.



**Figura 14:** Modelo compartimental de la gripe H1N1.

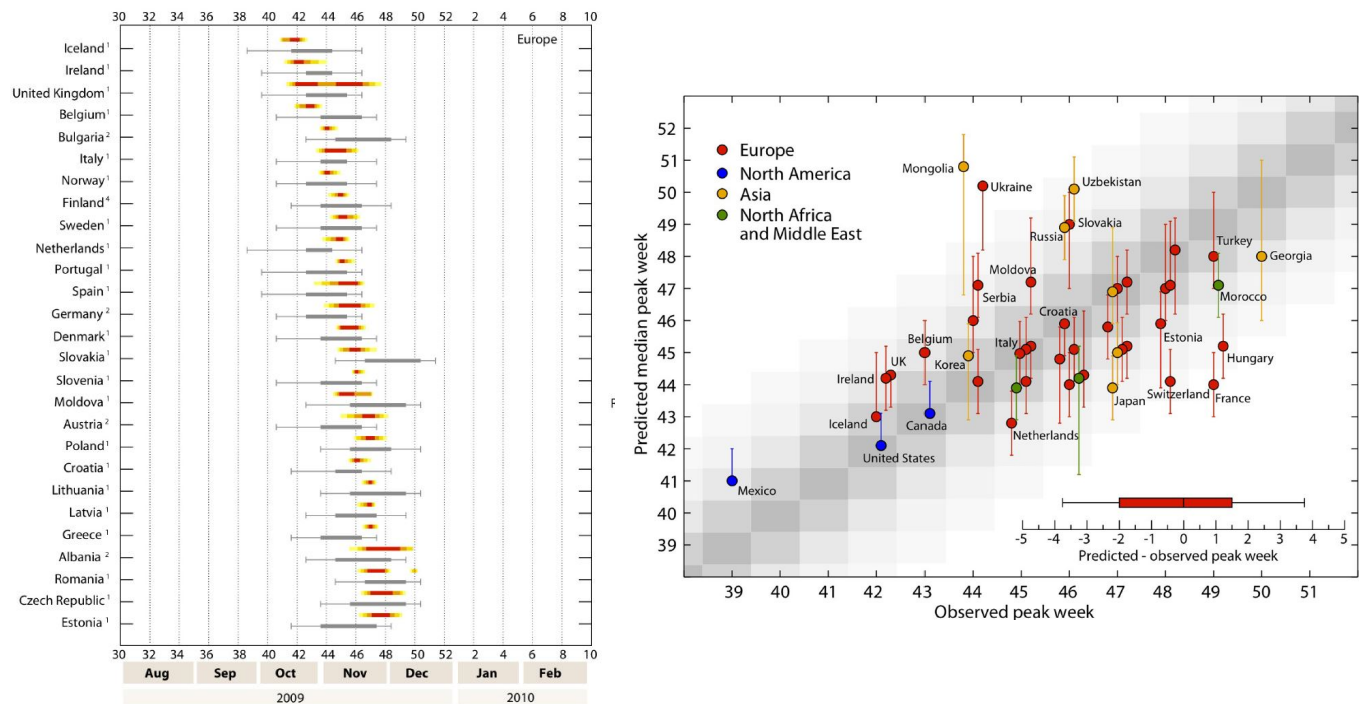
El modelo desarrollado permite determinar el retraso que producirían reducciones de tráfico extremas (figura 14). Vemos que incluso con una reducción de un 90% el retraso es menor de 20 días. Esto puede dar el tiempo suficiente a los gobiernos como para prepararse para la llegada de la epidemia, pero no será suficiente como para permitir desarrollar una vacuna. El tiempo estimado entre desarrollar, producir y distribuir una vacuna es de unos 6 meses.

Uno de los resultados más interesantes del artículo es el haber podido calcular con un error de menos de dos semanas la fecha concreta en la que se producirá el máximo de la epidemia en varios países.

Así, en la figura 15 se puede ver como el modelo predice correctamente que los picos se producen entre Octubre y Noviembre. Este resultado es muy importante para validarlo ya que según la estacionalidad que hemos impuesto el pico se debería producir a mediados de Enero. El que no lo haga significa que el resultado es consecuencia del movimiento de las personas y de la correcta estimación de los parámetros de la enfermedad.

Cabe destacar que la variación de 2 o 3 semanas de la predicción puede deberse no solo a las aproximaciones realizadas sino también a los factores que no se han tenido en cuenta debido a la falta de datos o al carácter muy localizado de los mismos.

El mejor ejemplo de ello es Francia. Los datos analizados muestran que la epidemia comenzó a propagarse por el país pero su crecimiento se detuvo durante unas dos semanas antes de continuar. Las fechas de este suceso coinciden precisamente con un periodo vacacional de los colegios. Así pues, el hecho de que



**Figura 15:** Izquierda: se representa en color el punto en el que se produjo el pico de la epidemia en varios países europeos, mientras que la predicción del modelo corresponde a las líneas grises. Derecha: representación del pico predicho por el modelo en función del observado.

el pico predicho por el modelo fuese dos semanas antes del que se observó puede deberse a las vacaciones de los colegios.

Como conclusión, vemos que a día de hoy es posible realizar predicciones muy buenas sobre procesos reales con los modelos que hemos desarrollado. Sin embargo, su precisión está condicionada al nivel de detalle con el que podamos conocer la movilidad y la distribución de la población.

## 5 Estudio de la persistencia en metapoblaciones

En la sección anterior hemos centrado nuestra atención en el modelo SIR debido a que apenas existen estudios sobre el modelo SIRS aplicado a metapoblaciones.

Motivados por esta situación, esta sección la vamos a dedicar a desarrollar un modelo computacional capaz de simular la difusión de una epidemia según el modelo SIRS aplicado sobre una metapoblación. Los resultados que obtengamos, aunque cualitativos, podrán servir como punto de partida para futuros desarrollos teóricos.

En el modelo SIR la duración de las epidemias es limitada ya que llega el momento en el que no quedan individuos susceptibles, por ello solo estudiábamos las condiciones necesarias para que se produjese un outbreak. Sin embargo, en el SIRS pueden mantenerse indefinidamente así que resultará más interesante estudiar las condiciones que permiten la permanencia/desaparición de las epidemias. En concreto, nos centraremos en la persistencia de una epidemia en función del tiempo medio que un individuo permanece inmune,  $\gamma^{-1}$ , y de la tasa de movilidad,  $p$ .

### 5.1 Modelo

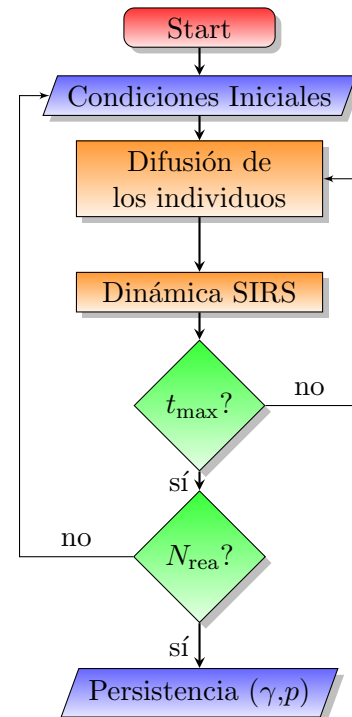
Con el objetivo de comprobar el efecto de la heterogeneidad de la red sobre el resultado realizaremos las simulaciones sobre varias redes diferentes. Construiremos tres redes aleatorias de 100, 1000 y 10000 nodos siguiendo el modelo ER[5] y tres redes scale-free con el mismo número de nodos mediante el modelo configuracional[37] y exponente  $\alpha = 3$ .

Consideraremos que la población total del sistema es de  $10^8$  de forma que la de cada subpoblación será, inicialmente,  $10^8/n$ , siendo  $n$  el número de nodos de la red. Como condición inicial introduciremos 50 individuos infectados en una subpoblación elegida al azar.

Supondremos difusión homogénea con tasa de movilidad  $p$ . Si el nodo  $j$  tiene grado  $k_j$ , en cada paso de tiempo, en promedio, viajarán  $N_j p / k_j$  individuos hacia cada uno de sus vecinos de forma que el número total de gente que se desplaza en cada nodo sea, en promedio,  $N_j p$ . En la simulación el número de individuos que van a cada nodo vendrá determinado por una distribución multinomial con probabilidades  $P_i = p / k_j$ .

Dentro de cada nodo la dinámica de la epidemia vendrá descrita por un modelo SIRS con homogeneous mixing. Los individuos pasarán de susceptible a infectado con una probabilidad  $P_{S \rightarrow I} = 1 - \left(1 - \frac{\beta}{N_j}\right)^{I_j}$ , de infectado a inmune con  $P_{I \rightarrow R} = \mu$  y de inmune a susceptible con  $P_{R \rightarrow S} = \gamma$ . El número concreto de susceptibles que pasan a infectados se calcula empleando una distribución binomial sobre una población de  $S_j$  individuos con probabilidad  $P_{S \rightarrow I}$ . Con el resto de clases se procede de forma análoga.

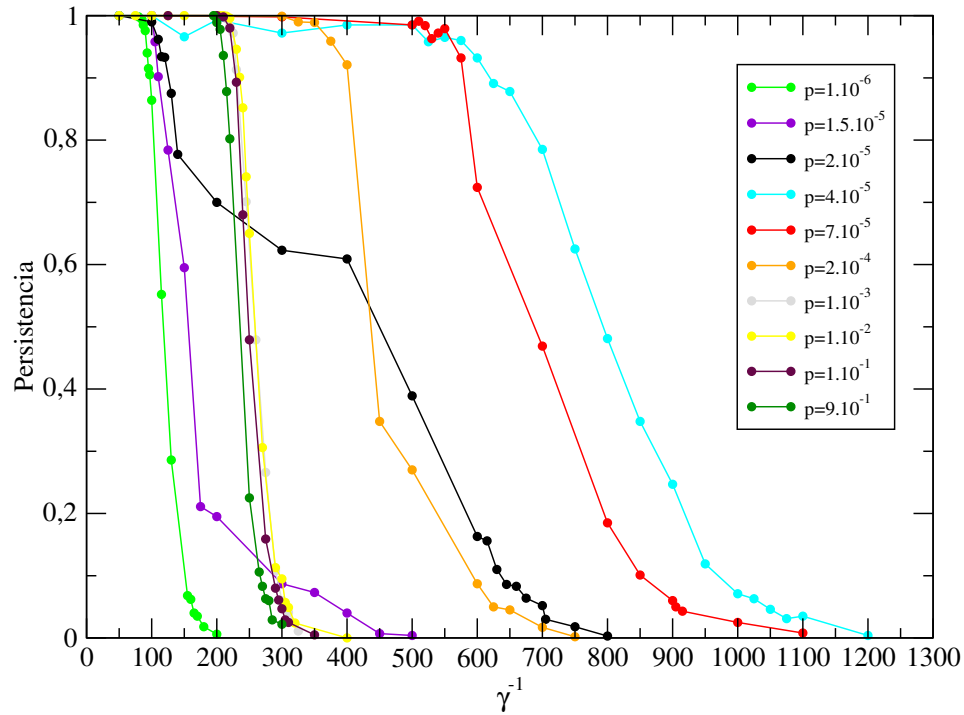
La simulación se detiene una vez que se ha alcanzado el equilibrio. Si queda por lo menos una persona infectada se considera que la persistencia es 1, si no queda nadie es 0. Este proceso se repite 1000 veces para cada valor de  $\gamma$  y  $p$  y se calcula el promedio con el objetivo de reducir el error estadístico. El tiempo de ejecución en las redes de 1000 nodos es de unas 20 horas para cada punto, por lo que en las redes de 10000 nodos el proceso solo se repite 100 veces (si mantuviésemos las 1000 realizaciones cada punto tardaría aproximadamente 8 días).



**Figura 16:** Diagrama de flujo de la simulación.

## 5.2 Resultados

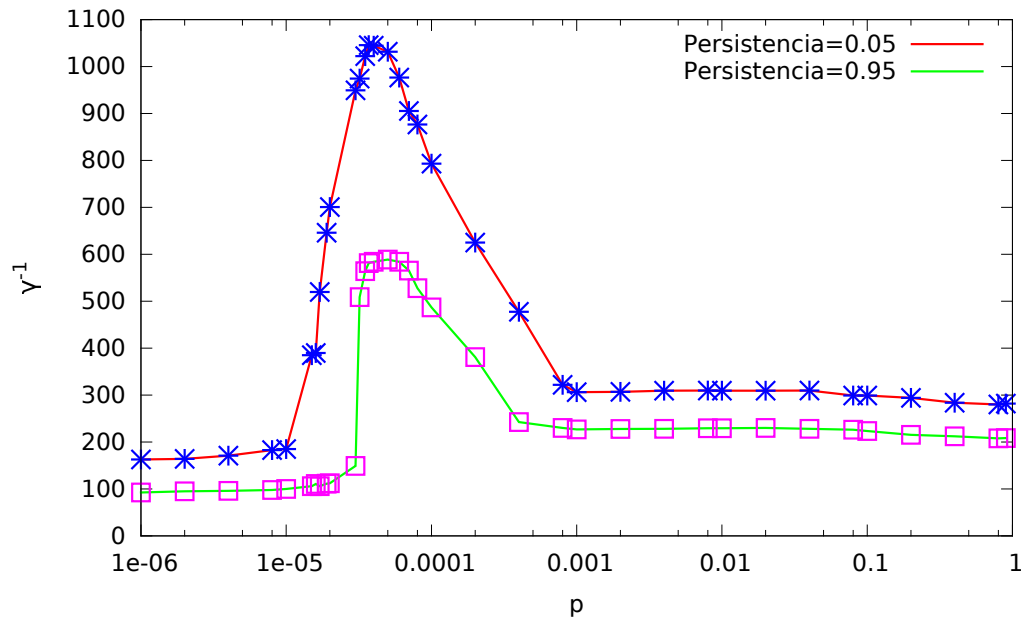
En primer lugar dejamos fijo el valor de  $p$  y variamos  $\gamma^{-1}$ . Realizando esto para varios valores de  $p$  se obtiene la figura 17.



**Figura 17:** Persistencia frente a duración de la inmunización en la red SF de 1000 nodos.

La gráfica muestra un comportamiento interesante. Al aumentar la movilidad se necesita un periodo de inmunización mucho mayor para erradicar la epidemia, como era de esperar. Sin embargo, se alcanza un máximo a partir del cual sucede lo contrario, a mayor movilidad más fácil es que se extinga.

Para analizar este comportamiento más cómodamente, vamos a buscar en qué puntos  $(p, \gamma^{-1})$  la probabilidad de persistencia es de un 5% y de un 95%. El resultado se muestra en la figura 18.



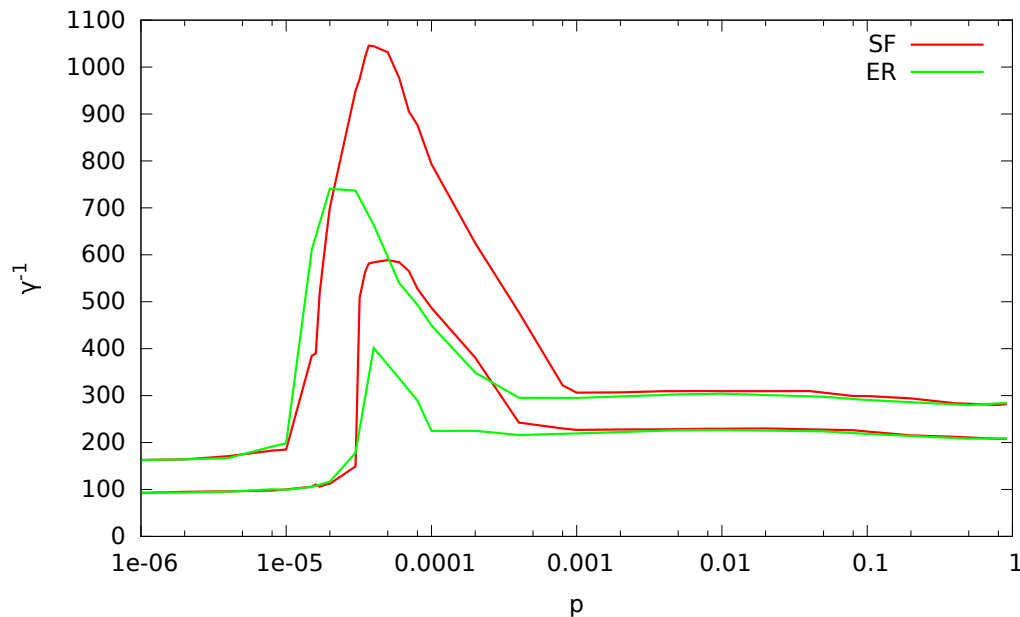
**Figura 18:** Persistencia frente a duración de la inmunización y movilidad en la red SF de 1000 nodos.

Así, podemos considerar que en cualquier punto  $(p, \gamma^{-1})$  que esté por encima de la línea roja la epidemia se ha extinguido. De la misma forma, en cualquier punto que esté por debajo de la verde la epidemia permanece siempre.

En el límite en el que  $p \rightarrow 0$  el sistema equivaldría a un montón de poblaciones desconectadas entre sí. Vemos como este comportamiento se mantiene así hasta que se alcanza un cierto valor crítico de la movilidad en el que empieza a haber una fracción apreciable de intercambio de individuos entre las poblaciones. Este intercambio hace que, aunque el periodo de inmunidad sea muy elevado, haya un suministro continuo de individuos susceptibles entre las subpoblaciones, lo que permite a la epidemia persistir.

Sin embargo, se alcanza un valor máximo a partir del cual la tendencia se invierte. Esto se puede achacar a que el intercambio es tan grande que en prácticamente todas las subpoblaciones habrá un outbreak. Así, la epidemia se propagará muy rápidamente y descenderá el número de susceptibles en todas las poblaciones, por lo que un valor pequeño de  $\gamma^{-1}$  será suficiente para eliminar la epidemia.

A valores muy elevados de la movilidad los efectos de la red se deberían ver disminuidos, recuperando en el límite  $p \rightarrow 1$  los resultados de una única población con  $10^8$  individuos.



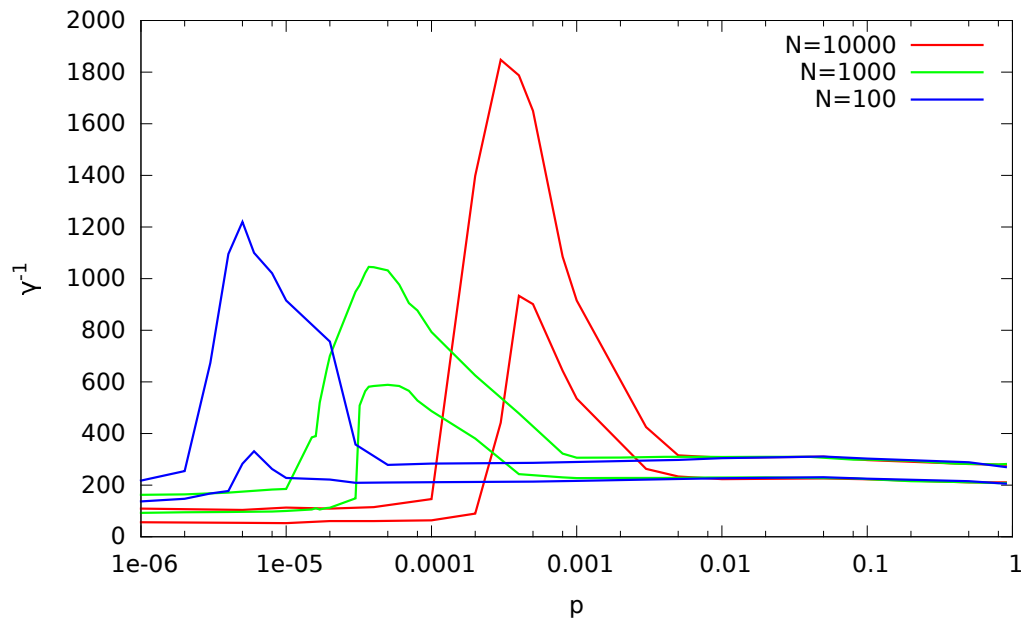
**Figura 19:** Persistencia frente a duración de la inmunización y movilidad en las redes SF y ER de 1000 nodos.

En efecto, tal y como se ve en la figura 19 a movilidades bajas y altas recuperamos el mismo resultado en la red SF y en la ER.

Además, la máxima inmunidad es mayor en la red SF que en la ER, por lo que tiene que estar relacionado con la estructura de la red. Será más fácil que la epidemia persista en un hub ya que al tener muchos vecinos tendrá un gran intercambio de individuos. Así, por una parte recibirá una gran cantidad de susceptibles y por otra propagará la epidemia a muchas subpoblaciones. Se recuperan resultados análogos al comparar las redes de 100 y de 10000.

Por último, comparamos los resultados para las redes SF, figura 20. La población total en todas ellas es de  $10^8$ , por lo que en las más grandes la población de cada nodo será menor. Así, cuando la epidemia está localizada en unas pocas subpoblaciones (movilidad baja) la persistencia es menor en las redes grandes. Al tener menor población se reducirá antes el número de susceptibles, de forma que será necesario un valor pequeño de  $\gamma^{-1}$  para que la epidemia pueda mantenerse.





**Figura 20:** Persistencia frente a duración de la inmunización y movilidad en las redes SF.

Para movilidades elevadas, como comentábamos anteriormente, se pierden los efectos de la red ya que se tiende a la situación de una única población con  $10^8$  individuos.

Estos resultados son muy interesantes ya que abren nuevos mecanismos para combatir una epidemia. Así, aunque a simple vista pudiera parecer contradictorio, podríamos aumentar las probabilidades de erradicar una epidemia incrementando el intercambio de personas sea cual sea la red. Esta medida, lógicamente, sería mucho más rentable en términos económicos y logísticos que restringir el tráfico[38].

Como comentábamos anteriormente estos resultados abren la puerta a nuevos estudios teóricos. En concreto, resultaría interesante determinar los tres puntos críticos de la movilidad en función de la estructura de la red.

## 6 Conclusiones

A lo largo del trabajo hemos podido comprobar la presencia de los comportamientos emergentes que mencionábamos en la introducción y su estrecha relación con la estructura de las redes complejas:

- En el apartado de introducción a las redes hemos visto que la mayoría de ellas siguen un comportamiento libre de escala.
- Al aplicar los modelos epidemiológicos a una red hemos podido ver el primer comportamiento colectivo consecuencia de la interacción entre los individuos: el umbral epidémico desaparece en las redes libres de escala.
- En los modelos metapoblacionales vuelve a aparecer un ejemplo de comportamiento colectivo: el global invasion threshold depende explícitamente de la red y de la intensidad de la interacción (la movilidad) entre las subpoblaciones.
- En el estudio de la persistencia hemos comprobado la gran dependencia que existe con la movilidad y la estructura de la red.

Podemos concluir, por tanto, que la difusión de una epidemia es un ejemplo claro de lo que hemos denominado sistemas complejos. Así, si queremos ser capaces de predecir e incluso detener la propagación de una epidemia no nos bastará con estudiar las características de la enfermedad, sino que deberemos conocer con precisión los patrones de movilidad y las estructuras sociales de la población, tal y como se ha podido comprobar en el estudio de la persistencia en metapoblaciones.

## Bibliografía

- [1] Mark Newman. *Networks: an introduction*. Oxford University Press, 2010.
- [2] S Gómez, A Arenas, J Borge-Holthoefer, S Meloni, and Y Moreno. Discrete-time Markov chain approach to contact-based disease spreading in complex networks. *EPL (Europhysics Letters)*, 89(3):38009, 2010.
- [3] Béla Bollobás. *Modern graph theory*, volume 184. Springer, 1998.
- [4] Supriyo Datta. *Electronic transport in mesoscopic systems*. Cambridge University press, 1997.
- [5] Erdős P. and Rényi A. On random graphs i. *Publ. Math. Debrecen*, 6:290–297, 1959.
- [6] Statista. Average number of facebook friends of u.s. users in 2014, by age group. <http://www.statista.com/statistics/232499/>.
- [7] Albert-László Barabási and Jennifer Frangos. *Linked: The New Science Of Networks Science Of Networks*. Basic Books, 2002.
- [8] Réka Albert, Hawoong Jeong, and Albert-László Barabási. Internet: Diameter of the world-wide web. *Nature*, 401(6749):130–131, 1999.
- [9] Albert-László Barabási and Réka Albert. Emergence of scaling in random networks. *Science*, 286(5439):509–512, 1999.
- [10] John MD Coey. *Magnetism and magnetic materials*. Cambridge University Press, 2010.
- [11] Size of the world wide web. <http://www.worldwidewebsite.com/>.
- [12] Albert-László Barabási and Eric Bonabeau. Scale-free networks. *Scientific American*, 2003.
- [13] Guido Caldarelli. Scale-free networks: complex webs in nature and technology. *OUN Catalogue*, 2007.
- [14] Mapping the internet. <http://nicolasrapp.com/?p=1180>.
- [15] Manfred S Green, Tiberio Swartz, Elana Mayshar, Boaz Lev, Alex Leventhal, Paul E Slater, and Joshua Shemer. When is an epidemic an epidemic? *The Israel Medical Association Journal: IMAJ*, 4(1):3–6, 2002.
- [16] David Easley and Jon Kleinberg. Networks, crowds, and markets. *Cambridge Univ Press*, 6(1):6–1, 2010.
- [17] Alain Barrat, Marc Barthélemy, and Alessandro Vespignani. *Dynamical processes on complex networks*, volume 574. Cambridge University Press Cambridge, 2008.
- [18] The Pennsylvania State University. Epidemic - the dynamics of infectious diseases. Coursera, 2013.
- [19] Albert-László Barabási. Class 17: Epidemic modeling. University Lecture, 2012.
- [20] Stefano Boccaletti, Vito Latora, Yamir Moreno, Mario Chavez, and D-U Hwang. Complex networks: Structure and dynamics. *Physics reports*, 424(4):175–308, 2006.
- [21] James D. Murray. *Mathematical Biology: I. An Introduction (Interdisciplinary Applied Mathematics) (Pt. 1)*. Springer, 2007.
- [22] James Holland Jones. Notes on  $\mathcal{R}_0$ . Stanford University, May 2007.
- [23] Beniamino Guerra and Jesús Gómez-Gardeñes. Annealed and mean-field formulations of disease dynamics on static and adaptive networks. *Physical Review E*, 82(3):035101, 2010.

- [24] Yamir Moreno, Romualdo Pastor-Satorras, and Alessandro Vespignani. Epidemic outbreaks in complex heterogeneous networks. *The European Physical Journal B-Condensed Matter and Complex Systems*, 26(4):521–529, 2002.
- [25] Romualdo Pastor-Satorras and Alessandro Vespignani. Epidemic spreading in scale-free networks. *Physical review letters*, 86(14):3200, 2001.
- [26] Neil M Ferguson, Matt J Keeling, W John Edmunds, Raymond Gani, Bryan T Grenfell, Roy M Anderson, and Steve Leach. Planning for smallpox outbreaks. *Nature*, 425(6959):681–685, 2003.
- [27] Andrea Apolloni, Chiara Poletto, José J Ramasco, Pablo Jensen, Vittoria Colizza, et al. Metapopulation epidemic models with heterogeneous mixing and travel behaviour. *Theoretical Biology and Medical Modelling*, 11(1):3, 2014.
- [28] Eckehard Schöll. *Nonlinear spatio-temporal dynamics and chaos in semiconductors*, volume 10. Cambridge University Press, 2001.
- [29] Vittoria Colizza, Romualdo Pastor-Satorras, and Alessandro Vespignani. Reaction–diffusion processes and metapopulation models in heterogeneous networks. *Nature Physics*, 3(4):276–282, 2007.
- [30] Vittoria Colizza and Alessandro Vespignani. Epidemic modeling in metapopulation systems with heterogeneous coupling pattern: Theory and simulations. *Journal of theoretical biology*, 251(3):450–467, 2008.
- [31] Sandro Meloni, Nicola Perra, Alex Arenas, Sergio Gómez, Yamir Moreno, and Alessandro Vespignani. Modeling human mobility responses to the large-scale spreading of infectious diseases. *Scientific reports*, 1, 2011.
- [32] T Déirdre Hollingsworth, Neil M Ferguson, and Roy M Anderson. Will travel restrictions control the international spread of pandemic influenza? *Nature medicine*, 12(5):497–499, 2006.
- [33] Ben S Cooper, Richard J Pitman, W John Edmunds, and Nigel J Gay. Delaying the international spread of pandemic influenza. *PLoS Medicine*, 3(6):e212, 2006.
- [34] Michele Tizzoni, Paolo Bajardi, Chiara Poletto, José J Ramasco, Duygu Balcan, Bruno Gonçalves, Nicola Perra, Vittoria Colizza, and Alessandro Vespignani. Real-time numerical forecast of global epidemic spreading: case study of 2009 a/h1n1pdm. *BMC medicine*, 10(1):165, 2012.
- [35] Global epidemic and mobility (GLEAM) computational model. <http://www.mobs-lab.org/global-epidemic-and-mobility-model.html>.
- [36] SocioEconomic Data and Application Center at Columbia University. Gridded population of the world (GPW). <http://sedac.ciesin.columbia.edu/data/collection/gpw-v3>.
- [37] Michele Catanzaro, Marián Boguñá, and Romualdo Pastor-Satorras. Generation of uncorrelated random scale-free networks. *Physical Review E*, 71(2):027103, 2005.
- [38] Alberto Aleta, Andreia Hisi, Chiara Poletto, Sandro Meloni, Vittoria Colizza, and Yamir Moreno. in preparation.

