



Universidad Zaragoza

Nuevos modelos de inmunización de poblaciones estructuradas

Juan Castillo Sanz

Trabajo fin de grado
Facultad de Ciencias
Grado en Física

Tutores

Tutor: Yamir Moreno Vega
Tutor: Alejandro Tejedor Cubero

2024

Índice

1. Introducción.	1
2. Marco teórico: teoría de redes.	2
2.1. Nociones básicas de la teoría de redes.	2
2.2. Medidas de centralidad.	3
2.3. Modelos de redes.	4
3. Marco teórico: modelos compartimentales.	7
3.1. Modelo SIR. Campo medio.	7
3.2. Modelo SIR. Dependencia con la estructura de la población	10
4. Modelo y metodología	13
4.1. Modelo de población.	13
4.2. Modelo epidemiológico	14
4.3. Metodología de simulación	14
5. Resultados y discusión.	15
5.1. Vacunación con prevalencia baja	17
5.2. Vacunación con prevalencia alta	18
6. Conclusiones.	20

1. Introducción.

Tal y como menciona P.W. Anderson en el artículo *More is different* [1], no siempre resulta posible reconstruir el comportamiento de sistemas de varios elementos extrapolando las propiedades individuales. En palabras de Marx, "*diferencias meramente cuantitativas, más allá de cierto punto, se convierten en cambios cualitativos*" [2]. Debido a esto, en las últimas décadas una cantidad significativa de investigadores han abandonado la búsqueda de los modelos reduccionistas tradicionales con el objetivo de explicar los fenómenos que surgen debido a la interacción entre componentes. Es así como surge una nueva ciencia interdisciplinar: la ciencia de los sistemas complejos.

Una de las áreas que más se ha beneficiado de esta nueva ciencia es la epidemiología. A pesar de que tradicionalmente se haya asociado esta con el ámbito de las ciencias biomédicas, existe también una larga relación con la física. Gracias a esta, se han podido desarrollar nuevos modelos cuantitativos e incorporar en el análisis nuevas herramientas como puede ser la ley de acción de masas usada en los modelos compartimentales. Estas técnicas y modelos nos permiten obtener resultados más potentes. Sin embargo, como la reciente crisis del COVID-19 ha mostrado, es necesario continuar con la investigación, no solo con el objetivo de entender mejor la propagación de patógenos, sino también con el fin de desarrollar métodos que aprovechen eficientemente los recursos limitados.

La vacunación de los individuos de la población es una de las formas actualmente más usadas en la prevención de la propagación de enfermedades. Gracias a estas, hemos conseguido reducir enfermedades como la difteria o el sarampión e incluso erradicar la polio. No obstante, no siempre es posible que toda la población pueda ser vacunada y puede ser necesario establecer un orden de importancia de los individuos con el objetivo de minimizar el impacto en el conjunto de la sociedad.

Aunque existen multitud de métricas que miden esta importancia, las desarrolladas desde la ciencia de los sistemas complejos tienen como objetivo cortar las vías de transmisión más relevantes. Para ello, hacen uso de las propiedades topológicas de la red de contactos subyacente a la población por la que puede transmitirse la enfermedad. Es en este último campo en el que nos centraremos en este trabajo, por lo que nos será necesario introducir algunos conceptos de teoría de redes.

Concretamente, estudiaremos la eficacia de una de las medidas más recientes: DomiRank [3], la cual ha demostrado ya ser eficiente destruyendo distintos tipos de redes y frenando otros procesos de difusión. Con el objetivo de comprobar la utilidad de esta nueva métrica en el proceso de vacunación, tras introducir algunos modelos y conceptos de teoría de redes (sección 2) y epidemiología (sección 3), realizaremos simulaciones de diferentes estrategias de vacunación en redes con diversas topologías (sección 4 y 5).

2. Marco teórico: teoría de redes.

La teoría de redes constituye el marco teórico a través del cuál se modelizan muchos sistemas complejos de interés. Mediante esta, es posible representar el sistema bajo estudio centrándonos exclusivamente en la estructura de las conexiones entre componentes. Gracias a esto, resulta más sencillo comprender el efecto que tiene esta estructura en los diferentes procesos que se desarrollan sobre ella.

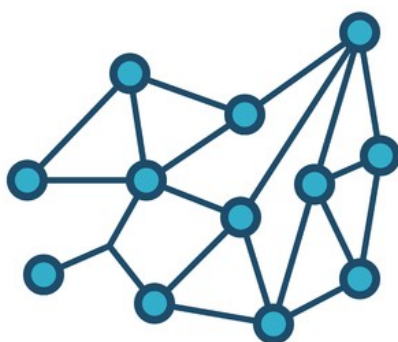


Figura 1: Ejemplo de un grafo pequeño. Mediante un grafo de este tipo podríamos representar desde una red de ordenadores en una oficina hasta la red ferroviaria de una región pequeña.

2.1. Nociones básicas de la teoría de redes.

Los elementos básicos de una red o grafo son los nodos y los links. Mediante los primeros representaremos los constituyentes básicos de nuestro sistema; mientras que los segundos simbolizarán las conexiones entre estos. En el caso que nos concierne, por ejemplo, los nodos van a representar a los individuos de la población y la presencia de links indicará que existe contacto y, por tanto, la posibilidad de contagio directo entre dos de ellos.

Diremos que un grafo es simple cuando entre dos nodos cualesquiera existe como máximo un link y no existen links que conecten un nodo consigo mismo. Puesto que las interacciones pueden tener direccionalidad, debemos distinguir entre grafos dirigidos (aquellos en los que los links tienen orientación) y no dirigidos (aquellos en los que los links no tienen orientación). Distinguiremos además entre redes con links pesados (en las cuales cada link tiene asociado un peso que indica su intensidad o probabilidad de interacción) y no pesados. Por las características del sistema que estamos analizando, en este trabajo únicamente tendremos en cuenta redes simples no pesadas y no dirigidas.

Existen diversas formas de almacenar la información relevante de una red. Una de ellas es a través de la matriz de adyacencia A . Para grafos no pesados y no dirigidos, el elemento de matriz a_{ij} valdrá 1 si existe un link entre el nodo i y el nodo j y 0 si no existe. A pesar de que computacionalmente existen formas más eficientes de almacenar un grafo, algunas herramientas y derivaciones matemáticas importantes, así como aplicaciones y resultados prácticos, tienen la matriz de adyacencia en su origen [4], [5].

2.2. Medidas de centralidad.

En la mayoría de ocasiones, las redes presentan heterogeneidades; tales como que los nodos no tengan el mismo número de conexiones. La existencia de estas provoca que no todo los nodos tengan la misma relevancia dentro de la estructura de la red. Las medidas de centralidad son magnitudes que nos determinan dicha importancia de acuerdo a diferentes criterios. El uso de las medidas de centralidad juega un papel muy relevante en el estudio de la dinámica, ya que permiten identificar aquellos nodos particularmente relevantes para el desarrollo del proceso sobre la red.

En la literatura existen multitud de medidas de centralidad. Sin embargo, cada una de estas medidas tiene en cuenta aspectos topológicos diferentes y, por tanto, no serán igual de relevantes en todos los procesos [6]. Así, la elección de la medida de centralidad relevante es fundamental en el diseño de estrategias de mitigación. En este trabajo nos centraremos en dos medidas diferentes: la centralidad de grado y DomiRank.

- **Centralidad de grado.** La centralidad de grado es, probablemente, la medida de centralidad más básica. En esta, a cada nodo i se le asigna como valor de centralidad el número de conexiones, también conocido como grado k_i . A pesar de ser una de las medidas más simples, se ha comprobado que está relacionada con la propagación de epidemias [4].
- **DomiRank.** DomiRank cuantifica la dominancia del nodo con respecto a sus vecinos [3]. Nodos con alto valor según DomiRank tienden a estar conectados a otros con bajo valor. Tal y como se demuestra en el artículo en el que se propone esta medida, DomiRank ofrece una forma eficiente de destruir la red y juega un papel relevante frenando otros procesos de difusión.

Matemáticamente, en redes no pesadas, DomiRank se define como la solución estacionaria de la ecuación (1), donde Γ_i se corresponde con el valor DomiRank del nodo i -ésimo y α, β, θ son números reales positivos.

$$\frac{d\Gamma_i(t)}{dt} = \alpha \left(\theta k_i - \sum_j a_{ij} \Gamma_j(t) \right) - \beta \Gamma_i(t) \quad (1)$$

La solución de la ecuación anterior a un tiempo infinito nos la da la ecuación 2, donde la entrada i del vector $\mathbf{\Gamma}$ es el valor DomiRank correspondiente al nodo i . Como se puede apreciar, el parámetro θ únicamente define la escala de $\mathbf{\Gamma}$, por lo que podemos fijarlo a 1 sin perder generalidad. El criterio de convergencia está condicionado al valor del cociente $\sigma = \alpha/\beta$. Concretamente, para que DomiRank esté bien definido, $\sigma \in (0, -\frac{1}{\lambda_N})$, donde λ_N es el autovalor mínimo de la matriz de adyacencia.

$$\mathbf{\Gamma} = \theta\sigma (\sigma A + \mathbb{I}_{N \times N})^{-1} A \mathcal{I}_{N \times 1} \quad (2)$$

Donde A es la matriz de adyacencia, $\mathbb{I}_{N \times N}$ es la matriz identidad de tamaño $N \times N$ y $\mathcal{I}_{N \times 1}$ es un vector columna cuyas entradas valen todas 1.

Sin embargo, σ no solo juega un papel fundamental en la convergencia de (1), sino que también controla el nivel de competencia entre nodos que nos establece su dominancia, definiendo una escala de la distancia a la que el nodo puede influir. Si $\sigma \rightarrow 0$, la competencia entre nodos desaparece y DomiRank solo tiene en cuenta las propiedades locales; es decir, el grado del nodo. En cambio, si $\sigma \rightarrow -\frac{1}{\lambda_N}$, la competición entre nodos se vuelve muy importante y DomiRank pasa a tener en cuenta la estructura global de la red (vease la figura 2).

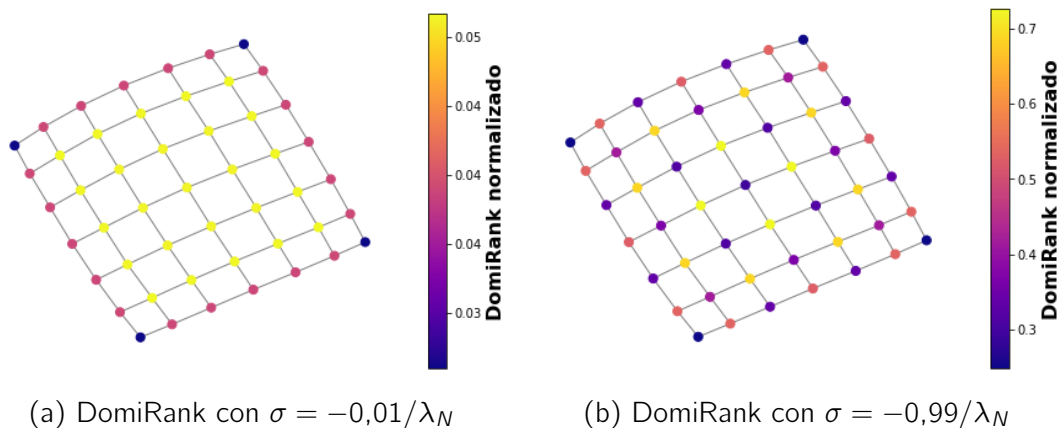


Figura 2: Cálculo de DomiRank sobre una red cuadrada regular en 2D para distintos valores de σ . Para σ bajo (imagen izquierda), el valor de DomiRank únicamente distingue por grado del nodo. En cambio, para σ alto, DomiRank es capaz de observar las propiedades globales asignando importancia a cada nodo según la distancia al borde de la red.

2.3. Modelos de redes.

No todas las redes tienen las mismas características. Con el objetivo de reproducir las diferentes propiedades observadas en los sistemas reales, varios modelos de redes complejas han sido propuestos. En este trabajo, nos centraremos en tres de los modelos más usados.

- Redes aleatorias. Modelo Erdős-Rényi [7].** En este tipo de redes partimos de un conjunto de N nodos, de tal forma que tenemos $\frac{N(N-1)}{2}$ posibles links. Para crear el grafo, activamos cada uno de estos posibles enlaces con probabilidad p . Este tipo de redes será adecuado para redes cuya formación este muy influenciada por procesos aleatorios.

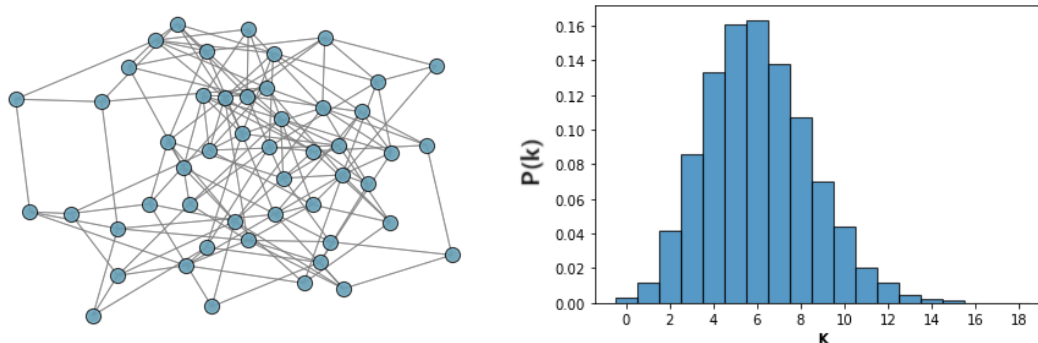


Figura 3: Ejemplo de un grafo aleatorio. La figura de la izquierda muestra una representación de este tipo de grafos. La imagen de la derecha muestra la distribución de probabilidad de k en este tipo de redes.

Es sencillo ver entonces que la distribución de probabilidad del número de enlaces se corresponde con una distribución binomial, la cual podrá aproximarse a una distribución de Poisson cuando el valor de la probabilidad p sea lo suficientemente bajo. Un calculo simple nos determina que la conectividad media de cada uno de los nodos será entonces $\langle k \rangle = p(N - 1) \approx pN$.

- Redes libres de escala. Modelo Barabasi-Albert [8].** Aunque el modelo Erdős-Rényi resulta interesante por su simplicidad matemática, la conectividad en las redes sociales no está gobernada por procesos aleatorios. En este tipo de redes, la distribución de grado sigue una ley potencial tal que $p(k) \propto k^{-\gamma}$, donde habitualmente $2 \leq \gamma \leq 3$ [9]. Una característica importante en este tipo de distribuciones es que $\langle k^2 \rangle$ diverge.

En este tipo de redes, la mayoría de los nodos tienen pocas conexiones mientras que unos pocos nodos están muy conectados con el resto. A los nodos con muchas conexiones se les denomina hubs. Como veremos en la sección 3, la existencia de estos hubs juega un papel fundamental en las dinámicas de contagio.

Para construir esta red, debemos empezar con una red aleatoria en la que todos los nodos tengan al menos una conexión. Tras esto, vamos añadiendo de uno a uno nuevos nodos que formarán m conexiones con la red. Para que aparezcan los hubs, la probabilidad de formar un enlace con el nodo i será proporcional a su grado. Cuando el número de nodos sea lo suficientemente grande, este proceso generará redes cuya distribución de grado escala como k^{-3} y que tienen $\langle k \rangle = 2m$.

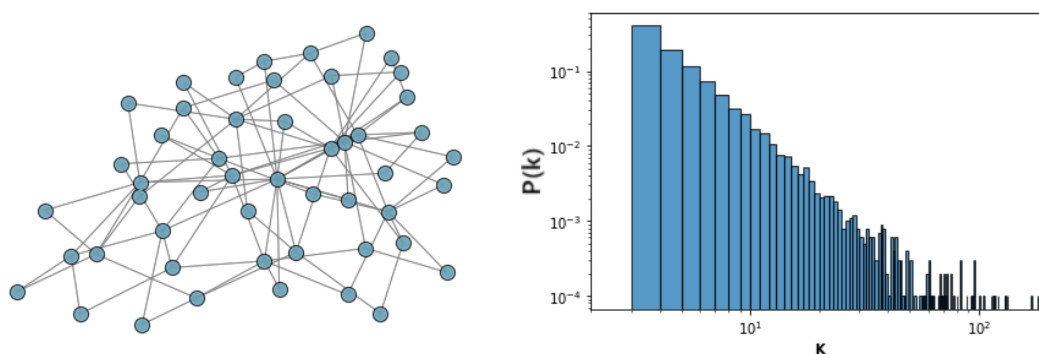


Figura 4: Ejemplo de un grafo libre de escala. La figura de la izquierda muestra una representación de este tipo de grafos. La imagen de la derecha muestra la distribución de probabilidad de k en este tipo de redes.

- Random Geometric Graph** (Grafo geométrico aleatorio)[10]. En numerosas ocasiones, las interacciones entre dos individuos diferentes de una población se ven influenciadas por la distancia entre ellos. Debido a esto, resulta interesante introducir modelos de redes que tengan en cuenta aspectos geométricos como el Random Geometric Graph. Para construir un grafo de este tipo, debemos distribuir aleatoriamente N puntos en un hipercubo de lado 1 en un espacio d -dimensional (nosotros nos limitaremos a 2 dimensiones). Tras esto, conectamos cada uno de esos puntos con el resto de puntos que se encuentren dentro de una d -esfera de radio r .

Como consecuencia de incorporar efectos espaciales, los RGG darán lugar a agrupaciones de nodos altamente conectados entre sí en las zonas con mayor densidad de nodos. A este tipo de agrupaciones se les conoce como estructuras de comunidad, las cuales no aparecen en los modelos Barabasi-Albert o Erdős-Rényi. Debemos tener en cuenta no obstante que la existencia de estas estructuras de comunidad introduce correlaciones que complicarán futuros desarrollos matemáticos.

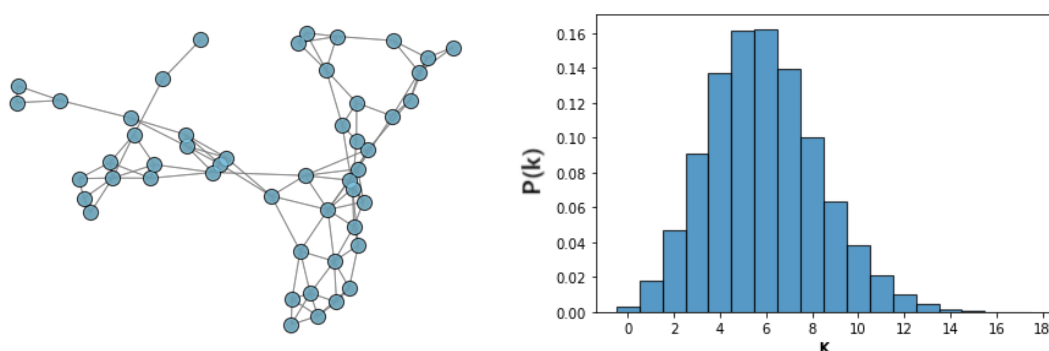


Figura 5: Ejemplo de un random geometric graph. La figura de la izquierda muestra una representación de este tipo de grafos. La imagen de la derecha muestra la distribución de probabilidad de k en este tipo de redes.

Puede demostrarse que la distribución de probabilidad para valores de N y r intermedios se asemeja a una distribución de Poisson. En dos dimensiones, de forma naive podemos estimar el grado medio de un nodo como el número de nodos por la fracción del área cubierta por un círculo de radio r , es decir $\langle k \rangle = N\pi r^2$.

3. Marco teórico: modelos compartimentales.

En la mayoría de procesos infecciosos, el individuo pasa por varios estados diferentes y las transiciones entre estos estados están provocadas por procesos claramente definidos. Debido a esto, históricamente se han introducido clases en las que podemos agrupar a las diferentes personas de la población de tal forma que las variables que caracterizan nuestro sistema sean el porcentaje de la población que se encuentra en cada una de las categorías. A este tipo de modelos se le denomina modelos compartimentales.

3.1. Modelo SIR. Campo medio.

El modelo SIR constituye uno de los modelos epidemiológicos más usados. Fue propuesto en 1927 por W. K. Kermack y A. G. McKendrick, quienes dividieron a la población en tres grupos: susceptibles (S), infectados (I) y recuperados (R) [11]. El paso del estado S al estado I se da con una tasa λ cuando haya un contacto con una persona infectada. Por otra parte, el paso de infectado a recuperado se da con una tasa μ sin necesidad de cumplir ninguna condición.

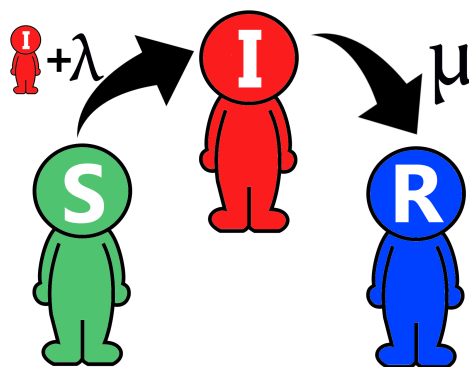


Figura 6: Cadena de Markov del modelo SIR. La población queda dividida en tres clases. Se puede pasar de la clase S a la I con una tasa λ cuando hay un contacto con un individuo de la clase I y de la clase I se puede pasar a la clase R con una tasa μ .

El modelo tiene tres importantes hipótesis:

1. Durante el tiempo que dure la dinámica, el tamaño de la población se mantiene constante. Es decir, no hay nacimientos, muertes no debidas a la enfermedad o migraciones de población.
2. El tamaño de la población es lo suficientemente grande como para poder considerar el límite termodinámico. Esto nos permite despreciar fluctuaciones estadísticas y considerar únicamente ecuaciones deterministas para la evolución de los promedios.
3. Los individuos de la población tienen el mismo comportamiento social; es decir, cualquier individuo puede relacionarse con el resto de la población y todos van a tener el mismo número k de contactos por unidad de tiempo. A esta hipótesis se le conoce como hipótesis de población homogénea o bien mezclada.

Bajo estas hipótesis, las ecuaciones que describen la evolución de las fracciones de población en cada uno de los estados son las ecuaciones (3)-(5), donde las variables S , I , R son dichas proporciones. Como ya se ha mencionado en la introducción, estas ecuaciones están inspiradas en la ley de acción de masas. Puesto que los individuos de la población están en una única de las tres clases, se debe cumplir que $S + I + R = 1$ y, por lo tanto, la ecuación (5) es en realidad redundante.

$$\frac{dS}{dt} = -\lambda k S I \quad (3)$$

$$\frac{dI}{dt} = \lambda k S I - \mu I \quad (4)$$

$$\frac{dR}{dt} = \mu I \quad (5)$$

La única forma de lograr una solución estacionaria del sistema de EDOs anterior es con $I_\infty = 0$ y se logra independientemente del valor de S_∞ y R_∞ . Así, tras pasar un tiempo largo, no quedarán individuos infectados en la población. No obstante, esto no nos indica si en el que el sistema ha tardado en alcanzar este punto de equilibrio, el número de infectados ha aumentado significativamente con respecto al valor inicial.

Para estudiar esto, se toma la ecuación (4), y se reescribe de la siguiente forma:

$$\frac{dI}{dt} = \mu (\mathcal{R}_0 S - 1) I \quad (6)$$

donde $\mathcal{R}_0 = \frac{\lambda k}{\mu}$ es el número reproductivo básico. Notese que \mathcal{R}_0 incorpora tanto información de la enfermedad como del comportamiento de la población. Otro forma también habitual de redefinir dicha ecuación es mediante $\mathcal{R}_E(t) = \mathcal{R}_0 S(t)$. A la vista de la ecuación (6), para que la fracción de individuos infectados aumente, $\mathcal{R}_0 S - 1 \geq 0$ o, equivalentemente $\mathcal{R}_E - 1 \geq 0$.

3.1 Modelo SIR. Campo medio.

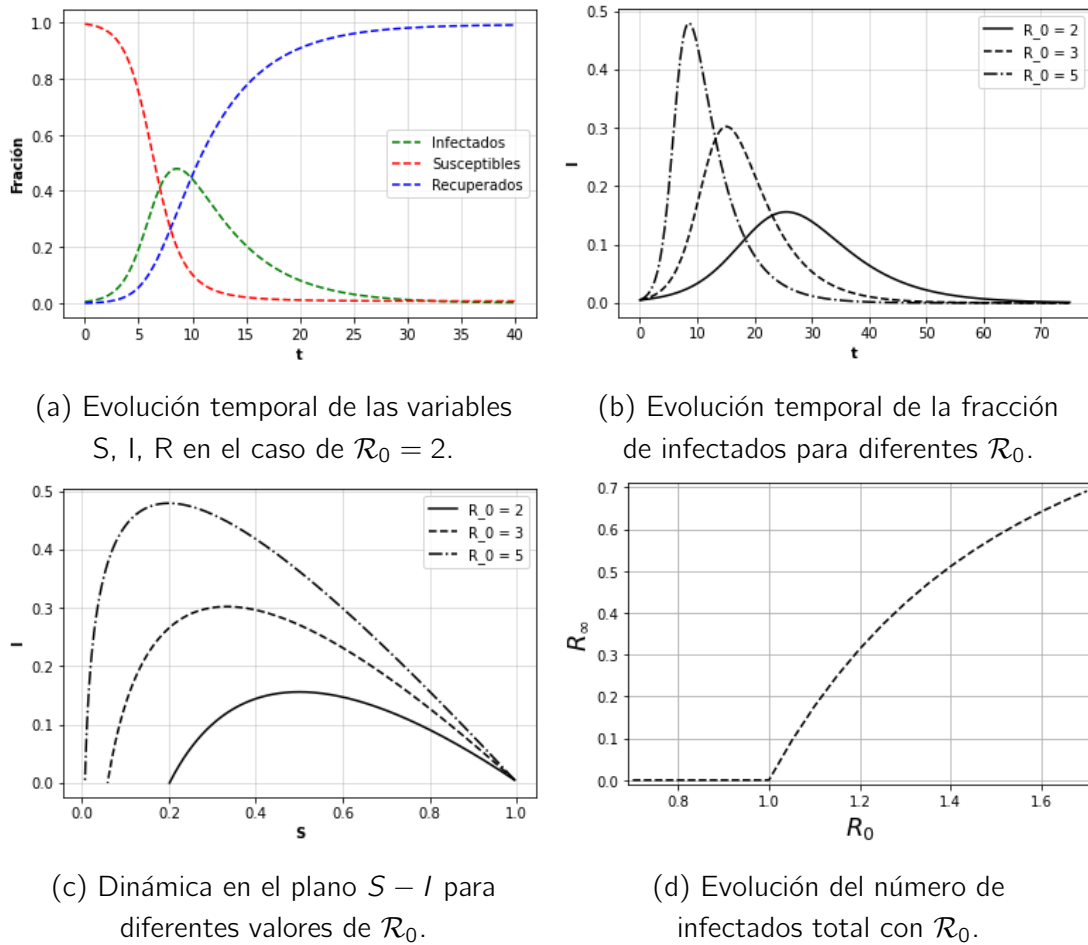


Figura 7: Dinámica del modelo SIR. Los resultados se han obtenido usando el algoritmo de integración Runge-Kutta de 4º orden. Las condiciones iniciales usadas son $S(0) = 0,995$, $I(0) = 0,005$ y $R(0) = 0$ excepto para la gráfica 7d, en la cual se han usado las condiciones $S(0) = 0,9995$, $I(0) = 0,0005$ y $R(0) = 0$.

Aunque la condición anterior sea válida en cualquier instante del tiempo, resulta especialmente interesante en el inicio de la epidemia. Si $\mathcal{R}_0 > \frac{1}{S(0)}$, el número de infectados inicial aumentará y la infección se propaga. En cambio, si $\mathcal{R}_0 < \frac{1}{S(0)}$ el número de infectados inicial disminuirá y la infección acaba muriendo tras un corto plazo de tiempo. Así, \mathcal{R}_0 determina por completo el comportamiento del sistema en función de las condiciones iniciales. Puesto que al inicio de la epidemia la mayoría de la población es susceptible en general, $S(0) \approx 1$ y se considera que el \mathcal{R}_0 crítico es 1.

Aunque la solución explícita de S , I y R en función del tiempo es compleja de obtener, podemos obtener cómo evoluciona $I(S)$. Para ello, aplicando la regla de la cadena y utilizando las ecuaciones (3) y (4), se obtiene la ecuación 7.

$$\frac{dI}{dS} = \frac{dI}{dt} \frac{dt}{dS} = \frac{\mu I - \lambda k S I}{\lambda k S I} = \frac{1}{\mathcal{R}_0 S} - 1 \quad (7)$$

La EDO anterior es resoluble por integración directa obteniendo la solución dada por la ecuación (8). Esta solución es útil puesto que nos permite obtener resultados como el número de infectados máximo o el número de personas infectadas total.

$$I(S) = I(0) + S(0) - S(t) + \frac{1}{\mathcal{R}_0} \ln \left(\frac{S(t)}{S(0)} \right) \quad (8)$$

Para obtener la fracción de infectados máxima, sustituimos el valor de $S(t)$ obtenido al imponer $\frac{dI}{dS} = 0$. El resultado de este procedimiento es

$$I_{max} = I(0) + S(0) - \frac{1 + \ln(\mathcal{R}_0 S(0))}{\mathcal{R}_0} \quad (9)$$

Por otra parte, para obtener el número de recuperados al final del proceso (que coincidirá con el número de infectados total y denominaremos prevalencia final) imponemos la condición de que $I_\infty(S_\infty) = 0$. Por simplicidad, podemos usar también que $I(0) + S(0) = 1$ y que $S_\infty + R_\infty = 1$. Al hacer esto en la ecuación (8) obtenemos la ecuación (10) como ecuación que nos da el valor de R_∞ .

$$R_\infty = \frac{1}{\mathcal{R}_0} \ln \left(\frac{1 - R_\infty}{S(0)} \right) \quad (10)$$

3.2. Modelo SIR. Dependencia con la estructura de la población

Aunque el modelo de campo medio resulta especialmente útil como herramienta con la que obtener resultados analíticos, no es capaz de reproducir las dinámicas reales [12]. Esto se debe a que las redes reales muestran interacciones heteróneas entre diferentes nodos ([13], [14]), lo que lleva a que la hipótesis de población bien mezclada no se cumpla.

Con el objetivo de tener en cuenta el efecto de las inhomogeneidades, debemos introducir en el modelo la estructura de la red [15]. Cada individuo estará representado por un nodo de la red y puede estar en uno de los estados S , I o R . Un nodo i podrá contagiar a otro nodo j con probabilidad λ si el nodo i está infectado, j es susceptible y existe un link que conecta i con j . Un nodo infectado podrá pasar al estado R con probabilidad μ .

Al introducir la red, resulta conveniente modificar las variables con las que trabajamos. En lugar de considerar las proporciones globales S , I , R , vamos a considerar las proporciones condicionadas a que el nodo tenga k contactos S_k , I_k y R_k .

$$\frac{dS_k}{dt} = -\lambda k S_k \Theta_k(t) \quad (11)$$

$$\frac{dI_k}{dt} = \lambda k S_k \Theta_k(t) - \mu I_k \quad (12)$$

$$\frac{dR_k}{dt} = \mu I_k \quad (13)$$

Además del ya mencionado cambio de variables S, I, R por otras que dependen de la conectividad del nodo, en las ecuaciones se introduce una función $\Theta_k(t)$, la cual nos dice la probabilidad de que uno de los k links que salen del nodo apunten a un nodo infectado. Si denotamos como $P(k|k')$ a la probabilidad de que un nodo con conectividad k' esté conectado con un nodo con grado k , $\Theta_k(t)$ puede calcularse según la definición 14.

$$\Theta_k(t) = \sum_{k'} P(k'|k) I_{k'}(t) \quad (14)$$

Supongamos ahora que tenemos una red sin correlaciones. Entonces, la definición 14 cambia a

$$\Theta^{NC}(t) = \frac{\sum_{k'} k' P(k') I_{k'}(t)}{\langle k \rangle} \quad (15)$$

Al igual que pasaba con las ecuaciones de campo medio, la situación estacionaria del sistema anterior es $I_{k\infty} = 0$. Esto nos indica que, tras un tiempo lo suficientemente grande, no tendremos infectados en la población, pero no nos da información sobre si va a producirse un aumento del número de infectados significativo a lo largo de todo el proceso.

Para estudiar esto último, nos situamos en el caso sin correlaciones y comenzamos integrando la ecuación (11). Si consideramos que, en la situación inicial $S_k(0) \approx 1$, se tiene la siguiente solución:

$$S_k(t) \approx \exp \left\{ -\lambda k \int_0^t \Theta^{NC}(t') dt' \right\} = \exp \{ -\lambda k \phi(t) \} \quad (16)$$

donde hemos definido una función auxiliar

$$\phi(t) = \int_0^t \Theta^{NC}(t') dt' = \frac{\sum_{k'} k' P(k') \int_0^t I_{k'}(t') dt'}{\langle k \rangle} \quad (17)$$

A través de la ecuación (13), podemos resolver la integral de la expresión anterior para obtener el siguiente resultado:

$$\phi(t) = \frac{\sum_{k'} k' P(k') R_{k'}(t)}{\mu \langle k \rangle} \quad (18)$$

Cuando se alcance el estado estacionario, podemos calcular la prevalencia final a través del número de personas susceptibles que quede en la población:

$$R_\infty = 1 - S_\infty = 1 - \sum_k P(k) S_{k\infty} \approx 1 - \sum_k P(k) e^{-\lambda k \phi_\infty} \quad (19)$$

donde, a partir de la ecuación 18, obtenemos que

$$\phi_\infty = \frac{\sum_{k'} k' P(k') (1 - S_{k'\infty})}{\mu \langle k \rangle} \approx \frac{1}{\mu} \left(1 - \frac{1}{\langle k \rangle} \sum_k k P(k) e^{-\lambda k \phi_\infty} \right) \quad (20)$$

La situación $\phi_\infty = 0$ siempre es solución de la ecuación (20). Al sustituir dicha solución en la ecuación (19), se obtiene $R_\infty \approx 0$; es decir, estamos en el caso en el que no hay aumento del número de infectados.

No obstante, existe otra posible solución distinta de 0 cuando se cumple la ecuación (21). Esto implica que el número de individuos contagiados ha aumentado y, por tanto, que la epidemia se ha propagado.

$$\left. \frac{d}{d\phi_\infty} \left(\frac{1}{\mu} - \frac{1}{\mu \langle k \rangle} \sum_k k P(k) e^{-\lambda k \phi_\infty} \right) \right|_{\phi_\infty=0} = \frac{\lambda \langle k^2 \rangle}{\mu \langle k \rangle} > 1 \quad (21)$$

Otra forma de escribir la condición (21) mediante el número reproductivo básico puede observarse en la ecuación (22). Este es un resultado importante, puesto que diferencia por completo la dinámica en la red Erdős-Rényi y en la red Barabasi-Albert.

$$\mathcal{R}_0 > \frac{\langle k \rangle^2}{\langle k^2 \rangle} \quad (22)$$

Recordemos que, en redes Barabasi-Albert lo suficientemente grandes, $\langle k^2 \rangle$ diverge. Esto conlleva que, para que haya propagación de la infección \mathcal{R}_0 debería ser mayor que 0, lo cual siempre se cumple. En cambio, las redes Erdős-Rényi siguen distribuciones de Poisson en las que $\langle k^2 \rangle = \langle k \rangle^2 + \langle k \rangle$, por lo que sigue existiendo umbral. Además, si $\langle k \rangle$ es lo suficientemente grande, $\langle k^2 \rangle \approx \langle k \rangle^2$ y recuperamos el umbral $\mathcal{R}_0 = 1$ de la situación de campo medio. Esta discusión puede verse ejemplificada en la figura 8.

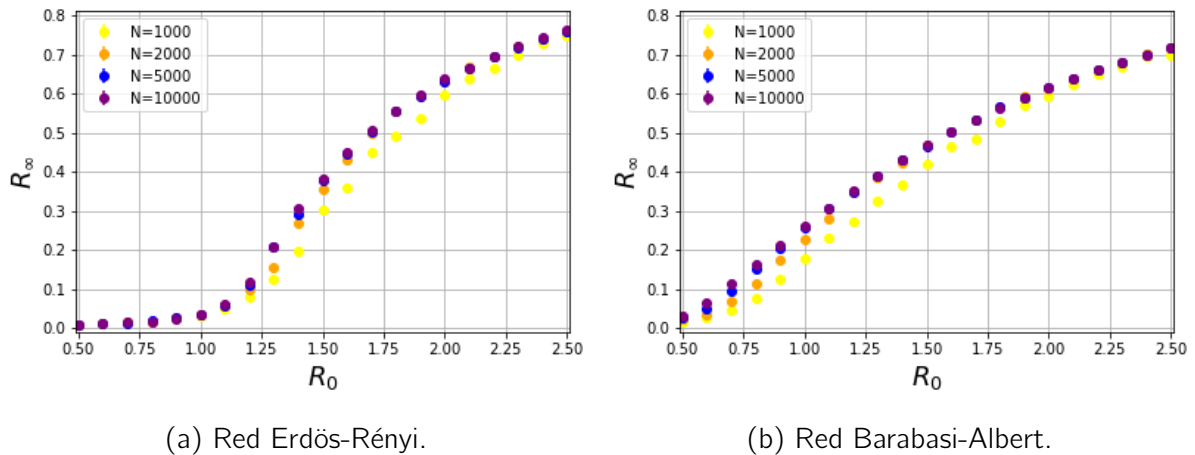
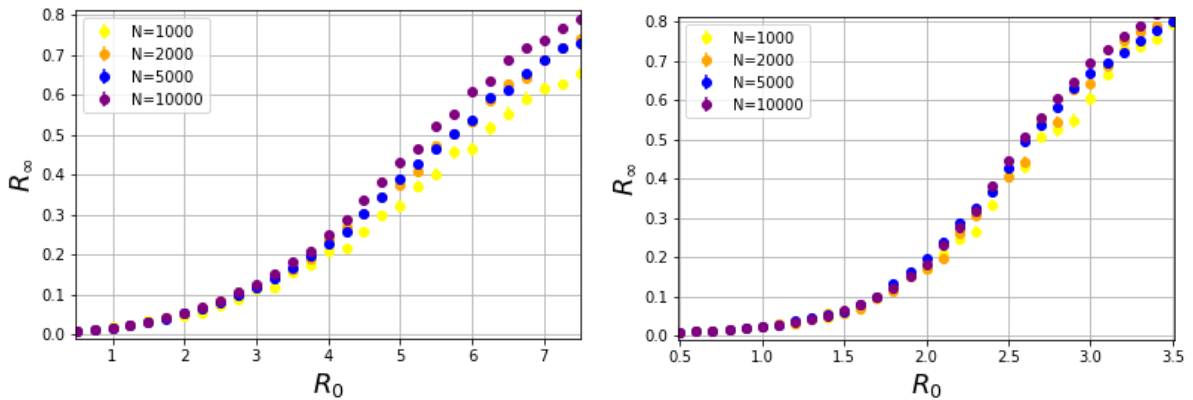


Figura 8: Evolución de la prevalencia final con el número reproductivo básico para diferentes topologías y tamaños de red. En todos los casos $\langle k \rangle \approx 6$. Los resultados se han obtenido mediante simulación computacional. En las figuras se puede ver la existencia de umbral en la red Erdős-Rényi y la ausencia de este en la red Barabasi-Albert.

La red Random Geometric Graph, al igual que la red Erdős-Rényi, sigue una distribución de Poisson. Por tanto, siguiendo el razonamiento anterior, debería también presentar una transición en $\mathcal{R}_0 \approx 1$. No obstante, en este tipo de redes se introducen correlaciones debido a la manera en la que se construye [16] y, por tanto, no se cumplen las hipótesis del desarrollo anterior.

A pesar de esto, sí se observa la existencia de una transición de fase [17]. De esta transición de fase se observa que tiene características similares a la teoría de percolación, puesto que las agrupaciones de nodos infectados son similares a los clusters formados en esta. Además se observa que el umbral no tiene un valor fijo como ocurría en la red Erdős-Rényi o en el modelo de campo medio sino que depende de la conectividad promedio: a mayor conectividad promedio, menor es el umbral (véase la figura 9).



(a) Random Geometric Graph con $\langle k \rangle = 6$. (b) Random Geometric Graph con $\langle k \rangle = 12$.

Figura 9: Evolución de la prevalencia final con el número reproductivo básico en una random geometric graph. En las figuras se puede ver la existencia de una transición de fase, cuyo umbral no coincide para los dos valores de $\langle k \rangle$.

4. Modelo y metodología

4.1. Modelo de población.

Los modelos de población que utilizaremos serán los presentados en la sección 2.3: red Erdős-Rényi, red Barabasi-Albert y Random Geometric Graph. Las poblaciones que simularemos serán redes de entorno a 5000 individuos y con $\langle k \rangle$ aproximadamente 6. En el caso de la red Random Geometric Graph simularemos también una población con $\langle k \rangle = 12$, con el objetivo de promover la formación de estructuras de comunidad y estudiar el efecto que tienen estas en la vacunación con DomiRank.

4.2. Modelo epidemiológico

Como modelo epidemiológico, utilizaremos el modelo SIR-V, el cual es una variante del modelo SIR introducido en la sección 3 al que se le añade una nueva clase: V, que representará a las personas vacunadas. Vamos a suponer que las personas vacunadas lo están desde el principio de la dinámica y que no pueden perder su inmunidad. De esta forma, la cadena de Markov representada en la figura 6 no cambia.

El único parámetro que prefijaremos en relación con la epidemia será la tasa de recuperación μ , la cual tomará un valor de 0,2. El valor de λ se variará sistemáticamente adaptando su valor al objetivo del análisis. En el proceso de vacunación usaremos en total 4 estrategias diferentes: una vacunación aleatoria, una vacunación por grado del nodo, una vacunación utilizando DomiRank con σ bajo (usaremos $\sigma = 0,05$) y otra con σ alto (usaremos $\sigma = 0,9$).

4.3. Metodología de simulación

Para la simulación usaremos un proceso de agentes. En cada paso temporal tenemos nuestra red con los nodos en estado S, I, R o V pudiendo variar de acuerdo al modelo presentado en la sección 3.1.

La red se crea según una de las 4 topologías presentadas. Tras crearla, nos quedaremos con la giant component; es decir, la agrupación de nodos conectados más grande posible. Para todo lo relacionado con la creación y gestión de la red usaremos la biblioteca Python *NetworkX*.

Tras la creación de la red, escogemos la métrica de vacunación que utilizaremos en dicha simulación y vacunamos a una fracción V de los nodos. Una vez vacunados, escogemos aleatoriamente un 0.5% de los nodos totales entre los no vacunados que actuarán como los infectados iniciales.

Desde el punto de vista de la aplicación del modelo, en cada paso temporal, crearemos una réplica exacta del estado de la red antes de hacer ninguna actualización de los estados. Basándonos en el estado de la red réplica, actualizaremos el estado de los nodos según las dinámicas propuestas por el modelo SIR. Esto se hace para evitar que si el nodo i se infecta en el paso t , sus vecinos puedan también infectarse en el paso t . Este proceso se repite hasta que no queden infectados en la red.

Al acabar la simulación, contaremos el número de nodos en el estado R. Este valor será el R_∞ obtenido. Para mejorar la bondad de los resultados, se realizarán varias simulaciones. Aunque el número concreto depende de la topología de la red y el fenómeno que se quiere estudiar, el número de simulaciones está en torno a las 100.

5. Resultados y discusión.

Puesto que hasta ahora no hemos hecho ningún análisis del efecto que tiene la existencia de nodos vacunados sobre la difusión de la enfermedad, vamos a comenzar haciendo un análisis general. Para ello nos apoyaremos en la figura 10, la cual muestra la evolución de la prevalencia al final de la simulación para escenarios con diferentes valores de \mathcal{R}_0 y diferentes estrategias de inmunización.

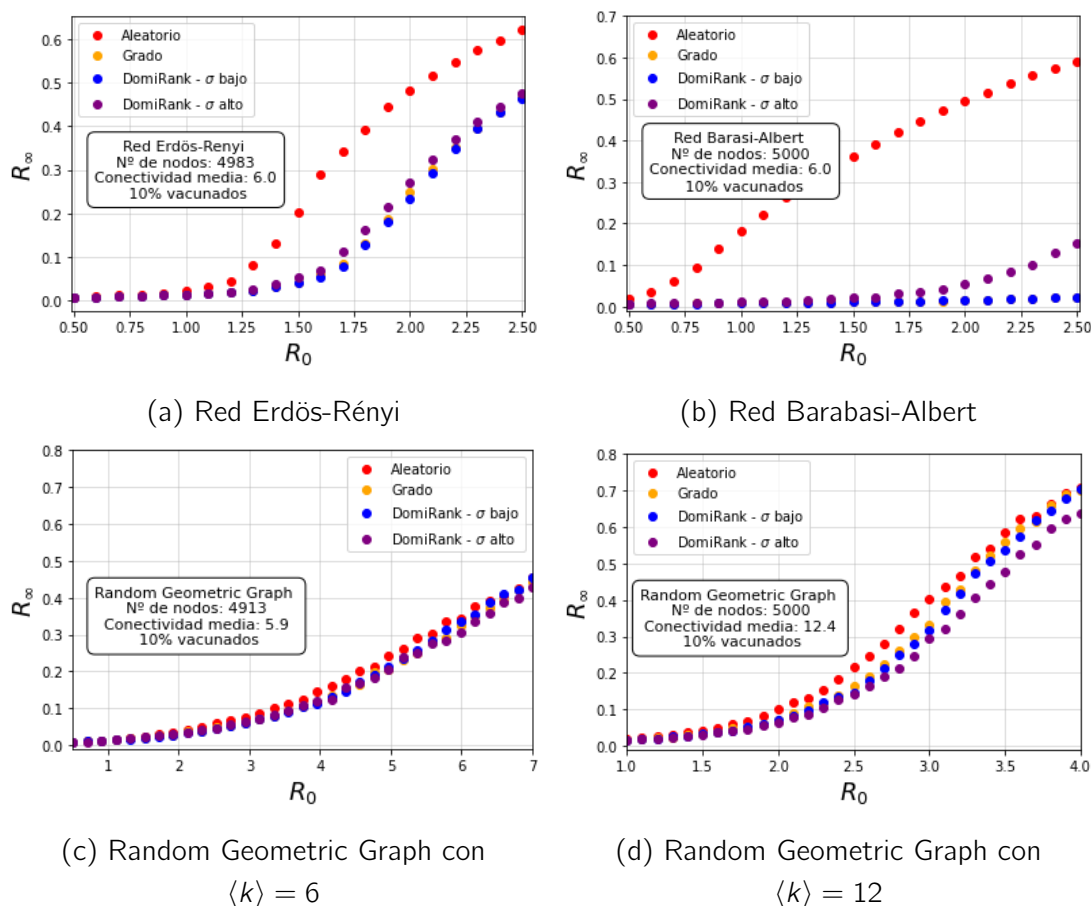


Figura 10: Evolución de la prevalencia epidemiológica con el número reproductivo básico \mathcal{R}_0 en presencia de vacunados para diferentes modelos de población. Nótese que las escalas de los ejes de abscisas no coinciden para las 4 topologías. La fracción de nodos vacunados en los 4 casos es el 10% de la población total.

Si comparamos los resultados de dicha figura con los obtenidos en las gráficas 8 y 9 vemos claramente como el efecto de la vacunación provoca una disminución de la prevalencia, tal y como era esperable. Sin embargo, más relevante aún que la disminución de la prevalencia es el aumento del umbral epidemiológico. Así, gracias a la vacunación, enfermedades que deberían propagarse por la población y causar un impacto notable en esta no logran difundirse. De esta forma se concluye que la vacunación puede actuar de forma eficaz como una medida de control contra diversas enfermedades.

5. Resultados y discusión.

Otro efecto que podemos observar en la figura 10 es la eficiencia de las distintas estrategias de vacunación dirigidas (basadas en alguna medida de centralidad) frente a la vacunación aleatoria. Aunque este fenómeno se observa en las 4 topologías, se aprecia más claramente en la red Erdős-Rényi y, principalmente, Barabasi-Albert. De hecho, en este último caso, una vacunación dirigida es necesaria si se quiere eliminar por completo la presencia de la enfermedad sin vacunar a un alto número de la población [18].

La eficacia de las estrategias de vacunación dirigidas se debe a que estas atacan directamente a la estructura de la red fragmentándola y haciendo más difícil la propagación de enfermedades. Por el contrario, una vacunación aleatoria ataca únicamente al número de personas susceptibles iniciales sin alterar la estructura, de tal forma que no hay una destrucción de las vías preferentes por la que el patógeno puede difundirse.

En el caso de las redes Random geometric graph, la diferencia entre ambos tipos de estrategias de inmunización no es tan significativa como en las otras dos topologías. Esto está provocado por la presencia de estructuras comunitarias. La existencia de estas hace que la destrucción de las vías de difusión preferentes no sea tan significativa al existir otras conexiones similares que han quedado intactas.

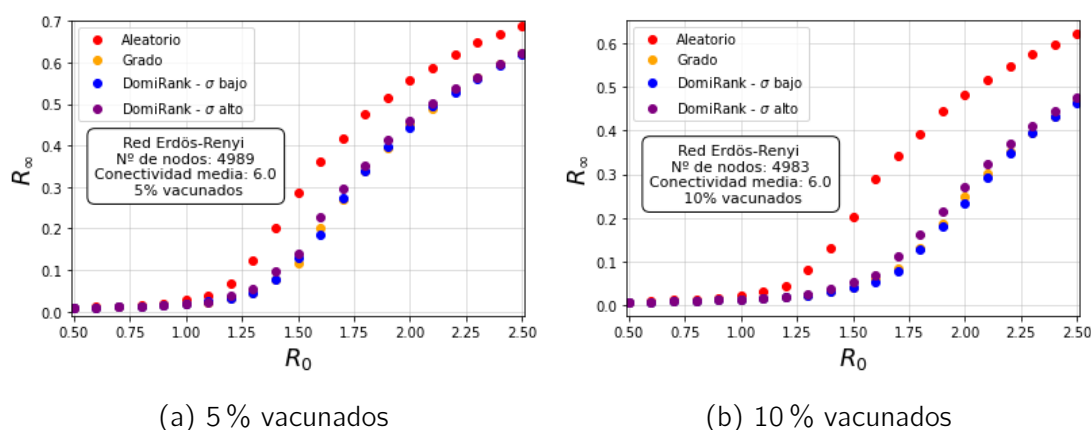


Figura 11: Comparación de la evolución de la prevalencia con el número reproductivo básico para diferente porcentaje de vacunados. Las simulaciones se han hecho sobre una red Erdős-Rényi.

En la figura 11 se aprecia el efecto que tiene vacunar a un mayor porcentaje de la población. Conforme este porcentaje aumenta, el efecto que tiene la epidemia en la población disminuye. Además, al aumentar el porcentaje de vacunados vemos también que el umbral epidemiológico aumenta. Así, resulta interesante definir el HIT (herd immunity treshold), el cual nos mide el porcentaje de la población que es necesario vacunar para evitar que una enfermedad con determinado \mathcal{R}_0 se propague. También se puede apreciar que la diferencia entre la vacunación dirigida y no dirigida aumenta con el número de vacunados.

5.1. Vacunación con prevalencia baja

Tras haber estudiado el efecto global que tiene la vacunación, vamos a estudiar en mayor detalle la eficiencia de las estrategias de vacunación basadas en las diferentes medidas de centralidad. Para ello, dependiendo del valor de \mathcal{R}_0 , separaremos las epidemias en dos clases: de baja prevalencia y de alta prevalencia. Consideraremos que la epidemia es de baja prevalencia si, al final del proceso sin vacunados, la población afectada está entorno al 20 – 30 % de los individuos totales. Por otro lado, consideraremos que la epidemia tiene una prevalencia alta si en condiciones de no vacunados, afecta al 70 % de la población total.

Con el objetivo de facilitar la comparación, entre las diferentes estructuras y estrategias, vamos a representar la prevalencia de la enfermedad con una fracción V de vacunados normalizada a la prevalencia sin vacunados en dicha red. Los resultados obtenidos para escenarios de baja prevalencia se encuentran en la figura 12.

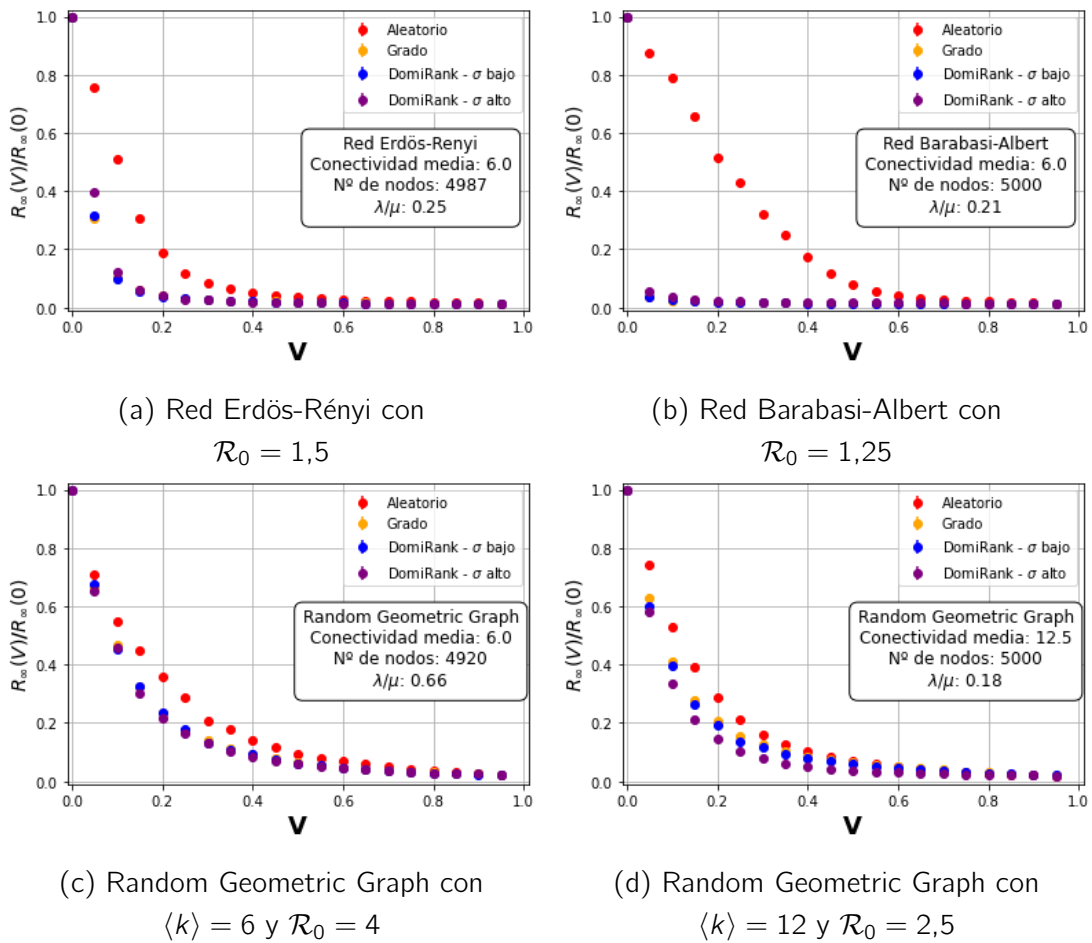


Figura 12: Evolución de la prevalencia epidemiológica con la fracción de vacunados para un régimen de baja prevalencia ($R_\infty \approx 0,25$ en el caso sin vacunados).

En primer lugar, cabe destacar que los resultados mostrados en la figura 12 nos corroboran los resultados que habíamos obtenido en la sección anterior. Por una parte, vemos que el incremento de la fracción de nodos vacunados provoca una disminución de la prevalencia. Esta disminución se produce de forma continua hasta llegar asintóticamente a un punto en el que R_∞ vale 0. Así, verificamos de nuevo la eficacia de las vacunas como forma de reducir el impacto de una enfermedad. Por otra parte, comprobamos que la vacunación no dirigida es menos eficiente que el resto de estrategias, independientemente del número de vacunados y de la topología de la población. Destaca el caso de la red Barabasi-Albert, en la cual el HIT está entorno al 60 % con una estrategia aleatoria de vacunación frente al 5 % de las estrategias dirigidas.

En lo referente a las estrategias de vacunación dirigidas, encontramos resultados dispares entre las cuatro topologías. Tanto en las redes Erdős-Rényi como Barabasi-Albert las estrategias basadas en medidas locales como lo son grado y DomiRank con σ bajo resultan más eficaces que la estrategia basada en DomiRank con σ alto. En el caso de la red Barabasi-Albert esto era lo esperable, ya que los elementos de la red que juegan un papel más relevante como propagadores son los hubs y estos están caracterizados por su alto grado. En el caso de la red Erdős-Rényi, la eficacia de estas medidas podría deberse a que este tipo de redes, al ser aleatorias, carecen de una estructura bien definida en la escala que alcanza la enfermedad en estos niveles de prevalencia. Al no tener este tipo de estructuras a las que atacar, esperamos que DomiRank con σ alto no fuera una buena medida.

No obstante, la situación que se da en las redes Random Geometric Graph es justo la contraria. En estas dos topologías, la métrica basada en DomiRank con σ alto ofrece mejores resultados que las medidas locales. Se aprecia también como la diferencia entre ambos tipos de estrategias se hace más significativa al pasar de $\langle k \rangle \approx 6$ a $\langle k \rangle \approx 12$. Así, encontramos que la existencia de estructuras de comunidad y favorece el uso de DomiRank como forma de evitar que la epidemia se propague a lo largo de la red.

5.2. Vacunación con prevalencia alta

Pasemos por último a analizar los resultados obtenidos para epidemias con mayor impacto en la población global. Los datos obtenidos para escenarios de alta prevalencia aparecen recopilados en la figura 13.

Comparando estos resultados con los obtenidos para enfermedades de baja prevalencia, vemos que estos son muy similares. De nuevo tenemos que en las redes Erdős-Rényi y Barabasi-Albert las estrategias basadas en propiedades locales del nodo son más eficaces mientras que en las redes random geometric graph la métrica más eficiente es DomiRank con σ alto. Así, aún a pesar de que la enfermedad tenga mayor alcance en la red, las propiedades topológicas relevantes siguen siendo las mismas.

5.2 Vacunación con prevalencia alta

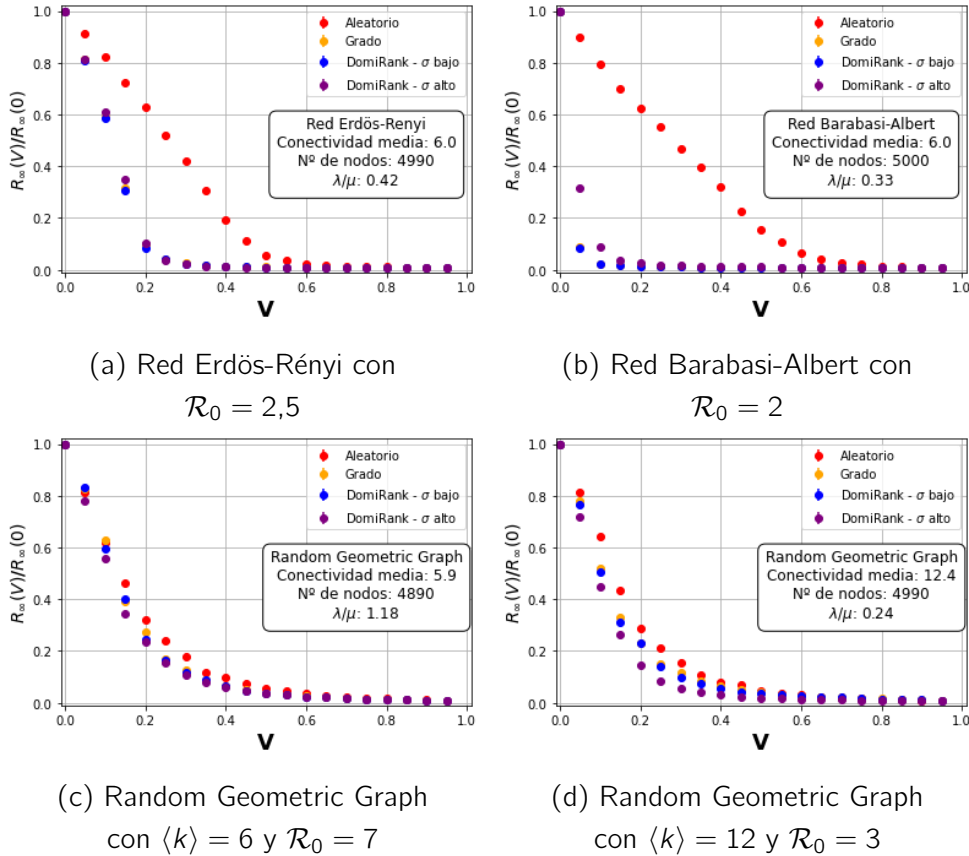


Figura 13: Evolución de la prevalencia con la fracción de vacunados para un régimen de alta prevalencia ($R_\infty \approx 0,7$ en el caso sin vacunados).

Sin embargo, no todo es igual. El aumento del número reproductivo básico se traduce, en general, en una mayor desigualdad entre las diferentes estrategias. Así, aunque las estrategias óptimas para cada modelo de población no hayan cambiado con respecto al caso de baja prevalencia, estas se han vuelto más óptimas en comparación con el resto. Vemos así que encontrar una estrategia adecuada resulta fundamental cuando hablamos de enfermedades capaces de afectar a un gran número de personas, tanto por el coste personal y social como por el coste económico.

No obstante, a pesar de lo comentado en los párrafos anteriores, en la figura 13a se observa como la diferencia entre DomiRank con σ alto y las métricas locales en el caso concreto de la red Erdős-Rényi ha disminuido. Aunque esto bien podría tratarse de un resultado válido únicamente para dicha simulación; podría ser también un indicativo de que a muy altos valores de \mathcal{R}_0 DomiRank con un σ alto sea la medida de centralidad que mejores resultados ofrece, lo cual se ve reforzado con la figura 14. Esto se produce debido a que, dado el alto valor de \mathcal{R}_0 , la epidemia tiene un alcance global en la red y DomiRank con σ alto captura la información estructural relevante en el proceso dinámico.

6. Conclusiones.

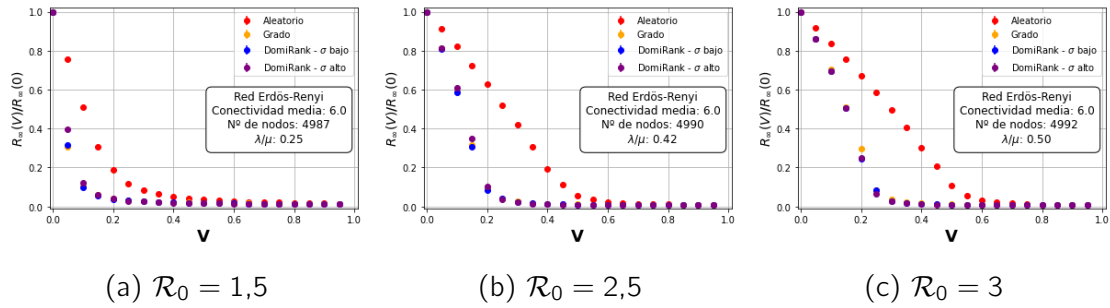


Figura 14: Comparación de las diferentes estrategias de inmunización sobre redes Erdős-Rényi para diferentes valores de \mathcal{R}_0 . Se observa como al aumentar \mathcal{R}_0 , DomiRank con σ alto funciona mejor que grado.

6. Conclusiones.

El propósito de este trabajo era el estudio de la efectividad de la nueva métrica de centralidad como base de una estrategia de vacunación frente a epidemias. Mediante técnicas computacionales, hemos comparado el efecto que tiene esta nueva técnica de inmunización con las estrategias basadas en el grado de los nodos y vacunación aleatoria.

Los resultados obtenidos para diversas topologías reflejan que DomiRank propone una alternativa a estas estrategias tradicionales igual de eficaz. En algunos casos incluso, ofrece mejores resultados que el resto de alternativas.

Sin embargo, la verdadera utilidad de DomiRank reside en la posibilidad de ajustarse a las características del patógeno y de la red. En redes caracterizadas por las propiedades individuales de cada nodo, DomiRank con σ bajo recupera los buenos resultados obtenidos mediante la vacunación por grado. En cambio, cuando la red está caracterizada por las estructuras comunitarias, DomiRank con σ alto destaca sobre sus competidores.

A pesar de los resultados positivos obtenidos en este trabajo, todavía quedan puntos que resolver en futuros estudios. Entre ellos, podríamos encontrar replicar este estudio con otros modelos epidemiológicos que recogen otras características diferentes. Más interesante podría ser aplicar DomiRank sobre algunas redes reales sobre las que se tienen datos, como poblaciones de algunos países o algunas redes digitales. Algunos autores han apuntado a que el factor actual limitante no es tanto el acceso a las vacunas como la opinión de la población frente a estas, por lo que sería conveniente realizar un análisis de un modelo como el propuesto en este trabajo acoplado a un modelo de opinión. Sin embargo, quizás la línea de investigación más necesaria sería idear una forma de poder conocer aquellos nodos con valor DomiRank más alto sin conocer la versión completa de la matriz de adyacencia.

Disponibilidad del código.

El código usado para simular la dinámica del SIR se encuentra disponible en este [repositorio](#).

Referencias

- [1] P.W. Anderson. More is different. *Science*, 177:393-396, 1972.
- [2] K. Marx. Capital: The process of capitalist production.
- [3] Y. Moreno E. Fofoula-Georgiou C. Kasmir M. Engsig, A. Tejedor. Dominant centrality reveals structural fragility of complex networks via node dominance. *Nature communications*, 2024.
- [4] M. Barthélémy A. Barrat and A. Vespignani. Dynamical processes on complex networks. *Cambridge University Press, New York*, 2008.
- [5] M. Newman. Networks: An introduction. *Oxford University Press, Oxford*, 2010.
- [6] G. F. de Arruda et al. Role of centrality for the identification of influential spreaders in complex networks. *Phys. Rev. E* 90, 032812, 2014.
- [7] A. Rényi P. Erdős. On random graphs i. *Publicationes mathematicae debr.* 6, 290, 1959.
- [8] Z. Néda E. Ravasz A. Schubert T. Vicsek A.-L. Barabási, H. Jeong. Evolution of the social network of scientific collaborations. *Physica A* 311. pag 590 – 614, 2002.
- [9] R. Albert A. L. Barabasi. Emergence of scaling in random networks. *Science* 286, 509, 1998.
- [10] J. Dall. M. Christensen. Random geometric graphs. *Physical Review E.* 66 (1 Pt 2): 016121, 2002.
- [11] A. G. McKendrick W. K. Kermack. A contribution to the mathematical theory of epidemics. *Proceedings of the Royal society of London. A* 115 700–21, 1927.
- [12] A. Vespignani R. Pastor-Satorras. Epidemic spreading in scale-free networks. *Phys. Rev. Lett.* 86, 3200, 2001.
- [13] F. Liljeros et al. The web of human sexual contacts. *Nature*, 441, 907, 2001.
- [14] A.-L. Barabási R. Albert, H. Jeong. The diameter of the world wide web. *Nature*, 401, 130, 1999.

- [15] A. Vespignani Y. Moreno, R. Pastor-Satorras. Epidemic outbreaks in complex heterogeneous networks. *Eur. Phys. J. B* 26, 521–529, 2002.
- [16] M. Tomassini A. Antoni. Degree correlations in random geometric graphs. *Phys. Rev. E* 86, 037101, 2012.
- [17] A. Sensharma D. Saha, S. Mitra. Critically spanning epidemic outbreak cluster in randomgeometric networks. *Physica A* 629, 2023.
- [18] A. Vespignani R. Pastor-Satorras. Inmunization of complex networks. *Phys. Rev. E* 65, 036104, 2002.