

El Modelo SIR aplicado a Redes Complejas

Carlos Elishah Candanosa Salazar¹ y Maximiliano Rojas Zaputovich¹

¹Universidad Nacional Autónoma de México

Noviembre de 2025

Resumen

En el presente escrito se muestra una generalización del modelo epidemiológico SIR (Susceptible-Infectado-Recuperado) clásico, pero aplicado a redes complejas, específicamente redes libres de escala tipo Barabási-Albert, con el objetivo de simular la propagación de enfermedades sin recurrir a idealizaciones como aquella de una sociedad bien mezclada, siempre con el objetivo de reducirse al caso clásico en dicho límite y para evolución temporal continua (a diferencia de la discretización que se utilizó en el presente trabajo). En ese sentido, se analizó el impacto de la topología de la red en la dinámica de contagio y de igual forma, se identificó el impacto de infectar a nodos con diferentes centralidades sobre el desarrollo de la pandemia, en particular, utilizando las métricas de Grado, Intermediación (Betweenness), Cercanía (Closeness), Eigenvector, Katz y Carga (Load). Las simulaciones numéricas revelaron que, a diferencia del modelo clásico que tiende a sobreestimar el alcance de la infección, el enfoque de redes permite observar fenómenos de contención local, que es precisamente lo que se observa en la vida real. Asimismo, se demostró que la centralidad del paciente cero influye significativamente en la velocidad de propagación y el tiempo hasta el pico epidémico, pero tiene un impacto marginal en la magnitud total acumulada de la pandemia. Estos resultados sugieren, por supuesto, que, en redes libres de escala, la estructura global y las características del agente infeccioso predominan sobre las características individuales del nodo inicial.

Keywords: *Modelo SIR, Redes Complejas, Barabási-Albert, Centralidad, Dinámica de Epidemias, Agentes Infecciosos.*

1. Introducción

Desde tiempos inmemoriales, la humanidad ha enfrentado un sinfín de obstáculos en su desarrollo: yendo desde la escasez de recursos y guerras mundiales, hasta problemas más contemporáneos como el cambio climático, desigualdad económica o incluso crisis devenidas del rápido crecimiento tecnológico inherente a nuestra especie. Sin embargo, desde la prehistoria, ha habido un enemigo con el cual nunca hemos dejado de luchar; uno que no sólo es inherente a la humanidad, sino a cualquier ser vivo presente en la Tierra: la enfermedad [2].

Es por ello que la modelación de infecciones es fundamental para la correcta comprensión de su comportamiento en la sociedad: esto permitiría planear estrategias para combatirlas, tácticas para que los hospitales puedan atender a la mayor cantidad de personas durante un brote, y claro, formas de prevenir el contagio, lo cual desde el 2020 ha sido crucial para minimizar pérdidas acaecidas por la Covid-19 [5]. Existen muchos modelos que han tratado, con cierto éxito, de modelar el espaciamiento de enfermedades, como el SI, que modela la transmisión de la enfermedad por medio de personas susceptibles e infectados (de ahí las siglas), que no es realista, puesto que nadie se recupera una vez la infección ha sucedido, o modelos más certeros como el SIR, que se discutirá más adelante y es en el que se sustentará este proyecto, en donde a la población se le divide en susceptibles, infectados y recuperados, donde ya se incluye curación de la enfermedad, o incluso el SIRS, que añade reinfección, que permite modelar epidemias a grandes escalas temporales (véase [13]), pero todos estos modelos se basan en el concepto de una población bien mezclada, es decir, que todos los individuos están en contacto con todos los demás, lo cual ciertamente es un obstáculo de estos modelos, pues definitivamente la sociedad no es un conjunto de personas que se conocen todos entre sí, sino que en ella se crean lazos o vínculos afectivos con individuos que consideran afines, lo cual modificaría la transmisión de la enfermedad. [1].

Las redes complejas nos permiten solucionar este problema. Dado que son la manera más natural de representar interacciones entre individuos de manera fidedigna, aplicar la fenomenología y el marco teórico, dentro de los alcances de un trabajo como el presente, de modelos como el clásico SIR, sin llegar al extremo de idealizar a la población a estudiar como un conjunto bien mezclado de personas, que es lo que precisamente se ha hecho por años, especialmente desde el surgimiento de la Covid-19 (véase [3]), es el principal objetivo que se tiene en este escrito, de manera que la modelación de una epidemia

o incluso una pandemia sea tan realista como sea posible, incluyendo conceptos como la contagiosidad y curabilidad de una enfermedad, así como su transmisión a lo largo de cierto grupo en función de la **importancia** del agente 0 en el mismo, esto es, por ejemplo, cuánta gente conoce y con cuánta interactúa, que evidentemente se relaciona al concepto de **centralidad en la ciencia de redes**.

Otro problema con los modelos clásicos, aparte de la "buena mescolanza", es aquel de la visibilidad. Si bien los modelos clásicos son excelentes para medir la cantidad de personas en cada compartimento y cómo evolucionan en estos mismos, indudablemente carecen de un factor visual: no se puede inferir en qué zonas está habiendo más contagios, ni en qué tipo de situaciones suceden estos. Por lo tanto, es con una red compleja que se puede solucionar este problema, ya que se podrá entender bajo qué clase de relaciones sociales, a través, por ejemplo de la métrica de betweenness para simular puentes sociales, se tiene que hay más propensión de contagio (personas con muchos conocidos, personas con conocidos con muchos conocidos, y así sucesivamente). En [6] ya se intentó hacer algo similar, pero se tuvo el problema de no mostrar la red (y por lo tanto, no brinda información visual acerca de la propagación de la enfermedad). Así que, si bien adapta la teoría del SIR al lenguaje de las redes complejas (básicamente lo que hace es un modelo SIR por cada centralidad de grado), vuelve a caer en el problema de no proveer un recurso visual para entender más a fondo la propagación de la enfermedad.

En pocas palabras, este trabajo pretende mezclar los modelos antes discutidos, o variaciones de los mismos, junto con la teoría de redes, y verificar la importancia que tiene la distribución de infectados en la sociedad en relación con el desarrollo de la enfermedad. El incorporar estas relaciones, y no reducir a la sociedad como una sopa de individuos en constante interacción, es la principal utilidad del escrito, y los resultados del mismo, de corroborar lo observado en epidemias reales o asemejarse a lo propuesto por el modelo SIR clásico, serán una forma eficiente de modelar a una epidemia y de hacer planes para atacarla, e incluso, prevenirla.

2. Antecedentes

2.1. Modelo Compartimental en Modelación de Epidemias [7]

El Modelar Epidemias es una de las herramientas más cruciales en la época moderna, pues presenta una forma de predecir la evolución de un virus o bacteria en determinada población a estudiar. El Modelo Compartimental, en específico el SIR (por las siglas Susceptible-

Infectado-Recuperado) es precisamente la manera más clásica de poder simular matemáticamente el comportamiento general de un agente patógeno en un sector poblacional en particular. Se basa en lo que sigue: **Se supone que se tiene una población constante de N individuos bien mezclada** y esa población se divide en 3 sectores, **Susceptibles** a contagio, S , **Infectados**, I , y **Recuperados** tras la infección, R , (aunque esto es fácilmente extendible al caso de enfermedades letales, pues simplemente se hace un cuarto compartimiento de Fallecidos, y como tal la población seguiría siendo constante, pero en el paper que se toma de referencia para esta sección no se considera). Claramente, mientras más gente infectada haya en comparación con la población general, más sencillo será que la gente susceptible enferme, es decir, que decrecerá la población de susceptibles más rápido. Esto se traduce en que se tiene lo siguiente para la derivada de S en función del tiempo:

$$\frac{dS}{dt} = -\beta \frac{I}{N} S \quad (1)$$

Donde β es un parámetro de escalamiento. El caso para R es análogo: mientras más gente infectada hay, más rápido crecerá la cantidad de gente recuperada, aunque en este caso ya no hay que compararla con la población total, pues el proceso de recuperación es individual. Esto es:

$$\frac{dR}{dt} = \gamma I \quad (2)$$

Y claro, la tasa de cambio de gente infectada serán los negativos de las dos tasas anteriores, pues la gente susceptible que se pierde, se va al compartimiento de personas infectadas, y las personas recuperadas no son más que individuos que salieron de la infección, así que:

$$\frac{dI}{dt} = \beta \frac{I}{N} S - \gamma I \quad (3)$$

El paper hecho por Carlson, Rubin, Heikkilä y Postema hizo uso de estas ecuaciones para modelar la epidemia de Covid-19 durante 2021 a partir de datos reales de gente infectada en el año 2021 y su evolución con el tiempo (no consideraron a la letalidad un factor importante, puesto que la Covid-19 tiene un índice relativamente bajo de letalidad, 2%, y eso justo al comienzo de la pandemia). Propusieron a $N = 7.8 \times 10^9$, lo cual es la población mundial aproximada, y utilizaron datos ofrecidos por la OMS para la evolución temporal de la enfermedad para ajustar los parámetros β y γ a sus ecuaciones, obteniendo, en promedio dado que los parámetros también son funciones del tiempo, $\beta = 0.33\text{semana}^{-1}$ y $\gamma = 0.23\text{semana}^{-1}$ lo cual era consistente con la investigación documental que realizaron sobre la modelación de la pandemia, y siendo compatible con el desarrollo de la pandemia durante el año 2021, mostrando la aplicabilidad de su modelo, y de hecho concluyendo que su modelo es **el primero** que permite hacer una regresión al momento inicial de la pandemia y verificar la cantidad de individuos inicialmente infectados que dieron lugar a la pandemia. En general este fue un acercamiento valioso al estudio del desarrollo de la COVID-19 en el mundo, pero se considera que aún podría agregarse un aspecto: analizar que en general la población no está bien mezclada.

El propósito del presente escrito, pues, es trabajar mayoritariamente con simulaciones (dado que los trabajos en este ámbito así lo hacen, y sólo un puñado recurre a datos reales para corroborar que el modelo funcione) de enfermedades con diferentes parámetros β y γ incorporando poblaciones que no están bien mezcladas (a través de una red) y observar las curvas de infección al variar dichos parámetros, los enlaces y centralidades de la gente infectada, y el tamaño de la población total observando, entre otras cosas:

- Qué sucedería al aumentar mucho la contagiosidad de la enfermedad.
- Cómo depende la propagación de la enfermedad de la centralidad del agente 0.

- Cómo varía la propagación de la enfermedad al disminuir o aumentar la densidad de la red (para redes completas, se tendría que regresar al modelo SIR).

2.2. Redes Complejas en Modelación de Epidemias [8]

La razón principal por la cual una epidemia se convierte en una pandemia es la conectividad entre las diferentes regiones del mundo. Por lo tanto, una vez que una epidemia entra a una región, la distribución local de la población es crucial para decidir el grado en que crecerá la enfermedad, que es algo que en el paper anterior no se discutió.

El paper estudiado para esta sección propone una extensión al modelo SIR, el cual considera la distribución de población en todo el mundo. Para esto, además de los parámetros clásicos antes mostrados, se consideran los siguientes:

- Un parámetro α , llamado el *parámetro de conectividad social*, que indica qué tan socialmente están conectados los individuos.
- Otro parámetro $c_{(i,j)}$, que es la movilidad de los individuos de la región i a la j .

En general, la principal pregunta por ser contestada utilizando este modelo es:

- ¿Cómo el parámetro de “conectividad social” α afecta a los individuos de los diferentes compartimentos?

En contraste al paper anterior (que es el SIR tradicional), este modelo se opone a la idea de una sociedad “bien mezclada” y opta por una distribución obtenida a través de redes complejas. Una vez añadido el parámetro α , las ecuaciones SIR adquieren la siguiente forma:

$$\frac{dS}{dt} = -\beta \frac{I}{N} S - \frac{\alpha S n N k \frac{I}{N} \beta}{N + n N}$$

Donde el nuevo término, dado que nótese que el primero es el que ya conocíamos de la sección anterior, es para considerar los contagios adicionales que pueden surgir a raíz de personas de otras localidades (y por eso es que entra en juego el término α). En el caso de la población recuperada, dado que es un proceso individual y no depende de la conectividad entre personas, se da por:

$$\frac{dR}{dt} = \frac{\gamma I}{N}$$

La cual claramente no cambia, salvo el hecho de que está normalizada (por eso es diferente a la del paper anterior, pero matemáticamente es equivalente, redefiniendo a $\tilde{\gamma} = \frac{\gamma}{N}$, sólamente se hace para tener simetría en las ecuaciones). Por último, la tasa de crecimiento de infectados se da por:

$$\frac{dI}{dt} = \frac{\beta S I}{N} + \frac{\alpha S n N \frac{I}{N} \beta}{N + n N} - \frac{\gamma I}{N}$$

Donde, una vez más, los términos adicionales son para ajustar posibles contagios dependientes de la “sociabilidad”, y se escoge de tal manera que la población total permanezca constante.

Para la simulación con redes, se usó una red con 10,000 nodos, notando que el SIR ampliado se comporta como el SIR tradicional en dos situaciones:

1. Cuando $\alpha = 0$
2. Si no hay movilidad ($c_{(i,j)} = 0$)

Ya una vez que se tiene al modelo, lo que importan son las simulaciones (que es justamente en lo que se desea trabajar mayoritariamente en el proyecto). Este paper, por ejemplo, utiliza distintos parámetros para α , pero también de movilidad, $c_{(i,j)}$, notando que lo crucial aquí no son los resultados, sino la adaptación del modelo a redes complejas. La simulación realizada por los autores empieza a partir de un nodo aleatorio, pero el objetivo en el proyecto a realizar es su elección acorde a distintas métricas como centralidad de grado o de Katz,

esto es, que se planea llegar a un punto medio entre ambos modelos, debido a que hay parámetros que no son necesarios para los objetivos planteados en el mismo.

3. Marco Teórico

A continuación se presenta la teoría relevante para el proyecto.

3.1. Redes

Definición 1 (Red) Una red R es un par ordenado (V, E) donde V es el conjunto de nodos y $E \subseteq V \times V$ es el conjunto de aristas que conectan los nodos. Además, al menos para este proyecto, no habrá nodos autoconectados, es decir, para todo $v \in V$, $(v, v) \notin E$, y siempre tomaremos a las redes finitas, $|V| < \infty$. [4]

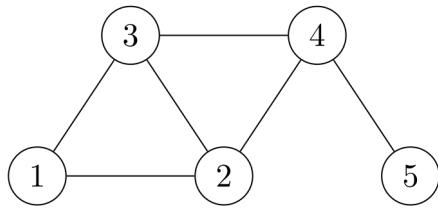


Figura 1. Red Simple.

Esto da pie a definir a las **métricas**, las cuales son esenciales ya que nos permitirán estudiar las diferentes “posiciones sociales” que pueden tener los individuos que conforman a la red.

3.2. Métricas

3.2.1. Centralidad de grado

Definición 2 Definimos la **centralidad de grado** de un nodo i , $\deg(i)$ como el número de aristas que inciden en él [4]. De manera formal, podemos definirla como la cardinalidad del conjunto de nodos adyacentes a i :

$$\deg(i) = |\{j \in V : (i, j) \in E\}|$$

O, equivalentemente:

$$\deg(i) = \sum_{j \in V} \chi_{(i,j) \in E}$$

Así pues, en lo que concierne al proyecto, un nodo con alta centralidad de grado representará a los individuos con mayor número de conocidos directos.

3.2.2. Centralidad de intermediación y carga

Antes de definir estas métricas, es necesario definir primero el concepto de camino más corto.

Definición 3 (Camino más corto) Dado un par de nodos (i, j) , un **camino más corto** entre ellos es una sucesión de nodos $(i = v_0, v_1, v_2, \dots, v_k = j)$ tal que para cualquier otra sucesión de nodos que empieza en i y termina en j , $(i = u_0, u_1, u_2, \dots, u_m = j)$ se cumple que $k \leq m$. Es decir, el camino más corto es aquel por el cual se pasan por la menor cantidad de nodos para llegar de i a j . [4]

Ya con esta definición, podemos proceder a definir nuestras métricas deseadas:

Definición 4 (Centralidad de intermediación) Dado un nodo i , su **centralidad de intermediación** (o **betweenness**, del inglés) se define como:

$$BC(i) = \sum_{j,k} \frac{\rho(j,i,k)}{\rho(j,k)}$$

Donde $\rho(j, k)$ es el número de caminos más cortos entre los nodos j y k , y $\rho(j, i, k)$ es el número de caminos más cortos entre j y k que pasan por i . [4]

Es decir, mide cuántos caminos más cortos pasan por un nodo determinado. No sólo eso, sino que lo hace en relación a cuántos caminos más cortos hay en general. En nuestro contexto, esto medirá a todas las personas que actúan como “puente”.

Una métrica similar, pero no equivalente, es la centralidad de carga. Esta, de cierta manera, también mide la cantidad de caminos más cortos que pasan por un nodo, pero toma en cuenta de manera más explícita la cantidad de posibles caminos más cortos que hay en la red, y las posibles bifurcaciones que podrían existir entre ellos, así como en la siguiente figura:

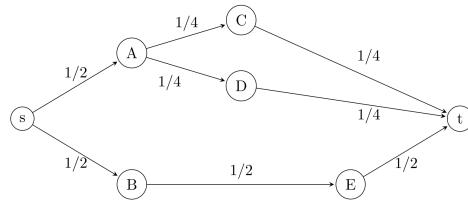


Figura 2. Distribución de peso a través de los diferentes caminos más cortos de s a t

Por lo tanto, tomando en cuenta la gráfica anterior como ayuda, la centralidad de carga se define de la siguiente manera:

Definición 5 (Centralidad de carga) Dado un nodo i , su **centralidad de carga** se define como:

$$LC(i) = \sum_{j,k} \theta_{j,k}(i)$$

Donde $\theta_{j,k}(i)$ es la cantidad de carga que envía i cuando se considera el camino más corto de j a k que pasa por i , así como en la figura de arriba. Podría aparecer este término múltiples veces, ya que nuestro nodo podría ser como el A de la gráfica antes dibujada. [11]

3.2.3. Centralidades espectrales

Estas métricas son aquellas que utilizan nociones del álgebra lineal para poder ser calculadas. Para ello, primero tenemos que definir la matriz de adyacencia de una gráfica:

Definición 6 (Matriz de adyacencia) Sea $G = (V, E)$ una red tal que sus nodos estén enumerados: $V = \{1, \dots, n\}$. Definimos entonces su matriz de adyacencia como:

$$A_{ij} = \begin{cases} 1 & \text{si } (i, j) \in E \\ 0 & \text{si } (i, j) \notin E \end{cases}$$

Así pues, la matriz de adyacencia es un “indicador” de cuáles aristas se encuentran en la red. [4]

Definición 7 (Centralidad de Katz) Dada una red G , su matriz de adyacencia A y un nodo i , definimos su **centralidad de Katz** como:

$$K_i = [\sum_{k=0}^{\infty} (\alpha^k A^k) e]_i$$

Donde $e = (1, \dots, 1)$ y $\alpha < \rho(A)$ (el radio espectral de A , el eigenvalor más grande de A). [4]

A grandes rasgos, esta centralidad mide la cantidad de caminos de tamaño arbitrario que tiene un nodo, tomando α como un factor de atenuación, de manera que los caminos cortos tienen más peso. Por lo tanto, en este proyecto, los individuos (nodos) con alta centralidad de Katz serán aquellos que tienen un gran alcance hacia personas que no conocen, a través de intermediarios.

A continuación definiremos la centralidad de eigenvector. Para ello, necesitamos del siguiente teorema:

Teorema 1 (Teorema de Perron-Frobenius) Sea A una matriz de entradas reales no negativas. Entonces

1. A tiene un eigenvalor máximo positivo, k_1 .
2. Existe un único eigenvector \mathbf{x} asociado a k_1 tal que todas sus entradas son positivas y $\|\mathbf{x}\|_1 = 1$. [4]

Ya con este teorema, podemos definir la centralidad de eigenvector de una manera muy sencilla:

Definición 8 (Centralidad de Eigenvector) Sea G una red y A su matriz de adyacencia. Definimos la centralidad de eigenvector del nodo i como \mathbf{x}_i , donde \mathbf{x} es el vector dado por el teorema de Perron-Frobenius. [4]

Dado que la centralidad de eigenvector está definida de manera implícita mediante la ecuación $A\mathbf{x} = k_1\mathbf{x}$, podemos notar que es equivalente a $\mathbf{x}_i = k_1^{-1} \sum_j A_{ij}x_j$, así que podemos ver que, salvo constantes, la centralidad de eigenvector es la suma de las centralidades de eigenvector de sus vecinos. Así pues, un individuo con alta centralidad de eigenvector será aquel que tiene muchos conocidos de alta influencia.

3.2.4. Centralidad de cercanía

La centralidad de cercanía mide qué tan cerca se encuentra un nodo con relación a todos los demás de la red. Los nodos con alta centralidad de cercanía tendrán muchos caminos cortos, mientras que aquellos con baja centralidad de cercanía tendrán caminos más largos. Formalmente, se define como:

Definición 9 (Centralidad de cercanía) Sea G una red con n nodos e i un nodo de G . Definimos la centralidad de cercanía de i como:

$$CC(i) = \frac{n - 1}{s(i)}$$

Donde $s(i)$ es la suma de todas las longitudes de los caminos más cortos que empiezan en i , es decir, si $d(i, j)$ es la longitud del camino más corto entre i y j , entonces $s(i) = \sum_j d(i, j)$. [4]

Como podemos notar, mientras más chico sea $s(i)$, más caminos cortos tendrá i , y a su vez, mayor centralidad de cercanía (pues estamos tomando el recíproco). Así pues, los individuos con alta centralidad de cercanía serán aquellos que tienen muy pocos intermediarios con los desconocidos que podrían conocer.

3.3. Redes de Barabási-Albert

A continuación se presenta un algoritmo para construir la red que se usará en el proyecto, la cual es del tipo Barabási-Albert. Esta sigue una construcción conocida como unión preferencial, la cual se basa en ir uniendo los nodos con aquellos de mayor grado. La construcción que daremos es la misma que se usó para generar la red del proyecto, aunque vale la pena tomar en cuenta que existen muchas variantes de esta construcción.

Definición 10 (Red tipo estrella) Una red tipo estrella es aquella que tiene un nodo central c que está conectado a todos los demás nodos, y ningún otro nodo aparte de c está conectado a otro que no sea el centro. Es decir, si tenemos n nodos, el nodo c tendrá grado $n - 1$, y todos los demás nodos tendrán grado 1 [4]. Un dibujo de este tipo de red se aprecia más abajo en el primer paso de la construcción de una red de Barabási-Albert.

Definición 11 (Construcción de red Barabási-Albert) La construcción hace uso de tres parámetros: un número inicial de nodos m_0 , un número de conexiones por nodo nuevo $m \leq m_0$, y un número total de nodos $n > m_0$. En el presente trabajo siempre se tomará $m_0 = m + 1$, pero no necesariamente debe ser así. Posteriormente, se sigue el siguiente algoritmo [10]:

1. Se comienza con una red inicial de m_0 nodos tipo estrella.
2. Se añade un nuevo nodo v a la red.
3. Se conecta v a m nodos existentes, donde la probabilidad de que v se conecte a un nodo i existente es de $\frac{\deg(i)}{\sum_j \deg(j)}$, es decir, proporcional al grado de i (j son los diferentes nodos existentes).

4. Se repiten los pasos 2 y 3 hasta alcanzar el número deseado de nodos.

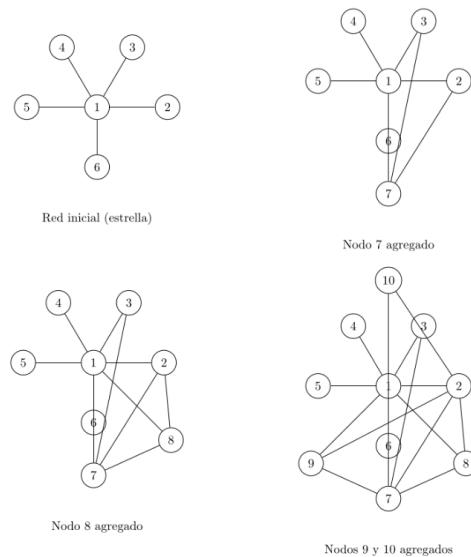


Figura 3. Evolución de la red Barabási-Albert. Conexiones preferenciales de 3 aristas por cada nodo agregado.

4. Metodología

El enfoque en este trabajo, es, sobre todo, **probabilístico**. Supongamos una población constante de N individuos representada por una red G , donde los nodos hacen las veces de individuos y los enlaces contactos entre ellos, y una enfermedad que la amenaza. Cada nodo, u , puede tener tres estados, a los que se les asignó una clave, S , I o R , a usar en adelante:

- **Susceptible (S):** El nodo puede contagiarse de la enfermedad. Al número de nodos S al tiempo t , en adelante, se le denotó por $S(t)$
- **Infectado (I):** El nodo ya ha sido contagiado, y está en proceso de recuperación. Al número de nodos I al tiempo t , en adelante, se le denotó por $I(t)$
- **Recuperación (R):** El nodo se ha recuperado totalmente y ha desarrollado inmunidad. Al número de nodos R al tiempo t , en adelante, se le denotó por $R(t)$

Notése que este modelo es idéntico al SIR descrito con anterioridad. Así, se supuso que la enfermedad afecta a la población y cambia su estado en cada lapso Δt : para ejemplificar este concepto, en la realidad esto puede significar un ajuste en las bases de datos de la Secretaría de Salud semanal, siendo Δt una semana.

Ahora bien, se propuso que por cada nodo I en la red, y por cada paso temporal, existe una **probabilidad de contagio a su vecino**, definida a través de un **parámetro de infección** $\beta > 0$, dada por:

$$p = 1 - e^{-\beta \Delta t} \quad (4)$$

y una **probabilidad de recuperación**, con **parámetro de recuperación** $\gamma > 0$, dada por:

$$q = 1 - e^{-\gamma \Delta t} \quad (5)$$

Es decir, que todos los vecinos S de cada nodo I en la red tendrán una probabilidad p de sufrir un contagio, y cada nodo I una q de recuperarse. Obsérvese que, si $\beta = 0$, $p = 0$, es decir, que la enfermedad nunca se transmitirá, y si $\beta \rightarrow \infty$, $p \rightarrow 1$, esto es, que la enfermedad no se puede evitar; y el caso para q es análogo.

El lector seguramente se estará preguntando por qué hemos de definir así estas tasas, a lo cual se le puede dar una respuesta simple.

En primer lugar, como se definieron, para cualquier $\beta > 0$ y $\gamma > 0$, se tiene que $p, q \in [0, 1]$ aprovechando que la exponencial propuesta en $[0, \infty)$ está acotada por 1 y es positiva, por lo que en efecto hablamos de probabilidades, al menos en magnitudes, pero además **se deben recuperar las ecuaciones SIR de este modelo en una población bien mezclada y en el caso límite $\Delta t \rightarrow 0$** . En dicha situación las ecuaciones 4 y 5 se reducirían, aproximando a primer orden, a:

$$p = 1 - e^{-\beta \Delta t} \approx 1 - (1 - \beta \Delta t) = \beta \Delta t$$

y:

$$q = 1 - e^{-\gamma \Delta t} \approx 1 - (1 - \gamma \Delta t) = \gamma \Delta t$$

Con estas probabilidades, es posible calcular los valores de expectación para cambios de población, esto es, los valores esperados para los cambios en el número de nodos en el tiempo $t + \Delta t$ para cada uno de los compartimentos S, I o R dado su estado en el tiempo t . En el caso de los nodos R :

$$\langle \Delta R \rangle = qI(t) \approx \gamma I(t)\Delta t \iff \frac{\langle \Delta R \rangle}{\Delta t} = \gamma I(t) \quad (6)$$

Pues cada nodo I tiene una probabilidad q de recuperarse, es decir, de aumentar la población en R . Para los nodos S :

$$\langle \Delta S \rangle = -p(S(t)I(t)) \approx -\beta I(t)S(t)\Delta t \iff \frac{\langle \Delta S \rangle}{\Delta t} = -\beta I(t)S(t) \quad (7)$$

El producto $S(t)I(t)$ viene de que por cada nodo I **en una población bien mezclada** se puede contagiar a toda la población de nodos S , por lo que el número de contactos $I - S$ se da por $S(t)I(t)$. Al multiplicar por la probabilidad de contagio p , se tiene la tasa de cambio esperada para ΔS , pero **también se multiplica por un menos** dado que al infectarse, la población **sale** del compartimento S , a diferencia del caso anterior donde **entra** al R . Finalmente:

$$\frac{\langle \Delta I \rangle}{\Delta t} = -\left(\frac{\langle \Delta R \rangle}{\Delta t} + \frac{\langle \Delta S \rangle}{\Delta t} \right) = \beta I(t)S(t) - \gamma I(t) \quad (8)$$

Simplemente porque la población total **es constante** con lo cual $\langle \Delta R \rangle + \langle \Delta S \rangle + \langle \Delta I \rangle = 0$. Y nuevamente tomando el límite $\Delta t \rightarrow 0$, las ecuaciones 6, 7 y 8 se transforman en **derivadas temporales**:

$$\begin{cases} \frac{dR}{dt} = \gamma I \\ \frac{dS}{dt} = -\beta IS \\ \frac{dI}{dt} = \beta IS - \gamma I \end{cases}$$

Que el lector podrá notar **es un sistema de ecuaciones diferenciales que corresponde al modelo SIR clásico** (salvo una $\beta = N\tilde{\beta}$ para ser idéntico al mismo, lo cual no afectaría al modelo en cuestión, pues N es constante). Es por ello que las expresiones probabilísticas propuestas en este trabajo (ecuaciones 4 y 5) son consistentes con el modelo SIR, específicamente con su fenomenología.

Así pues, la metodología es evidente: se generó una red **aleatoria** y se infectó a 1 o más nodos dentro de la misma, y a los demás se les dejó susceptibles a la enfermedad. Se inició un **ciclo de simulación** para que la red evolucione en pasos temporales Δt , es decir, que el tiempo t se parametrizó como $t = n\Delta t$, con $n = 0$ el paso, o estado, inicial, y siempre se consideró a $\Delta t = 1$.

En cada paso temporal, se iteró en cada vecino de cada nodo infectado, cada cual tendrá una probabilidad p de infección para el siguiente paso temporal (en código, se generó un número aleatorio r entre 0 y 1, y si $r \leq p$, el nodo se infectará en el paso siguiente al que se esté iterando, donde en dicho paso mantendrá su estado S dado que este proceso sólo incidirá en el tiempo subsecuente). Posteriormente, por cada nodo infectado se generó otro número aleatorio $s \in [0, 1]$, de tal manera que si $s \leq q$, dicho nodo se “recuperó” (de nuevo, esto sólo afectará **en el siguiente paso temporal, en el que**

se esté iterando, todo quedará igual). Al final, se hizo un recuento de $S(n\Delta t), R(n\Delta t)$ y, principalmente, de $I(n\Delta t)$, y con ello, se generó una evolución temporal de los individuos infectados, susceptibles y recuperados, **que es precisamente el objetivo del trabajo**. La simulación **terminó una vez que no existió ninguna persona infectada**.

Quedan aún algunos cabos sueltos. La simulación tiene como parámetros a β y a γ , esto es, los parámetros de infección y recuperación, los cuales serán ajustados conforme a la situación y se le hizo saber al lector cuáles son usados en determinado momento. Aún así, resta definir:

¿Qué red aleatoria se utilizó?: Se utilizó una red aleatoria de Barabási-Albert, dado que, como se mencionó con anterioridad, es una red que facilita la generación de **hubs** y es de libre escala, lo cual con certeza es un comportamiento propio de la sociedad vista como un todo. Esta red permitió generar nodos (individuos) con muchos y pocos contactos entre sí y es particularmente útil para modelar interacciones a larga distancia, con lo cual para modelar pandemias esto es significativamente aplicable, pues en la vida real por más débil que una enfermedad sea, si afecta a un individuo muy conectado (incluso es posible pensar que los nodos de la red sean centrales como aeropuertos o estaciones), se esparrería muy fácilmente.

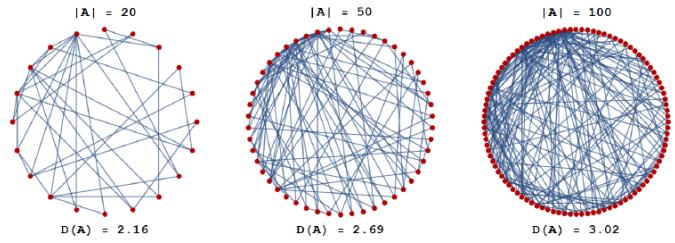


Figura 4. Ejemplo de una red de Barabási-Albert. Los parámetros N y m se variaron a lo largo del trabajo.

y claro :

¿Qué nodos fueron los primeros en infectarse?: Aquí entra la novedad del estudio: los primeros nodos infectados, en adelante **pacientes cero**, variaron de acuerdo a la **centralidad a estudiar**. Se eligió primero una centralidad que se quisiera analizar en una pandemia (como la centralidad de grado para determinar qué personas tienen más o menos conexiones con otras) y se infectó inicialmente a aquellos individuos (un número que se le informó al lector oportunamente) con mayor y menor centralidad, para obtener el desarrollo de la epidemia a lo largo del tiempo.

Por último, es conveniente definir una cantidad conocida como **infecciones cumulativas**. En pocas palabras, no es más que la suma de infecciones que han ocurrido desde el inicio de la pandemia hasta el tiempo t . Dado que el modelo propuesto es estocástico, conviene en primera instancia encontrar un **promedio de infecciones al tiempo t** al hacer N simulaciones. Esto se expresa como:

$$\mu_t = \frac{1}{N} \sum_{n=1}^N I_n(t) \quad (9)$$

Y con ello, las infecciones cumulativas al tiempo $t = n\Delta t$ se darán por:

$$C(t) = \sum_{k=0}^n \mu_{k\Delta t} \quad (10)$$

Esto por supuesto no refleja cuántos infectados hay por el agente patógeno en determinado tiempo, simplemente porque una infección que, por ejemplo, infecte a una sola persona por 10 unidades de tiempo tendrá más infecciones cumulativas que una infección que afecte a 9 personas sólo por una unidad de tiempo, con lo cual esta cantidad no refleja bien ese concepto. Lo que sí hace es **medir**

cuánto tiempo duró la pandemia pues una infección con mayores infecciones cumulativas significa que tuvo cierta cantidad de infectados, por pequeña que sea, por varias unidades de tiempo, esto es, que aquejó a la población por un lapso significativo.

Una vez tenemos todas estas herramientas definidas es prudente entonces mencionar específicamente qué se hizo: en primer lugar, se realizó una simulación con parámetros $N = 1000$, $m = 2$, $\beta = 0.3$, $\gamma = 0.1$ infectando a un solo nodo con la mayor centralidad de grado para mostrar que el modelo propuesto en efecto es compatible con la versión clásica del SIR, y una vez hecho eso, se llevó a cabo la misma simulación, aunque ahora con parámetros $N = 50$, $m = 1$, $\beta = 0.3$, $\gamma = 0.1$, para mostrar explícitamente los estadios de la red para diferentes tiempos a lo largo de la pandemia, y finalmente se generaron simulaciones con diferentes parámetros (N , m , β y γ) para observar fenomenológicamente su influencia en la red.

Una vez hecho esto, se realizaron simulaciones infectando a 4 nodos dependiendo de su centralidad, en una red de 50 nodos con parámetros de infección $\beta = 0.1$ y $\gamma = 0.1$ (de forma que contagio y recuperación fueran igualmente probables), así como $m = 2$ para observar la propagación de la enfermedad en el estadio inicial y final de la pandemia, y las diferencias en ellos entre diferentes métricas. Se eligieron 6 métricas en particular: grado, betweenness, cercanía, eigenvector, katz y carga, todas con los parámetros predefinidos de networkx (dado que no se privilegió a ningún nodo sobre otro al calcular las métricas, más bien, se usaron las métricas para determinar si serían pacientes cero o no). Como se mencionó, los 4 pacientes cero se eligieron de acuerdo a los nodos con mayor y menor centralidad en cada caso (aunque no los primeros 4 ni los últimos 4, siempre hubo un “shift” de 10 para que no se eligieran a los mismos nodos en cada caso, pues en la red utilizada, y en la sociedad en general, los nodos con mayor centralidad en una métrica son usualmente los mismos en otra).

Por último, y para sustentar lo observado en el apartado anterior, para cada una de las métricas, y eligiendo 3 pacientes cero con mayor y menor centralidad, se realizaron 100 simulaciones con parámetros $N = 1000$, $m = 2$, $\beta = 0.3$ y $\gamma = 0.1$ para hacer un promedio de las curvas de infección a lo largo del tiempo, así como un promedio de las curvas de infecciones cumulativas; con el objetivo de observar la variación en las mismas dependiendo de los pacientes cero para cada métrica, tanto en cantidad de contagios, como en el tiempo del pico de la epidemia y las diferencias en las infecciones cumulativas entre ambos casos.

En cuestión del código, se utilizaron las librerías **networkx**, **matplotlib**, **numpy** y **random**, para trabajar con las redes, graficarlas, obtener datos de ellas y operarlos, y generar los números aleatorios respectivamente. El código se puede obtener de este enlace: [Código](#).

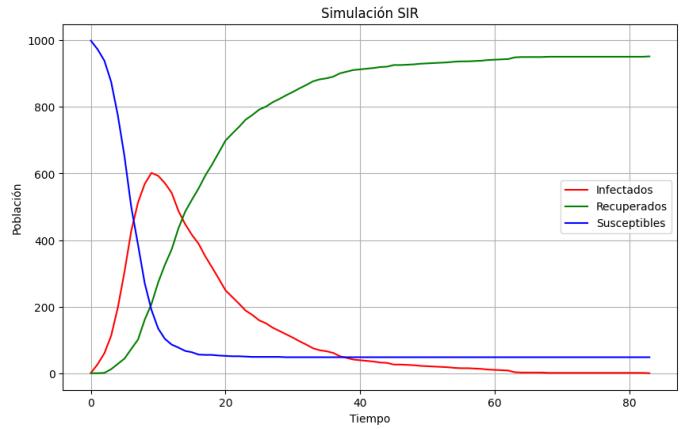


Figura 5. Simulación SIR para una red poco conectada

Esta gráfica es, fenomenológicamente, **idéntica** a lo que se observaría para el modelo SIR clásico en una población bien mezclada:

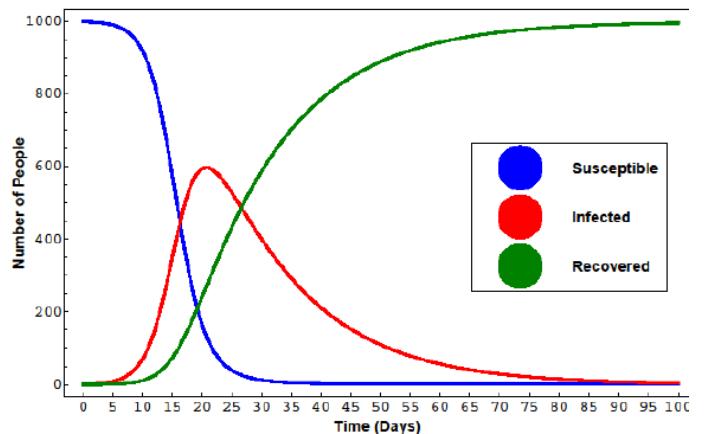


Figura 6. Simulación SIR en una población bien mezclada. Recuperado de [9]

5. Resultados y Discusión de Resultados

5.1. Fenomenología General del Modelo

Claramente el primer resultado razonable que es prudente mostrar es que el modelo propuesto en efecto funciona. Utilizando una red de Barabási-Albert con $N = 1000$ y $m = 2$, eligiendo a $\beta = 0.3$, $\gamma = 0.1$ como los parámetros de infección junto con el paciente cero siendo el nodo con mayor centralidad de grado, y usando un paso de tiempo de 1 (que puede simular días, semanas, meses, o cualquier otro lapso razonable), la pandemia tiene la siguiente forma:

Claro, con menos suavidad en las curvas al ser una simulación en una gráfica, con lo que la cantidad de infectados en cada tiempo sólo puede ser un número entero. Antes de continuar con el impacto de las métricas en el desarrollo de la pandemia, es prudente notar, gráficamente, el desarrollo de la misma a lo largo de la red. eligiendo una red de 50 nodos y $m = 1$ (para que se formen “grupos” visualmente distinguibles en la red), con los mismos parámetros de infección que en la simulación pasada, se obtiene la siguiente evolución:

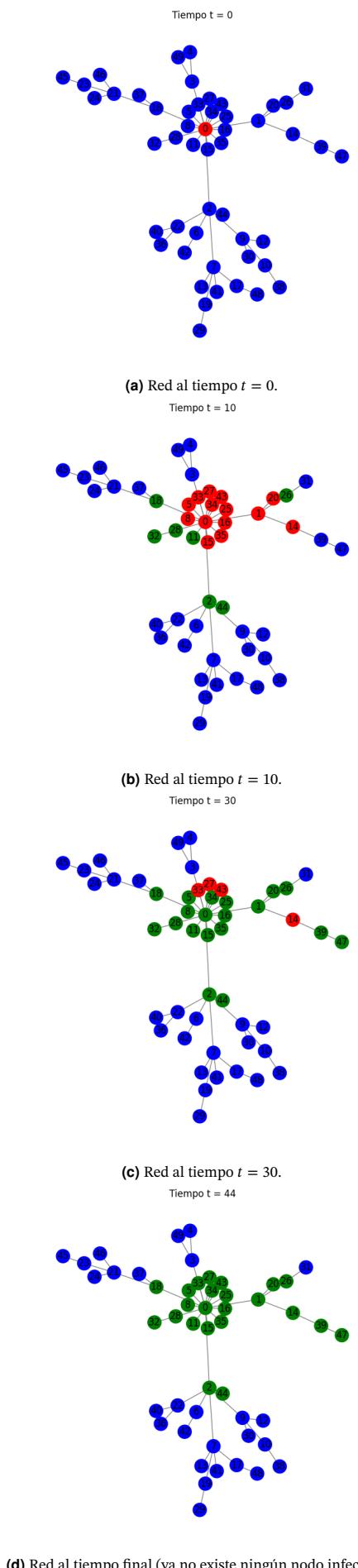


Figura 7. Evolución de la Población a lo largo del tiempo. Nótese que en el modelo presentado la epidemia en efecto se puede localizar en un grupo de la población en particular y no esparcirse en toda la sociedad.

Nuevamente el paciente 0 fue quien tenía mayor centralidad de grado de la red, pero nótese que en esta simulación la infección se limitó al grupo cercano del paciente 0, y contados miembros de otros grupos, lo cual tiene sentido: se infectó a un individuo que localmente está bien conectado, pero no necesariamente quiere decir que lo esté con toda la red (y es lo que se observó en México a lo largo de la pandemia de la COVID, vea la figura inferior), pero en general, esta es otra prueba de que el modelo generalizado a redes del clásico SIR cumple con lo que intuitivamente se espera de una epidemia, y más aún, se aborda un problema que el mismo tenía y que fue discutido en la introducción: **ahora es posible visualizar gráficamente la transmisión de un agente patógeno a lo largo de una sociedad, y observar a qué grupos les afecta en mayor medida.**

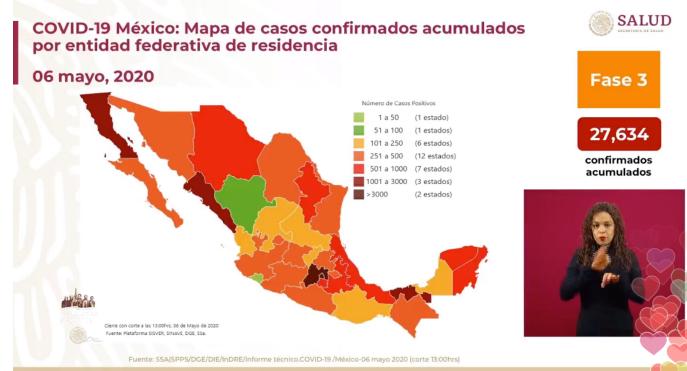


Figura 8. Población Infectada por estado en México en Mayo de 2020, nótese que la pandemia no se distribuye uniformemente por estado, sino que se concentra en unas regiones y se diluye en otras. Recuperado de [12].

Casos confirmados por alcaldía

Total en la CDMX: 1,266

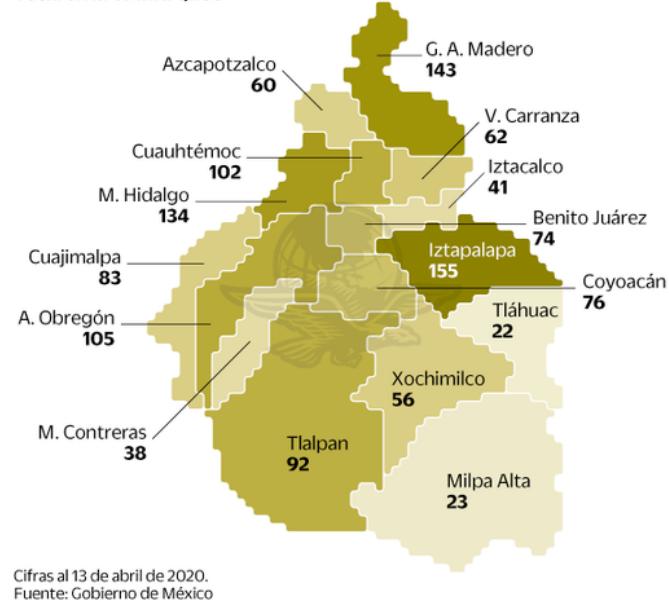


Figura 9. Población Infectada por alcaldía en México en Abril de 2020, nótese que la pandemia no se distribuye uniformemente por alcaldía, sino que se concentra en unas regiones y se diluye en otras. Recuperado de [12].

El último tema a abordar en esta sección es la influencia que tienen los parámetros m , β y γ en la simulación (realmente N no tiene influencia, pues es sólo la cantidad de individuos en la población, por lo que se fijará en 1000, y el paciente cero nuevamente es el nodo con mayor centralidad de grado). En cada gráfica **se indican los parámetros utilizados, pero en el pie de foto de cada una se resaltará**

el parámetro cuyo comportamiento se desea evidenciar.

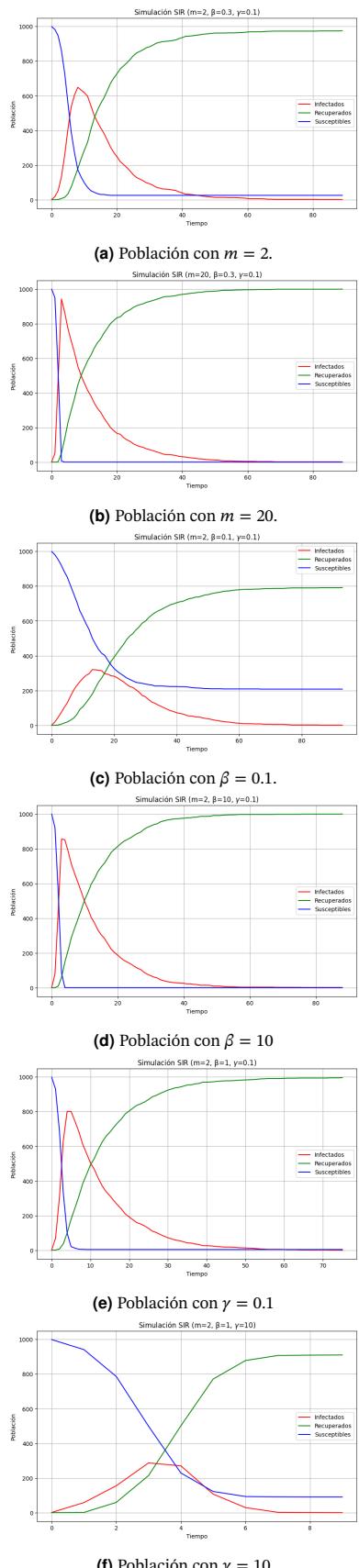


Figura 10. Evolución de la cantidad de infectados, recuperados y susceptibles a lo largo del tiempo con distintos parámetros de infección y de generación de enlaces.

En el primer y segundo renglones se observan, respectivamente,

una sociedad con pocos enlaces, y otra con muchos, siendo afectada por **exactamente el mismo agente** (es decir que β y γ son **invariantes**). Incluso si sólo hay un solo paciente cero, que está localmente bien conectado con la sociedad, la interrelación que existe en la misma claramente afecta al desarrollo de la epidemia, haciendo que mientras más contacto haya entre los individuos, más sencilla es la transmisión del agente patógeno, lo cual tiene sentido, y corrobora que la práctica de **confinamiento reduce significativamente los contagios a lo largo del tiempo, que es la razón por la que la Secretaría de Salud en 2020 recomendó dicha medida como el principal mecanismo de mitigación en la pandemia de Covid-19** [12].

Ahora bien, en los siguientes dos renglones se tiene la misma población (m es invariante) y se altera como tal a la infección (β). Obsérvese que de la ecuación 4, la probabilidad de infección es menor mientras β sea más cercana a 0, y viceversa. De ahí que en la primera gráfica, la cantidad de contagios en función del tiempo es mucho menor con $\beta = 0.1$ (la asombrosa cantidad de 200 personas salieron indemnes de la pandemia) y que en la de abajo prácticamente la totalidad de la población se contagió con $\beta = 10$. Esto se reduce a que el parámetro β en efecto modela apropiadamente a la infectividad del patógeno.

Por último, en el caso de la variación del parámetro de recuperación, obsérvese que para $\gamma = 0.1$, el pico de infección es bastante alto, llegando incluso a haber 800 personas infectadas al mismo tiempo, mientras que para $\gamma = 10$ la curva de infección se aplana significativamente, teniendo un total de 100 personas aún susceptibles al final de la infección. Este parámetro entonces también cuantifica apropiadamente la recuperación de la población, basados en la ecuación 5, que tiene el mismo comportamiento que la 4, y engloba tanto al sistema inmune de la misma, como los tratamientos hospitalarios que pueda haber en la sociedad.

5.2. Simulaciones variando Centralidades

Como se mencionó anteriormente, el objetivo principal de este trabajo, además de generalizar el método SIR al caso de redes complejas, es el de poder ver como influencia la elección de pacientes cero de acuerdo a sus centralidades. Esto permitiría, por ejemplo, ver qué impacto tiene la posición social de una persona en el desarrollo de la pandemia, lo cual sería radicalmente útil para ofrecer estrategias de prevención personalizadas para cada tipo de individuo.

Pragmáticamente, lo más sencillo sería elegir pacientes cero simplemente infectando cierta cantidad de nodos con la **mayor centralidad que se esté analizando en ese momento**, pero resulta poco provechoso, dado que generalmente, y especialmente con el tipo de red con la que trabajamos (y en la sociedad en general) los nodos más altos de una centralidad usualmente serán también los más altos en otra, por lo que no reflejaría adecuadamente el comportamiento de la red (no quiere decir que sea equivocado, sólo que prácticamente sería la misma simulación). Así, en lugar de hacerlo con los primeros nodos en términos de centralidad, podemos hacerlo con los **primeros nodos a partir del décimo**, simplemente porque en ese rango ya todas las centralidades tendrán un ranking diferente. Análogamente, también es posible elegir a los pacientes cero con base en los nodos **con menor centralidad**, también con ese "desfase", aunque ahora de abajo hacia arriba. Eligiendo como parámetros $\beta = 0.1$ y $\gamma = 0.1$ (para mejor visualización del desarrollo de la pandemia) en una red de 50 nodos con $m = 2$, y escogiendo a los cuatro nodos (con desfase) con mayor y menor centralidad para ser los pacientes cero, obtendríamos los siguientes estados iniciales y finales:

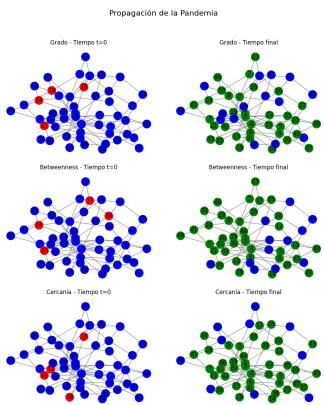


Figura 11. Simulación de Propagación para los nodos con mayor centralidad de grado, betweenness y cercanía.

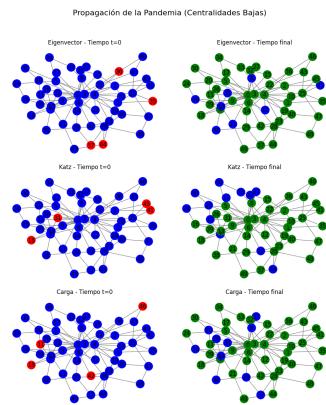


Figura 14. Simulación de Propagación para los nodos con menor centralidad de Eigenvector, Katz y Carga .

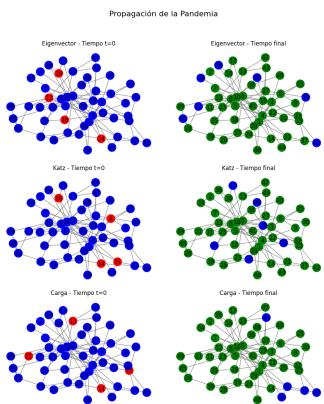


Figura 12. Simulación de Propagación para los nodos con mayor centralidad de Eigenvector, Katz y Carga .

Así como:

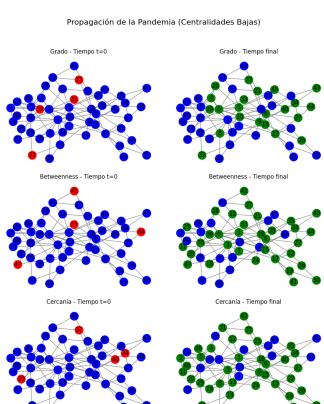


Figura 13. Simulación de Propagación para los nodos con menor centralidad de grado, betweenness y cercanía.

El análisis de estas gráficas resulta interesante. Ordenando las simulaciones según el nivel de afectación total de la población —medido por la cantidad de nodos verdes al final de la simulación, tras la erradicación del agente patógeno—, se observa que las centralidades que generaron mayor propagación fueron, en orden descendente y considerando a los pacientes cero como aquellos **con la mayor centralidad**: la centralidad de carga, de Katz, de eigenvector, de cercanía, de betweenness y, finalmente, la centralidad de grado. Este orden, particularmente en el tipo de red que fue elegida para modelar una sociedad, tiene muchísimo sentido, en el hecho de que los privilegios sociales que tiene cierto nodo, que se reflejan matemáticamente en “qué tan bien conectado esté en la red” tengan un impacto directo sobre los procesos de propagación que se desarrollan en la red al acontecer una pandemia: la **centralidad de carga**, como se mencionó, considera el flujo de información a través de todos los caminos posibles entre pares de nodos, no restringiéndose únicamente a los caminos más cortos. En redes libres de escala, esta centralidad tiende a identificar nodos ubicados estratégicamente entre comunidades o subestructuras densas, lo que los convierte en puntos óptimos para maximizar la cobertura de la propagación. La infección iniciada desde estos nodos tiene alta probabilidad de acceder a múltiples rutas simultáneamente, alcanzando regiones alejadas de manera eficiente.

Las **centralidades de Katz y de eigenvector**, por su parte, incorporan no solo la conectividad directa de un nodo, sino también la importancia de sus vecinos, lo cual resulta particularmente adecuado para modelar una sociedad humana: al fin y al cabo, si una persona, sin importar que tan aislada esté, conoce y contagia a otra muy conectada en la red, la pandemia seguro alcanzará proporciones mayúsculas. En otras palabras, estas métricas permiten identificar nodos que, aunque no sean los de mayor grado, están situados en posiciones de influencia estructural, capaces de desencadenar procesos de propagación de gran alcance a través de su red de vecinos altamente conectados.

La **centralidad de cercanía**, claro, mide la distancia promedio de un nodo respecto a todos los demás en la red. Si bien puede resultar efectiva en redes homogéneas, su utilidad en redes libres de escala, como la que se usa en este trabajo y, se cree, modelan adecuadamente a una sociedad, es menor, pues a pesar de que los nodos con alta centralidad de cercanía suelen coincidir con los grandes hubs, que están conectados a muchas regiones de la red a través de caminos cortos, debido a que la red está fuertemente centralizada alrededor de estos pocos nodos, la mayoría de los demás presentan una centralidad de cercanía baja y valores poco diferenciados entre sí. Como resultado, al seleccionar como paciente cero un nodo con alta centralidad de cercanía, se escoge típicamente un nodo bien posicionado, pero **no necesariamente con la mejor capacidad para desencadenar una propagación extensa**, ya que esta métrica no considera la densidad de conexiones directas ni la importancia estructural de los vecinos.

Por otro lado, la **centralidad de betweenness**, aunque relevante

para identificar nodos que median en la conexión entre diferentes regiones de la red, puede sobreestimar la capacidad de dichos nodos para generar contagios, dado que su valor está más relacionado con la función estructural de intermediación que con qué tanto realmente pueden influenciar a la red en una escala global. En consecuencia, su efectividad como estrategia de iniciación de la propagación resulta moderada.

Por último, la **centralidad de grado** resulta ser la menos efectiva en este escenario, puesto que, como se mencionó previamente, es una centralidad **local**, esto es, que un alto grado no garantiza una posición estructural favorable para la propagación global, ya que es posible que un nodo altamente conectado localmente se ubique en una región periférica o aislada de la red, con acceso limitado a otras comunidades relevantes.

En resumen, en una pandemia son los individuos con mayor centralidad de Katz y de Carga los que tienen la capacidad de iniciar procesos de propagación extensivos, en comparación con aquellas que se basan únicamente en propiedades locales o de intermediación puntual. Es por ello que medidas como el aislamiento o el cuidado de poblaciones de riesgo resultan tan eficientes en la prevención de una enfermedad.

En el caso de elegir a los pacientes cero con la **menor centralidad** aún así se observó la misma tendencia, lo cual puede parecer contraintuitivo. Sin embargo, es razonable que el orden relativo de las centralidades se mantenga, incluso al seleccionar los nodos con menores valores, porque las métricas no solo identifican a los nodos más influyentes, sino que también reflejan —aunque de forma menos precisa— la **estructura global** de la red en sus valores bajos. Esto implica que un nodo con baja centralidad de Katz puede seguir estando estructuralmente mejor posicionado para propagar que uno con bajo grado, preservando así el orden observado. No hay una simetría estructural que justifique un orden diferente, e incluso invertido (que es lo que intuitivamente se esperaría, dado que se creería el comportamiento opuesto a elegir a las mejores centralidades como pacientes cero), ya que las funciones de las métricas no son linealmente opuestas entre sí. En otras palabras, **en el caso de las peores centralidades conviene más hacer un análisis comparativo con esa misma métrica en el caso de las mejores centralidades, más que entre diferentes métricas**. Ese, de hecho, es el último análisis del presente escrito.

5.3. Gráficas Variando Centralidad

Si bien, las visualizaciones anteriores permiten hacerse una idea del comportamiento general del modelo basado en las centralidades antes discutidas, resulta imperativo hacer un análisis numérico apropiado del transcurso de la infección en función del tiempo, especialmente en cuestión de la comparación entre alta centralidad y baja centralidad de los pacientes cero para cada métrica en la misma red. Con ese objetivo, es posible hacer una cantidad alta de simulaciones, que en este trabajo se limitará a 100, y promediar sus resultados tal y como se discutió en el marco teórico, **para obtener un promedio de infecciones al tiempo t** . Adicionalmente, es posible encontrar **las infecciones cumulativas** al tiempo t , esto es, la suma de todas las infecciones que ha habido desde el inicio de la pandemia hasta dicho tiempo.

Dado que este análisis es más numérico, es posible utilizar parámetros más realistas para las simulaciones, a diferencia de la sección anterior donde se necesitaron parámetros pequeños para facilitar la visualización. Eligiendo una población de $N = 1000$ personas, con $m = 2$ en la red, y parámetros $\beta = 0.3$ y $\gamma = 0.1$, y nuevamente eligiendo los 3 nodos con mayor y menor centralidad en cada caso (con el desfase discutido previamente), se obtienen las siguientes figuras, donde el área sombreada es la desviación estándar de las simulaciones:

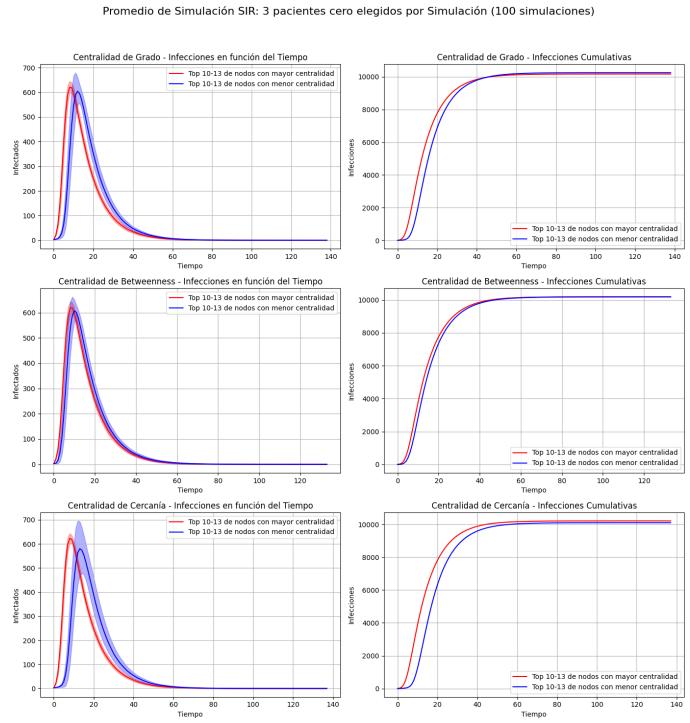


Figura 15. Infecciones Promedio y Cumulativas para centralidad de Grado, Betweenness y Cercanía, con pacientes cero definidos por los nodos con mayor y menor centralidad en la red.

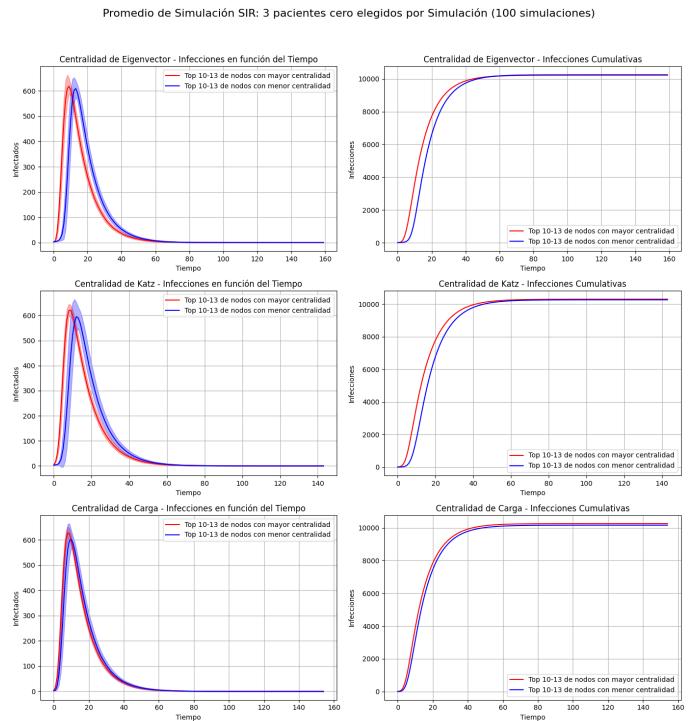


Figura 16. Infecciones Promedio y Cumulativas para centralidad de Eigenvector, Katz y Carga, con pacientes cero definidos por los nodos con mayor y menor centralidad en la red.

En el caso de la **centralidad de grado**, cuando los pacientes cero presentan una alta centralidad, la epidemia se propaga de forma más acelerada, alcanzando su pico de infecciones aproximadamente 5 unidades de tiempo antes que en el caso de baja centralidad. La magnitud máxima del brote es ligeramente mayor para los nodos con alto grado, lo que indica una mayor eficiencia en la transmisión inicial. No obstante, las infecciones acumuladas finales son prácticamente

idénticas en ambos casos, lo que sugiere que **la elección de los nodos iniciales influye principalmente en la dinámica temporal y no en el tamaño total del brote, y ni siquiera eso, puesto que la diferencia temporal es prácticamente despreciable**. Esto tiene sentido, puesto que la centralidad de grado, a pesar de su sencillez conceptual, es de hecho una métrica bastante inefficiente para analizar comportamientos globales de la red.

En el caso de la **centralidad de Betweenness**, el inicio de la infección en nodos con alta centralidad de intermediación también conduce a una propagación más rápida, con un pico de infecciones más temprano y una curva más simétrica en comparación con los nodos de baja betweenness. Sin embargo, las diferencias en la altura del pico, tanto en número de infecciones como en desfases temporales, son reducidas, y las curvas acumulativas convergen hacia un mismo número final de infectados. Esto indica que **a pesar de que en primera instancia la función de puente de dichos nodos debería acelerar la difusión entre comunidades sin modificar de forma sustancial la extensión total del contagio, aún así no representa una diferencia radical el elegir a pacientes cero con mayor o menor betweenness**, nuevamente regresando al caso antes discutido: esta métrica tiene más bien una función estructural de intermediación, a diferencia de una que cuantifique qué tanto realmente pueden influenciar a la red en una escala global.

Cuando hablamos de la **centralidad de cercanía**, al infectar a los nodos con mayor centralidad se tuvo un crecimiento inicial que reflejó una rápida diseminación global de la infección. Además, el pico ocurre significativamente antes y con una intensidad ligeramente superior respecto al escenario de baja cercanía, **aunque el número acumulado final de infecciones vuelve a ser casi equivalente** (aunque siempre teniendo más infecciones cumulativas en el caso de la curva con pacientes cero de mayor centralidad). Este resultado respalda la idea de que **la cercanía favorece la sincronización del contagio en toda la red, acelerando la fase de expansión sin alterar el desenlace final**.

Para la **centralidad de eigenvector**, la curva de infecciones para mayor centralidad alcanza su pico en el mismo nivel máximo que en el caso de baja centralidad, lo que indica que la magnitud del brote no depende, en una primera aproximación, de esta medida. Sin embargo, el pico ocurre de manera significativamente más temprana para los nodos de alta centralidad, mientras que en los de baja centralidad se observa un desplazamiento temporal hacia la derecha, reflejando un inicio y propagación más lentos. Las curvas acumulativas son prácticamente idénticas, alcanzando el mismo número final de infecciones, aunque en este caso la baja centralidad sobrepasa a la alta **que significa que la pandemia tiene mayor duración para pacientes cero con menor centralidad de eigenvector**, lo que confirma que la centralidad de eigenvector influye principalmente en la **velocidad** de propagación y no en la **extensión total** del brote. En conjunto, estos resultados sugieren que la estructura jerárquica de influencia capturada por esta medida modula el ritmo de difusión epidémica sin alterar su alcance final, que tiene sentido, pues la centralidad de eigenvector mide “que tan bien conectados están los vecinos de un nodo”, y tener un paciente cero con una centralidad de eigenvector alta o baja significa que alcanzará a un núcleo o hub de la red rápida o lentamente, respectivamente. Sin embargo, debido a que ese núcleo está fuertemente conectado, **la epidemia alcanza prácticamente a los mismos nodos finales** que cuando empieza en regiones periféricas, que es lo que se obtuvo. En resumen, esta centralidad modela la **velocidad de propagación de la epidemia: pacientes cero con alta centralidad generan una epidemia rápida, mientras que los de baja tardan más en alcanzar un pico de infección, pero a cambio la epidemia dura más**.

El comportamiento epidémico asociado a la **centralidad de Katz** es en gran medida análogo al observado para la centralidad de cercanía, aunque con diferencias sutiles en la magnitud de los brotes. Cuando los pacientes cero presentan alta centralidad de Katz, la in-

fección se propaga con mayor rapidez, alcanzando un **pico más alto y más temprano** que en el caso de baja centralidad. Además, el número total de infecciones acumuladas es **ligeramente mayor** en el escenario de alta centralidad, lo que sugiere una difusión más eficiente a través de la red.

Estos resultados reflejan que la centralidad de Katz, al ponderar las conexiones tanto directas como indirectas, refleja una noción de “influencia global” similar a la de la cercanía, pero más sensible a la estructura jerárquica de la red. En consecuencia, los nodos con alta centralidad de Katz no solo aceleran la propagación inicial, sino que también logran una penetración marginalmente más amplia en la misma, lo que los convierte en puntos particularmente efectivos para la diseminación epidémica, que es algo que ya se había observado desde la sección anterior.

Por último, en el caso de la **centralidad de carga**, las simulaciones muestran el **mayor número total de infecciones** entre todas las métricas analizadas, lo que indica una difusión especialmente extensa del contagio, lo cual se observó desde las simulaciones de la sección anterior. Sin embargo, las curvas correspondientes a nodos de **alta y baja centralidad de carga** son **prácticamente indistinguibles**: ambas presentan picos de igual altura, ocurren en tiempos similares y alcanzan valores acumulativos casi idénticos.

Este resultado sugiere que, aunque la estructura global de la red favorece una propagación más amplia del brote, la **posición inicial de los pacientes cero según la centralidad de carga tiene un impacto mínimo** en la dinámica epidémica. En otras palabras, la centralidad de carga parece capturar propiedades del flujo global de la red que determinan su susceptibilidad general a la infección, pero **no diferencia de forma efectiva entre nodos en cuanto a su capacidad individual de iniciar o acelerar la propagación**. Este caso, en pocas palabras, es el opuesto al de Eigenvector.

6. Conclusiones

Finalmente, se retoma la problemática planteada al inicio de este trabajo: la limitante del modelo SIR clásico al asumir una población bien mezclada, ignorando así la estructura inherente de la sociedad. No obstante, los resultados que se obtuvieron en el presente escrito demuestran que, pese a esta simplificación, el modelo SIR clásico mantiene su utilidad para estimar las cifras generales y el comportamiento macroscópico de una pandemia, y si bien el modelo de red a gran escala es ligeramente más preciso, para un análisis cualitativo ofrecerían resultados similares.

Sin embargo, se observó una diferencia crucial entre el modelo analítico y la implementación que aquí se ofrece para redes complejas: el primero tiende, claro, a sobreestimar el alcance de la infección, afectando invariablemente a la casi totalidad de la población bajo ciertos parámetros, mientras que la simulación en red permite escenarios más realistas donde grupos de individuos permanecen susceptibles, pero nunca son infectados, tal y como sucedió con la COVID-19 en México [12]. Esto sugiere que la incorporación de la topología de red ofrece una modelación más fidedigna al considerar el impacto de la organización social. Así, este trabajo logra integrar la robustez matemática del marco SIR con el realismo estructural de la teoría de redes.

Respecto a la fenomenología observada, es prudente destacar que los contagios pueden permanecer confinados en comunidades locales, incluso sin la intervención de medidas externas. Por otro lado, el análisis de las métricas de centralidad arrojó un resultado bastante contraintuitivo: a pesar de que la realización de análisis variando la centralidad del paciente cero presenta algunas diferencias, realmente esto tiene una influencia mínima sobre la magnitud total de la epidemia. Esto indica que la topología global de la red predomina sobre las características individuales del nodo inicial, comportamiento que se verificó tanto visual como analíticamente, observando que las curvas de infección promedio y cumulativa para pacientes cero con baja centralidad resultaron casi indistinguibles de aquellas con alta

centralidad, salvo en el tiempo para el pico de infecciones.

En consecuencia, el desarrollo final de una pandemia depende en mayor medida de las propiedades del agente patógeno y de la conectividad de la red que de la identidad o posición social del portador inicial. Este hallazgo reafirma la importancia crítica de las estrategias de mitigación, como el distanciamiento social, las cuales modifican la estructura de la red para frenar la propagación.

A la luz de los resultados que se arrojaron del presente estudio, se pueden identificar áreas de oportunidad para refinar el modelo en investigaciones futuras. En primer lugar, resultaría prudente contrastar las simulaciones con datos epidemiológicos reales (aplicando machine learning, por ejemplo) para validar cuantitativamente si la reducción de infectados aplicando los modelos y las estrategias de mitigación que se propusieron y observaron en el modelo de redes se ajusta a la realidad, y claro, ver los mejores ajustes para los parámetros de infección β y recuperación γ . Asimismo, sería pertinente extender el análisis a otras topologías de red (como redes de mundo pequeño o aleatorias) para verificar, o desmentir, la universalidad de las conclusiones respecto a las métricas de centralidad. Finalmente, una mejora sustancial consistiría en incorporar redes temporales, donde los enlaces entre individuos puedan crearse o romperse dinámicamente durante la simulación, otorgando un grado superior de realismo a la dinámica social frente a una emergencia sanitaria, que podrían hacer las veces tanto de variabilidad inherente a las relaciones humanas, como los cambios acaecidos por mero efecto de una pandemia.

■ Referencias

- [1] «The Milroy Lectures ON EPIDEMIC DISEASE IN ENGLAND—THE EVIDENCE OF VARIABILITY AND OF PERSISTENCY OF TYPE,» *The Lancet*, vol. 358, n.º 9294, págs. 1485-1490, 2001. DOI: 10.1016/S0140-6736(01)80187-2. dirección: [https://doi.org/10.1016/S0140-6736\(01\)80187-2](https://doi.org/10.1016/S0140-6736(01)80187-2).
- [2] A. Peña y O. Paco, «El concepto general de enfermedad. Revisión, crítica y propuesta. Segunda parte: Carencias y defectos en los intentos por lograr una definición general de enfermedad,» *Anales de la Facultad de Medicina, Universidad Nacional Mayor de San Marcos*, vol. 63, n.º 4, págs. 313-321, 2002. DOI: 10.15381/anales.v63i4.1513. dirección: <https://ojsvrip.unmsm.edu.pe/index.php/anales/article/view/1513>.
- [3] M. J. Keeling y K. T. D. Eames, «Networks and epidemic models,» *Journal of The Royal Society Interface*, vol. 2, n.º 4, págs. 295-307, 2005. DOI: 10.1098/rsif.2005.0051. dirección: <https://doi.org/10.1098/rsif.2005.0051>.
- [4] M. E. J. Newman, *Networks*. Oxford: Oxford University Press, 2018, ISBN: 9780198805090.
- [5] N. Parekh y A. L. Deierlein, «Health behaviours during the coronavirus disease 2019 pandemic: Implications for obesity,» *Public Health Nutrition*, vol. 23, n.º 17, págs. 3121-3125, 2020. DOI: 10.1017/S1368980020003031. dirección: <https://doi.org/10.1017/S1368980020003031>.
- [6] L. Bonnetaín, R. Goel, R. Sharma y A. Furno, «A mobility-based SIR model for complex networks: with case study of COVID-19,» *Social Network Analysis and Mining*, vol. 11, n.º 1, pág. 104, 2021. DOI: 10.1007/s13278-021-00814-3. dirección: <https://pmc.ncbi.nlm.nih.gov/articles/PMC8532458/>.
- [7] C. S. Carlson, D. M. Rubin, V. Heikkilä y M. Postema, «Extracting transmission and recovery parameters for an adaptive global system dynamics model of the COVID-19 pandemic,» en *2021 IEEE AFRICON*, Arusha, Tanzania, United Republic of, 2021, págs. 1-4. DOI: 10.1109/AFRICON51333.2021.9570946.
- [8] R. Goel, L. Bonnetaín, R. Sharma y A. Furno, «Mobility-based SIR model for complex networks: with case study of COVID-19,» *Social Network Analysis and Mining*, vol. 11, n.º 1, pág. 105, 2021, Epub 2021 Oct 22. DOI: 10.1007/s13278-021-00814-3.
- [9] M. F. J. Lafebre, J. A. S. Ávila, M. F. B. Armas, S. L. F. García y M. F. C. Morales, «Modelado de Propagación de Enfermedades Infecciosas: Modelos SIR, SIS y SEIR: Revisión Sistemática / Modeling the Spread of Infectious Diseases: SIR, SIS and SEIR Models: Systematic Review,» *LATAM Revista Latinoamericana de Ciencias Sociales y Humanidades*, vol. 6, n.º 1, págs. 2329-2658, 2025, Accedido: 2025-11-18. DOI: 10.56712/latam.v6i1.3524. dirección: <https://dialnet.unirioja.es/descarga/articulo/10080869.pdf>.
- [10] NetworkX Developers, *NetworkX barabasi_albert_graph — NetworkX Documentation*, https://networkx.org/documentation/stable/reference/generated/networkx.generators.random_graphs.barabasi_albert_graph.html, Accessed: 2025-11-18, 2025.
- [11] NetworkX Developers, *NetworkX load_centrality — NetworkX Documentation*, https://networkx.org/documentation/stable/reference/algorithms/generated/networkx.algorithms.centrality.load_centrality.html, Accessed: 2025-11-18, 2025.
- [12] Secretaría de Salud, Gobierno de México, *COVID-19: documentos y recursos oficiales*, <https://www.gob.mx/salud/documentos/covid19>, Accessed: 2025-11-18, 2025.
- [13] L. Verma, R. Meher, O. Nikan, A. A. Al-Saedi et al., «Numerical study on fractional order nonlinear SIR-SI model for dengue fever epidemics,» *Scientific Reports*, vol. 15, n.º 1, pág. 30677, 2025. DOI: 10.1038/s41598-025-16599-w. dirección: <https://www.nature.com/articles/s41598-025-16599-w>.