

ESCUELA POLITÉCNICA NACIONAL

FACULTAD DE CIENCIAS

ESTUDIO DE LAS CURVAS DE CONTAGIO GENERADAS POR EL
MODELO EPIDEMIOLÓGICO SIR SOBRE UNA RED COMPLEJA

TRABAJO DE TITULACIÓN PREVIO A LA OBTENCIÓN DEL TÍTULO DE FÍSICA
PROYECTO DE INVESTIGACIÓN

MARLLI ANDREA ZAMBRANO ZAMBRANO
marlli.zambrano@epn.edu.ec

Director: DR. RAMON XULVI-BRUNET
ramon.xulvi@epn.edu.ec

QUITO, AGOSTO 2022

DECLARACIÓN

Yo, MARLLI ANDREA ZAMBRANO ZAMBRANO, declaro bajo juramento que el trabajo aquí descrito es de mi autoría; que no ha sido previamente presentado para ningún grado o calificación profesional; y que he consultado las referencias bibliográficas que se incluyen en este documento.

A través de la presente declaración cedo mis derechos de propiedad intelectual, correspondientes a este trabajo, a la Escuela Politécnica Nacional, según lo establecido por la Ley de Propiedad Intelectual, por su Reglamento y por la normativa institucional vigente.

Marlli Andrea Zambrano Zambrano

CERTIFICACIÓN

Certifico que el presente trabajo fue desarrollado por MARLLI ANDREA ZAMBRANO ZAMBRANO, bajo mi supervisión.

Dr. Ramon Xulvi-Brunet
Director del Proyecto

AGRADECIMIENTOS

Quiero agradecer a todas aquellas personas que aportaron de una u otra manera en mi formación profesional y personal a lo largo de la carrera, en especial al Dr. Ramon Xulvi-Brunet, mi tutor, por todo el tiempo dedicado, los conocimientos impartidos, su infinita paciencia y por animarme a seguir en el camino de la investigación.

Quiero también agradecer de manera particular al Dr. Álvaro Dueñas, quien más que un excelente profesor, ha sido una guía y una mano amiga, sobre todo, en los momentos más críticos de la carrera. Al Dr. Leonardo Bacile, quien, en estos últimos años de pandemia, supo preocuparse por mis estudios y me motivó a seguir adelante.

Gracias a mis padres Camilo y Genith, por ser los principales promotores de mis sueños, gracias a ellos por cada día confiar y creer en mí y en mis expectativas. A mis hermanos Cristopher y Evelyn, no ha sido sencillo el camino hasta ahora, pero gracias a sus aportes, a su amor, a su inmensa bondad y apoyo, lo complicado de lograr esta meta se ha notado menos.

Mi más profundo agradecimiento a Daniel por todo su amor, su carisma y su incondicional apoyo no solo para el desarrollo de esta tesis, sino también para mi vida; eres mi inspiración y mi motivación.

A mis poli-amigos Pedrito, Jack y Yess, quienes hicieron de esta aventura algo extraordinario. También, al Club de Física Teórica EPN por brindarme un espacio para compartir sobre ciencia y enseñarme a enfrentar mis miedos.

Y por último, pero no menos importante, a mis Big-Friends: Darío, John y Jeff, quienes me regalaron un poquito de su tiempo para pulir el escrito de este trabajo y me brindaron muchos momentos increíbles.

'Fiziko estas kun vi' -Darío Arroyo

DEDICATORIA

A mi abuelito Pepe, mi padrino Abel y mi primo Andrés. Mis ángeles en el cielo.

A la ciencia.

“Nada en la vida es para ser temido, es sólo para ser comprendido. Ahora es el momento de entender más, de modo que podamos temer menos”

Marie Curie (1867 – 1934)

Índice general

1. Introducción	1
1.1. Antecedentes	1
1.2. Modelos Epidemiológicos	2
1.3. Introducción a las Redes Complejas	4
1.4. 'Outline' del Trabajo de Titulación	6
2. Fundamento Teórico	7
2.1. Clasificación de las redes complejas	7
2.1.1. Red Erdős-Rényi o Aleatoria	8
2.1.2. Red de libre escala: modelo Barabási-Albert	9
2.2. Ecuaciones de Kermack y McKendrick: Modelo SIR	11
2.3. Método Runge Kutta para resolución de EDOs	12
2.4. Algoritmo de Gillespie	13
3. Simulaciones Computacionales	17
3.1. Modelización de las Redes Complejas	17
3.1.1. Red Completa	17
3.1.2. Red Aleatoria	18
3.1.3. Red Barabási-Albert	18
3.2. Algoritmo de Gillespie para la resolución del modelo SIR	19
3.3. Algoritmo de Gillespie para la resolución del modelo SIR sobre una red	22
4. Resultados y Análisis	26
4.1. Degree distribution de las redes creadas	26

4.2. Gillespie vs Runge Kutta	27
4.3. Curvas de contagio del modelo SIR sobre una red compleja	30
4.3.1. Para la red completa	30
4.3.2. Para la red aleatoria	32
4.3.3. Para la red Barabási-Albert	33
5. Conclusiones	36
A. Algoritmo	38
A.1. Resolución por el método de Runge Kutta (RK4)	38
B. Códigos	40
B.1. Runge Kutta	40
B.2. Gillespie para resolución de EDOs	41
B.3. Creación de la Red Aleatoria	43
B.4. Creación de la Red Barabasi-Albert	44
B.5. Gillespie para resolución de EDOs sobre una red compleja	46
Bibliografía	54

Índice de figuras

1.1. Estructura de una red compleja con ejes dirigidos. Las flechas indican la direccionalidad de cada enlace.	5
2.1. Curvas de contagio obtenidas del modelo SIR con parámetros $\beta = 2,1 \times 10^{-4}$ y $\gamma = 0,5$ (estos valores son totalmente arbitrarios).	11
3.1. Intervalo de probabilidades.	19
4.1. (a)-Degree distribution de una red aleatoria de 5 mil nodos. (b)-Degree distribution de una red de libre escala de $2,011 \times 10^6$ millones de nodos en escala logarítmica. (Los puntos rojos son valores experimentales, tomados cada 10 iteraciones, con sus respectivas barras de error)	27
4.2. Curvas de contagio del modelo SIR utilizando el Método Runge Kutta y el algoritmo de Gillespie con $N = 50 \times 10^3$, $\beta = 1,02 \times 10^{-5}$ y $\gamma = 0,04$: (a) Susceptibles, (b) Infectados, (c) Recuperados. (Las barras de error se tomaron cada 200 iteraciones temporales con el fin apreciar de mejor manera la curva.)	28
4.3. Curvas de contagio del modelo SIR con $N = 5000$, $\beta = 0,13$ y $\gamma = 0,048$ utilizando el método Runge Kutta y Gillespie: (a) Susceptibles, (b) Infectados, (c) Recuperados. (Las barras de error se tomaron cada 200 iteraciones temporales con el fin apreciar de mejor manera la curva.)	29
4.4. Curvas de contagio del modelo SIR sobre una red completa de $N = 5000$, $\beta = 1,02 \times 10^{-4}$ y $\gamma = 0,04$: (a) Susceptibles, (b) Infectados, (c) Recuperados. (Las barras de error se tomaron cada 200 iteraciones temporales con el fin apreciar de mejor manera la curva.)	31

4.5. Curvas de susceptibles del modelo SIR sobre redes aleatorias con diferente número de enlaces, $N = 5000$, $\beta = 6,626 \times 10^{-5}$ y $\gamma = 0,0833$: (a) Susceptibles, (b) Infectados, (c) Recuperados. (Las barras de error se tomaron cada 200 iteraciones temporales con el fin apreciar de mejor manera la curva.)	32
4.6. Curvas de contagio del modelo SIR sobre una Red Barabasi con $N = 5000$, $\beta = 0,40$ y $\gamma = 0,035$: (a) Susceptibles, (b) Infectados, (c) Recuperados. (Las barras de error se tomaron cada 200 iteraciones temporales con el fin apreciar de mejor manera la curva.)	34
4.7. Curvas de contagio del modelo SIR sobre una Red Barabasi con $N = 2,011 \times 10^6$, $\beta = 0,40$ y $\gamma = 0,035$.(Las barras de error se tomaron cada 50×10^3 iteraciones temporales con el fin apreciar de mejor manera la curva.)	35

Resumen

Poder estudiar el comportamiento de una enfermedad infecciosa dentro de una población ha sido de vital importancia, en especial, estos últimos años. Una pandemia tiene diferentes repercusiones dentro de una sociedad, como restricciones sociales, de movilidad, el déficit económico y político, entre otros. Por esta razón, y desde hace ya muchas décadas, se ha tratado de desarrollar modelos epidemiológicos cuantitativos para estudiar la dinámica de transmisión de enfermedades.

Los modelos más exitosos han sido los basados en ecuaciones diferenciales ordinarias (EDOs). Este proyecto se enfocará únicamente en el modelo compartimental SIR, que divide a la población de acuerdo a su estado de salud: Susceptible, Infectado y Recuperado. Sin embargo, estos modelos tienen como hipótesis, poco realista, que cada par de individuos está en permanente contacto entre sí, es decir, se considera una población perfectamente mezclada.

Este trabajo de investigación busca estudiar las curvas de contagio generadas por el modelo SIR cuando se considera una población perfectamente mezclada y una que no lo está. Para ello, consideramos los contactos entre los individuos de la población como una red compleja y evaluamos la dinámica de la enfermedad sobre ella.

Utilizando un algoritmo estocástico levemente modificado (Algoritmo de Gillespie), se logra comprobar que para el estudio de propagación de enfermedades se debe considerar los medios de contagio, pues la dinámica de una epidemia depende de la estructura de contactos entre cada par de individuos de una población.

Abstract

Being able to study the behavior of an infectious disease within a population has been of vital importance, especially in recent years. A pandemic has different repercussions within a society, such as social restrictions, mobility, and economic and political deficit, among others. For this reason, and for many decades now, attempts have been made to develop quantitative epidemiological models to study disease transmission dynamics.

The most successful models have been those based on ordinary differential equations (ODEs). This project will focus on the SIR compartmental model, which divides the population according to their health status: Susceptible, Infected, and Recovered. However, these models have the unrealistic assumption that each pair of individuals is in permanent contact with each other, meaning that, they are considered a perfectly mixed population.

This research project seeks to study the contagion curves generated by the SIR model, ruling out this hypothesis and analyzing whether there is any relevant change in the contagion curves if a perfectly mixed population and one that is not, are considered. For this, we consider the contact between the individuals of the population as a complex network and evaluate the dynamics of the disease on it.

Using a slightly modified stochastic algorithm (Gillespie's algorithm), it is possible to verify that for the study of disease propagation, the means of contagion must be considered, since the dynamics of an epidemic depend on the structure of contacts between each pair of individuals of a population.

Capítulo 1

Introducción

En este Capítulo se presenta una breve descripción acerca de las redes complejas y modelos epidemiológicos. Se expone una rápida revisión del modelo epidemiológico SIR, que es el modelo que consideraremos en este proyecto. Además, se introducen los conceptos necesarios para entender qué son las redes complejas y cómo estas están ligadas a la propagación de enfermedades. Finalmente, en la Sección 1.4, se lista el contenido de los siguientes capítulos de este trabajo de titulación a modo de “outline”.

1.1. Antecedentes

Históricamente las enfermedades infecciosas han tenido un gran impacto dentro de las sociedades [1]. Cuando un gran número de personas de una región se infecta en un período de tiempo relativamente corto se denomina epidemia [1]. Por ejemplo, la peste negra fue la más devastadora de la historia y tuvo lugar en el siglo XIV en Europa, dejando entre 80 a 200 millones de fallecidos [2].

En los últimos años, estudiar la dinámica de transmisión de una enfermedad dentro de la sociedad ha sido de vital importancia [3]. Pandemias como el Covid-19 no solo afecta el ámbito social por las restricciones de acercamiento, sino también el área política y económica de una comunidad. Es por ello que poder pronosticar la evolución de un patógeno permite brindar a las autoridades sanitarias herramientas que posibiliten y faciliten la implementación de protocolos adecuados para sobrellevar la transmisión del

mismo [3, 4].

Dentro de las matemáticas, las ciencias físicas y la estadística existen diferentes herramientas y técnicas que permiten desarrollar e implementar modelos computacionales capaces de simular la propagación de enfermedades infecciosas [3]. Sin embargo, su desarrollo es complejo debido a la cantidad de variables que puede adquirir el sistema, como son el entorno del patógeno y sus huéspedes, la población expuesta; edad, sexo, entre otros [3, 5].

Los primeros modelos epidemiológicos comenzaron a surgir en 1766 cuando Daniel Bernoulli presentó un modelo matemático que describía los efectos de la variación de la viruela (un precursor de la vacunación) sobre la esperanza de vida de los individuos [5, 6]. Sin embargo, en 1927, Kermack y McKendrick publicaron una serie de artículos en los que describían la dinámica de la transmisión de enfermedades en términos de un sistema de ecuaciones diferenciales, siendo así, los pioneros en esta área [6, 7].

Existe una variedad de modelos epidemiológicos que permiten obtener predicciones sobre el avance de una enfermedad dentro de una población [5]. Cada uno de ellos tiene un abordaje diferente, presentando diferentes ventajas y desventajas que afectan la exactitud del estudio.

1.2. Modelos Epidemiológicos

En general, los modelos matemáticos pueden clasificarse de acuerdo a las diferentes características que definen su naturaleza [3]. La mayoría de los modelos deterministas clásicos están basados en el uso de ecuaciones diferenciales ordinarias brindándoles un carácter continuo [3, 6].

Los modelos presentados por Kermack y McKendrick son los primeros denominados compartimentales debido a que la población es dividida de acuerdo a su estado de salud respecto al patógeno introducido [3, 5, 6]. Por ejemplo, se puede clasificar a los

individuos en Susceptibles (S) ¹, Infecciosos (I), Recuperados (R), Expuestos (E) ² [5].

El presente proyecto se enfocará en el modelo epidemiológico SIR que viene descrito por el siguiente conjunto de ecuaciones:

$$\frac{dS}{dt} = -\beta SI \quad (1.1)$$

$$\frac{dI}{dt} = \beta SI - \gamma I \quad (1.2)$$

$$\frac{dR}{dt} = \gamma I \quad (1.3)$$

donde, γ es la tasa de recuperación (de tal manera que el período de tiempo medio de que una persona infectada se cure es $1/\gamma$) y β es el parámetro de infección (es la probabilidad de que un susceptible en contacto con un infeccioso se infecte). Estas ecuaciones deben cumplir que $S + I + R = N$, siendo N el número personas en la población [7].

Asimismo, existen varios modelos compartimentales, como el SI, SIS, SEIR, entre otros, donde variaciones del SIR son consideradas. Por ejemplo, en el modelo SEIR, a diferencia del modelo SIR, los individuos susceptibles cuando se infectan primero pasan al estado expuesto, E. Después de un período de incubación, pasan a ser personas con la capacidad de infectar a otros individuos, I [8–10].

La solución a estos sistemas de ecuaciones diferenciales da como resultado las curvas de contagio, que permiten identificar como varía el número de personas susceptibles, infectadas y recuperadas a lo largo del tiempo.

A estos modelos se les considera modelos globales ya que estudian la dinámica de la población en su conjunto, es decir, consideran una población en la cual todos los individuos están en contacto uno con otros (población perfectamente u homogéneamente

¹Personas que pueden infectarse.

²Personas que están infectadas pero no son transmisoras todavía.

mezclada) [3, 7, 11]. En epidemiología, los 'contactos' entre individuos son el medio por el cual la enfermedad se transmite, por ejemplo, el VIH se transmite mediante la sangre o las relaciones sexuales, o el Covid-19 por las secreciones de la boca y la nariz [12, 13].

En consecuencia, a pesar de que estos modelos matemáticos permiten el análisis de la dinámica de una enfermedad, no captan de forma precisa lo que está pasando en realidad. ¿Cómo se podría mejorar estos modelos?. ¿Es posible considerar el medio de transmisión del virus y cómo afectaría el hecho de que no todos estén en contacto con todos? Este proyecto quiere contestar estas preguntas mediante el uso de las redes complejas.

El área de los sistemas complejos brinda herramientas que permiten el estudio de la conducta de propagación de un virus. La teoría de los sistemas complejos permite introducir las redes complejas como la 'red de contactos' entre los individuos y analizar la evolución de la enfermedad dentro de la población de manera más realista.

1.3. Introducción a las Redes Complejas

El estudio de las redes, en forma de teoría matemática de grafos, es uno de los pilares fundamentales de las matemáticas discretas [14]. Desde su nacimiento en 1736, cuando el matemático suizo Leonhard Euler encontró la solución al problema de los siete puentes de Königsberg ³, la teoría de las redes ha visto logros importantes en algunos contextos especializados, como por ejemplo en las ciencias sociales, los sistemas biológicos, físicos y químicos acoplados, las redes neuronales, entre otros [15, 16].

Una red compleja \mathcal{R} consiste de un conjunto de nodos \mathcal{V} y un conjunto de pares ordenados $\mathcal{E} \subset \mathcal{V} \times \mathcal{V}$. Cada pareja ordenada (v_i, v_j) es llamada eje o enlace dirigido del nodo v_i hacia el nodo v_j [3].

³El problema de los siete puentes de Königsberg se basa en encontrar un viaje de ida y vuelta que atraviese cada uno de los puentes de la ciudad prusiana de Königsberg exactamente una vez.

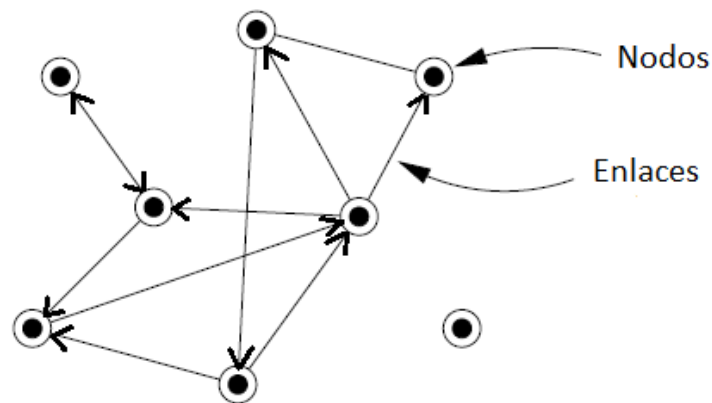


Figura 1.1: Estructura de una red compleja con ejes dirigidos. Las flechas indican la direccionalidad de cada enlace.

En la figura 1.1 se puede apreciar la estructura de una red con ejes dirigidos sencilla. Sin embargo, existe una gran variedad de características que pueden hacerlas más complejas [14, 16]. Por ejemplo, dependiendo del sistema a estudiar, los nodos y/o los enlaces pueden adquirir ciertas propiedades que los diferencien entre ellos [14].

En general, el estudio de las redes complejas puede dividirse en dos áreas diferentes y complementarias: Estructura y Dinámica [17]. El primer ámbito se centra en determinar las propiedades topológicas de la red, mientras que el segundo se enfoca en los fenómenos dinámicos que ocurren sobre ellas. También está la interrelación entre ambas áreas, donde se estudia cómo afecta la topología al proceso dinámico que está ocurriendo sobre ella. En esta última área se basará este proyecto. Se estudiará cómo afecta la estructura de la red en un proceso de propagación de una enfermedad.

Dentro del estudio de la estructura de una red existen algunas propiedades topológicas importantes como el degree distribution $P(k)$, que es la probabilidad de que un nodo escogido al azar tenga k conexiones, o el clustering coefficient, que se relaciona con la probabilidad de que dos nodos conectados directamente a un tercer nodo, estén conectados entre sí [14, 17].

Las redes en el mundo real, como la sociedad o Internet, son redes de libre escala, en las cuales el degree distribution viene dado, aproximadamente, por una ley de po-

tencia, $P(k) = k^{-\alpha}$ [18, 19]. Barabási y Albert fueron de los primeros en proponer un modelo capaz de obtener redes de libre escala [18]. Si se lograra observar cómo se propaga una enfermedad dentro de esta red, se tendría una aproximación más realista de la evolución de una epidemia. En este contexto, las personas o familias dentro de una sociedad vendrían representadas por nodos y, el medio de transmisión por enlaces. Los ejes serán dirigidos de la persona infectada a la sana.

Existen redes con ejes dinámicos, es decir, que varían en el tiempo, sea en su magnitud, su direccionalidad o su posición (que puedan desaparecer o que conecte un par de nodos diferentes cada cierto tiempo) [20]. Esto permite estudiar la dinámica de la enfermedad dentro de la sociedad de manera más precisa. Sin embargo, para este proyecto, consideraremos redes estáticas dado que solo queremos analizar como influye los contactos de las personas durante una epidemia, no construir un modelo totalmente realista.

Este trabajo de titulación tiene como objetivo principal observar si existe un cambio sustancial en las curvas de contagio del modelo epidemiológico SIR obtenidas si se considera un población perfectamente mezclada (solución clásica) o una población donde no todos estén conectados entre sí. Para ello, se construirán tres redes con topología diferente, entre ellas una red de libre escala, para estudiar como evoluciona la epidemia sobre ellas y comparar los resultados con la solución clásica.

1.4. 'Outline' del Trabajo de Titulación

Este trabajo consta de cuatro capítulos adicionales. En el Capítulo 2 se presentan los fundamentos teóricos necesarios para comprender la clasificación de las redes complejas, qué es una red aleatoria, una red Barabási-Albert y el modelo epidemiológico SIR. En el Capítulo 3 se plantea el desarrollo numérico utilizado para obtener las curvas de contagio sobre la red. Los resultados obtenidos se presentan en el Capítulo 4 y por último, el Capítulo 5 recopila las conclusiones del trabajo.

Capítulo 2

Fundamento Teórico

El presente capítulo expone varios conceptos importantes sobre las redes complejas y su clasificación (Sección 2.1). Se detalla, con mayor profundidad, las ecuaciones de Kermack y McKendrick; el modelo epidemiológico SIR. Además, se describen los algoritmos que permitirán obtener las curvas de contagio del modelo SIR.

2.1. Clasificación de las redes complejas

En el capítulo anterior se definió una red como un conjunto de puntos que pueden o no estar conectados entre sí. Para redes de decenas o cientos de nodos resulta relativamente sencillo dibujar la red con puntos y líneas, como se muestra en la figura 1.1 [14]. Sin embargo, para redes que contengan millones de nodos, las propiedades topológicas que permiten describirlas estadísticamente son la única manera de saber cómo están estructuradas sin necesidad de dibujarlas [14].

La topología de la red juega un papel importante al momento de estudiar procesos dinámicos sobre estas [21]. Las redes se pueden caracterizar por muchas medidas diferentes, cada una de las cuales refleja algunos rasgos particulares de la estructura estudiada [22]. Dado que, en este estudio no es tan importante la topología exacta, sino el hecho de que no todos los nodos están conectados con todos, solo nos centraremos en el degree distribution.

El degree o grado k_i corresponde al número de conexiones que tiene el nodo i [14, 22]. Se denominan a todos los nodos que estén conectados directamente al nodo i , los vecinos de i [17].

No todos los nodos de una red tienen el mismo degree. En este sentido, se define una función de distribución: el degree distribution $P(k)$, que da la probabilidad de que un nodo seleccionado aleatoriamente tenga exactamente k enlaces [23].

Por ejemplo, para redes aleatorias, $P(K)$ viene dada por una distribución binomial [23]. Por otro lado, para redes como el World Wide Web o el internet, se tienen redes libres de escala [14, 16, 23].

2.1.1. Red Erdős-Rényi o Aleatoria

Desde una perspectiva de modelado una red consta solo de nodos y enlaces. Sin embargo, decidir dónde colocar los enlaces entre los nodos, para que reproduzcamos la complejidad de un sistema real, es el verdadero desafío [24].

Las redes aleatorias fueron introducidas por Paul Erdős y Alfréd Rényi en 1959, las cuales consisten en un modelo estocástico con N nodos, donde cada par de nodos está conectado con probabilidad p . [23, 25–32]. En este modelo se asume que el número de nodos es fijo y que p es independiente del grado del nodo [28, 33].

Cada red aleatoria generada con los mismos parámetros N y p se ve ligeramente diferente, debido a la estocasticidad del modelo. No solo cambia el diagrama de conexiones, sino también el número de enlaces. El número esperado de enlaces en estas redes viene dado por la ecuación (2.1) [24].

$$\langle N_e \rangle = p \frac{N(N-1)}{2} \quad (2.1)$$

Es decir, $\langle N_e \rangle$ es el producto de la probabilidad de que dos nodos estén conectados, p , y el número de pares máximo que se pueden conectar, que es $N_{e_{max}} =$

$N(N - 1)/2$ (cuando $p = 1$) [23,24]. Si se llegase a tener el número máximo de enlaces se obtiene una red completa [23].

Dada la aleatoriedad de conexión, existen nodos con mayor grado que otros, inclusive nodos sin ningún vecino [24, 26]. Estas diferencias son capturadas por el degree distribution.

En una red aleatoria la probabilidad de que el nodo i tenga exactamente k enlaces es el producto de tres términos [34]:

1. La probabilidad de que k de sus enlaces estén presentes: p^k .
2. La probabilidad de que falten los enlaces restantes ($N - 1 - k$): $(1 - p)^{N-1-k}$.
3. El número de formas en que podemos seleccionar k enlaces de $N - 1$ enlaces potenciales que puede tener un nodo:

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

En consecuencia, el degree distribution de una red aleatoria sigue la distribución binomial 2.2.

$$P(k) = \binom{N - 1}{k} p^k (1 - p)^{N-1-k} \quad (2.2)$$

La forma de esta distribución depende del tamaño N y la probabilidad p .

2.1.2. Red de libre escala: modelo Barabási-Albert

A pesar de que se conocía que las redes que se pueden encontrar en la naturaleza, como la sociedad o Internet eran de libre escala, no se sabía por qué [35]. Barabási y Albert fueron los primeros en proponer un modelo capaz de explicar cómo se obtienen estas redes.

Una red de libre escala se caracteriza por tener un degree distribution que sigue una ley de potencias,

$$P(k) \sim k^{-\alpha} \quad (2.3)$$

donde α se denomina el exponente de grado y en la mayoría de las redes reales este suele tomar un valor entre 1 y 3 [33, 35].

La mayoría de las redes del mundo real describen sistemas abiertos que crecen mediante la adición continua de nodos nuevos [23, 33]. Además, los nodos exhiben una preferencia de conexión de acuerdo a sus grados [23].

Estas dos características: el crecimiento y la preferencia de conexión, permitieron a Barabási presentar un modelo capaz de construir redes de libre escala [23, 36]. En este sentido, el algoritmo del modelo Barabási-Albert es el siguiente [23, 33, 36]:

- **Crecimiento:** Se empieza con un pequeño número n_0 de nodos conectados aleatoriamente con c_0 enlaces. Luego, en cada paso de tiempo, se agrega un nuevo nodo con $m < n_0$ enlaces que vinculan el nuevo nodo con m nodos diferentes ya presentes en el sistema.
- **Preferencia de conexión:** Para unir un nodo viejo i con el nuevo nodo, se utiliza la probabilidad:

$$\Pi(k_i) = \frac{k_i}{\sum_j k_j} \quad (2.4)$$

donde k_i es el grado del nodo i y se cumple que $\sum_i \Pi(k_i) = 1$. Es decir que la preferencia de conexión se relaciona únicamente con el grado de los nodos.

Después de t pasos de tiempo, este procedimiento da como resultado una red con $N = n_0 + t$ nodos y $c_0 + m * t$ enlaces.

El degree distribution de este modelo viene dada por la ecuación (2.5) [33, 36]. Cuan-

do k es suficientemente grande $P(k)$ tiende a k^{-3} , teniendo redes de libre escala [31].

$$P(k) = \frac{2m(m+1)}{k(k+1)(k+2)} \quad (2.5)$$

Las ecuaciones (2.2) y (2.5), nos permitirán comprobar que las redes creadas en el Capítulo 3 son correctas.

2.2. Ecuaciones de Kermack y McKendrick: Modelo SIR

El modelo epidemiológico SIR fue introducido en 1927 por Kermack y McKendrick para estudiar las epidemias de peste y cólera en Londres y Bombay [37]. Sin embargo, fuera de la epidemiología, el modelo SIR también se usa ampliamente en el modelado de redes sociales en línea, marketing viral, difusión de ideas, propagación de virus informáticos, contagio de redes financieras, entre otros [37].

Las ecuaciones presentadas en el capítulo anterior: (1.1), (1.2) y (1.3) suponen una enfermedad a la que ningún individuo tiene inmunidad, una pequeña cantidad de infectados y una gran cantidad de individuos susceptibles [7, 37].

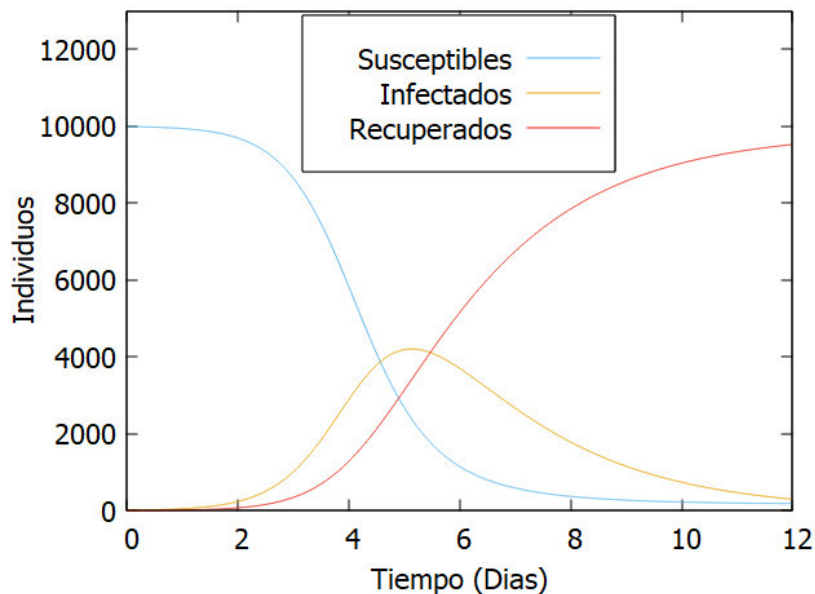


Figura 2.1: Curvas de contagio obtenidas del modelo SIR con parámetros $\beta = 2,1 \times 10^{-4}$ y $\gamma = 0,5$ (estos valores son totalmente arbitrarios).

Las curvas que dan solución a estas ecuaciones diferenciales, curvas de contagio, se pueden resolver mediante varios métodos numéricos como el Runge Kutta, método que se presentará brevemente en la siguiente sección. Por ejemplo, en la figura 2.1, se puede apreciar las soluciones para un sistema de $N = 10000$ personas y $I_0 = 10$ (Infectados iniciales) con los parámetros $\beta = 2,1 \times 10^{-4}$ y $\gamma = 0,5$, resueltas por el método Runge Kutta.

2.3. Método Runge Kutta para resolución de EDOs

Los métodos de Runge-Kutta son métodos iterativos usados para integrar sistemas de ecuaciones diferenciales ordinarias (EDOs) numéricamente. Se realiza evaluando derivadas en puntos intermedios entre varios pasos iterativos sucesivos [38, 39].

Para derivar este método, en primer lugar, se considera un sistema de n EDOs generales de primer orden, es decir, ecuaciones del tipo:

$$y'(t) = f(t, y(t)) \quad a \leq t \leq b, \quad (2.6)$$

La solución más general a este sistema viene dada de la forma [38, 40]:

$$y_{j+1} = y_j + \Delta t \sum_{i=1}^m b_i k_i \quad (2.7)$$

donde y_j es el vector de dimensión n que contiene las variables del sistema en el j -ésimo paso. Eligiendo diferentes valores de los parámetros m y b_i se obtiene una fórmula de cálculo de Runge Kutta de orden superior. En específico, para este proyecto se ocupará los parámetros para RK4; Runge Kutta de cuarto orden ($m=4$) [39],

$$y_{n+1} = y_n + \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4), \quad (2.8)$$

donde,

$$k_1 = hf(t_n, y_n) \quad (2.9)$$

$$k_2 = hf\left(t_n + \frac{1}{2}h, y_n + \frac{1}{2}k_1\right) \quad (2.10)$$

$$k_3 = hf\left(t_n + \frac{1}{2}h, y_n + \frac{1}{2}k_2\right) \quad (2.11)$$

$$k_4 = hf(t_n + h, y_n + k_3) \quad (2.12)$$

Es decir, se necesita cuatro valores de función en cada iteración de paso. Por otro lado, los k_i son vectores que se obtienen al evaluar el sistema en puntos intermedios (entre $[y_{j+1}, y_j]$ y entre $[t, t + \Delta t]$) y se conocen como etapas del método de Runge-Kutta [38, 41].

Mediante este método se podrán obtener las soluciones clásicas a las ecuaciones diferenciales del modelo epidemiológico SIR y posteriormente comparar estos resultados con las curvas obtenidas por el algoritmo de Gillespie, presentado en la siguiente sección.

La implementación computacional de este algoritmo se encuentra en el Anexo A.1.

2.4. Algoritmo de Gillespie

En 1976, Gillespie desarrolló un algoritmo estocástico capaz de calcular numéricamente la evolución temporal de prácticamente cualquier sistema químico espacialmente homogéneo [42]. Sin embargo, actualmente se utiliza para resolver otros procesos estocásticos dentro de sistemas sociales dinámicos, sistemas ecológicos e incluso sistemas

tecnológicos [43].

En la formulación estocástica, la evolución temporal no se describe mediante un conjunto de N pares de ecuaciones diferenciales para las concentraciones de especies (método determinista), sino mediante una única ecuación diferencial para una gran función de probabilidad en la que el tiempo y las N poblaciones de especies aparecen como variables independientes. A esta ecuación diferencial la denominan la ecuación maestra [42, 44, 45].

Gillespie basó su algoritmo en una entidad que llamó función de densidad de probabilidad de reacción,

$$P(\tau, \mu)d\tau, \quad (2.13)$$

la probabilidad al tiempo t de que la siguiente reacción en el volumen V ocurrirá en el intervalo temporal diferencial $(t + \tau, t + \tau + d\tau)$ y será la reacción R_μ , donde μ puede ser cualquiera de las M reacciones posibles que existen en el sistema ($\mu = 1, 2, \dots, M$). Aquí, τ es una variable continua y μ una discreta [42, 46].

Para hallar la expresión de $P(\tau, \mu)$, se empieza definiendo, para M reacciones, h_μ como el número de combinaciones de reactivos moleculares distintas para la reacción R_μ , que se encuentran presentes en V en el tiempo t . Y, c_μ , los parámetros de reacción que caracteriza a la reacción R_μ [42, 45].

Considerando que $h_\mu c_\mu d\tau$ es la probabilidad de que una reacción R_μ ocurra en V en el siguiente intervalo temporal diferencial $(t + \tau, t + \tau + d\tau)$, se define lo siguiente,

$$P(\tau, \mu)d\tau = P_0(\tau) \cdot h_\mu c_\mu d\tau. \quad (2.14)$$

De aquí se obtiene [42],

$$P_0(\tau) = \exp \left\{ - \sum_{l=1}^M h_l c_l \tau \right\}, \quad (2.15)$$

la probabilidad al tiempo t que ninguna reacción ocurrirá en el intervalo $(t, t + \tau)$ [42]. Cabe recalcar que todas las fórmulas presentadas dentro de este algoritmo asumen una probabilidad constante por unidad de tiempo [42].

Gillespie propuso dos variantes diferentes del algoritmo de simulación: el método directo y el método de la primera reacción [43]. En este trabajo se tomará como referencia el primero.

El método directo del algoritmo de Gillespie se construye a partir de $P(\tau, \mu)$ y para ello se considera que cualquier función de densidad de probabilidad de dos variables se puede escribir como el producto de dos funciones de densidad de probabilidad de una variable, como se describe a continuación,

$$P(\tau, \mu) = P_1(\tau) \cdot P_2(\mu|\tau) \quad (2.16)$$

donde,

$$P_1(\tau) = a = \sum_{\mu=1}^M a_{\mu} \exp\{-a\tau\} \quad (2.17)$$

$$P_2(\tau) = \frac{a_{\mu}}{a} \quad (2.18)$$

Aquí, $a_{\mu} = h_{\mu}c_{\mu}$, $a = \sum_{\mu=1}^M a_{\mu}$, $P_1(\tau)d\tau$ es la probabilidad que la siguiente reacción ocurrirá entre el tiempo $t + \tau$ y $t + \tau + d\tau$. Y, $P_2(\mu|\tau)$ la probabilidad que la siguiente reacción será R_{μ} dado que esta ocurre en el tiempo $t + \tau$ [42, 47].

Este algoritmo permite predecir el tiempo en el cual se dará un cambio en el sistema y qué tipo de cambio será. Para esto se considera la ecuación (2.17) obteniendo así [42, 45]:

$$\tau = \frac{1}{a} \ln \left(\frac{1}{r_1} \right) \quad (2.19)$$

donde r_1 es un número aleatorio en una distribución uniforme del 0 al 1.

Una vez encontrado τ , se selecciona el tipo de cambio mediante la condición dada por (2.18) [42, 45]:

$$\sum_j^{\mu-1} a_j \leq r_2 * a \leq \sum_j^{\mu} a_j \quad (2.20)$$

donde, r_2 es un número aleatorio en una distribución uniforme del 0 al 1.

En el Capítulo 3 se modificará ligeramente este algoritmo para adaptarlo al problema es este proyecto; encontrar las curvas de contagio sobre una red compleja.

Capítulo 3

Simulaciones Computacionales

El objetivo de este trabajo de titulación es estudiar las curvas de contagio generadas por el modelo SIR sobre una red compleja no completa y comprobar si existe alguna variación con los resultados obtenidos por el modelo clásico con EDOs. Para probar que el estudio realizado es correcto, primero se encontró las curvas de contagio sobre una red completa, dado que esto sería equivalente a considerar una población perfectamente mezclada. En este sentido, el resultado del algoritmo Gillespie modificado a la resolución de EDOs (Sección 3.2) debería coincidir con el que se obtendría sobre una red completa. En la Sección 3.3 se propone un algoritmo para encontrar las curvas de contagio sobre una red compleja cualesquiera, que supone una adaptación del algoritmo modificado de Gillespie.

3.1. Modelización de las Redes Complejas

3.1.1. Red Completa

Como se definió en el capítulo anterior, una red completa es aquella donde todos los nodos se encuentran conectados entre sí. Esto permite tener una red con N nodos y $N(N - 1)/2$ enlaces.

La creación de esta red es la más sencilla, pues solo se crea una matriz adyacente \mathcal{M} de dimensión $N \times N$, cuyas entradas son todas 1. Cada componente \mathcal{M}_{ij} de esta matriz

represente el enlace del nodo i al nodo j .

3.1.2. Red Aleatoria

Para la construcción de una red Erdős-Rényi o red aleatoria se siguen los siguientes pasos [25–30]:

1. Se definen N número de nodos.
2. Se selecciona un par de nodos y se genera un número aleatorio entre 0 y 1. Si el número es menor que p , se conecta el par de nodos seleccionado con un enlace, de lo contrario se dejan desconectados.
3. Repetir el paso anterior hasta construir los enlaces de interés recordando el límite dado por (2.1).

La matriz adyacente \mathcal{M} , en este caso, tendrá de entradas 1 cuando exista un enlace y 0 cuando no. Sin embargo, para una red no completa es conveniente, por temas de memoria computacional, definir una matriz \mathcal{R} de dimensión $L \times 2$, donde L es el número de enlaces totales. Cada fila i de esta matriz representa un enlace dirigido del nodo R_{i1} al nodo R_{i2} .

Para comprobar que la red que creamos es correcta, calculamos el degree distribution y comparamos con la ecuación (2.2). El código utilizado para crear esta red se presenta en al Anexo B.3.

3.1.3. Red Barabási-Albert

El algoritmo para la construcción de la red Barabási-Albert se detalló en el capítulo previo. Aquí vamos a explicar nuestra implementación del mismo, que nos permitirá generar las redes que usaremos en nuestro estudio.

Para el crecimiento de la red se añadió un nodo 'nuevo' en cada iteración j hasta completar un número límite N . En nuestra implementación del algoritmo, los nodos agregados se conectaron con $m = 4$ nodos ya existentes (nodos viejos) en el sistema con una preferencia de conexión dada por la ecuación (2.4).

Computacionalmente, para el cálculo e interpretación de las probabilidades de cada nodo en la iteración j , tomamos un intervalo de $[0, 1]$ (figura 3.1) dividido en S sub-intervalos asociados a los S nodos viejos existentes en la red. La longitud de cada sub-intervalo i representa la probabilidad de que el nuevo nodo se conecte con el nodo i . Entonces, para el nuevo nodo agregado a la red se toma un valor aleatorio r entre 0 y 1, si r se encuentra en el sub-intervalo i , se conecta al nodo i . Este proceso se repite hasta completar las m conexiones, sin olvidar actualizar Π .

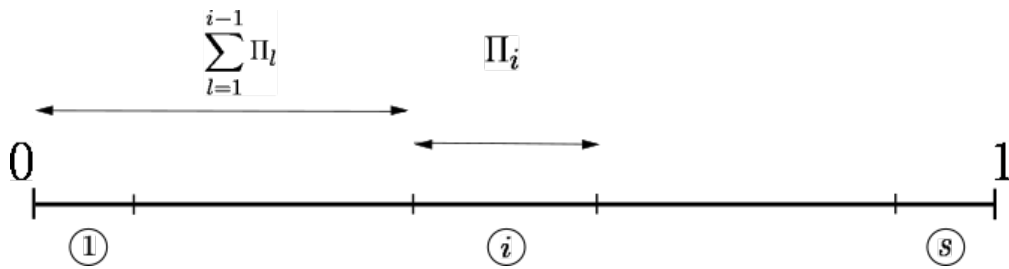


Figura 3.1: Intervalo de probabilidades.

3.2. Algoritmo de Gillespie para la resolución del modelo SIR

En esta sección se presentará una leve modificación al algoritmo de Gillespie descrito anteriormente para poder resolver las ecuaciones diferenciales del modelo epidemiológico SIR de forma estocástica, teniendo como hipótesis una población perfectamente mezclada.

Primero se debe identificar el tipo de 'cambios' que se dan dentro del sistema. En este caso existen dos:

1. **Proceso infeccioso:** Si una persona infectada se encuentra en contacto con una

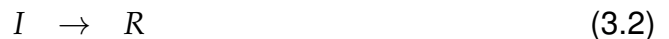
susceptible provocará la propagación de la enfermedad aumentando el número de infectados y disminuyendo el número de susceptibles. Se tiene entonces la siguiente 'reacción':



Recordemos que el intervalo temporal en el cual se dará un cambio en el sistema (ecuación (2.19)) depende del número de combinaciones de reactivos moleculares de cada 'cambio', h_μ , y de los parámetros c_μ .

Esta primera reacción o cambio tiene dos 'reactivos moleculares': S e I . Por lo cual el número de combinaciones $h_1 = SI$. Por otro lado, el parámetro que permite este cambio vendrá dado por el mismo que caracteriza las ecuaciones del modelo epidemiológico SIR: $c_1 = \beta$.

2. **Proceso de recuperación:** Si una persona infectada supera la enfermedad causará que el número de infectados disminuya y aumente el número de personas recuperadas. El 'cambio' será de la forma:



En este proceso se tiene un único 'reactivo molecular': I . En ese sentido, tenemos que definir $h_2 = I$. Al igual que el caso anterior, el parámetro de este proceso será igual al de las ecuaciones del modelo SIR: $c_2 = \gamma$.

Dados los valores iniciales de S , I y R , y los valores de β y γ , el algoritmo procede de manera iterativa:

1. Para poder encontrar el intervalo temporal en el cual se dará un cambio en el sistema e identificar este cambio, se calcula $a_\mu = h_\mu c_\mu$ en cada iteración temporal dada, j :

- $a_1 = \beta S_j I_j$
- $a_2 = \gamma I_j$

2. Se elige aleatoriamente un número r_1 entre 0 y 1. Una vez se obtenga r_1 se calcula el siguiente intervalo temporal τ en el cual se dará un cambio en el sistema utilizando la ecuación (2.19). En consecuencia, t_{j+1} será $t_{j+1} = t_j + \tau$.
3. Se elige aleatoriamente un número r_2 entre 0 y 1. Este número r_2 permitirá conocer que tipo de cambio se dará en el sistema, si una persona se infecta o si una persona infectada se recupera:

- Si $r_2 * a \leq h_1 c_1$ una persona se infecta, por lo cual:

$$S_{j+1} = S_j - 1 \quad y \quad I_{j+1} = I_j + 1$$

En esta 'reacción' $R_{j+1} = R_j$, es decir, permanece constante el número de personas recuperadas.

- Si $r_2 * a > h_1 c_1$ y $r_2 * a \leq a$ una persona se recupera, por lo cual:

$$I_{j+1} = I_j - 1 \quad y \quad R_{j+1} = R_j + 1$$

En esta reacción $S_{j+1} = S_j$, es decir, permanece constante el número de personas susceptibles.

Recordemos que j es el paso dinámico del algoritmo (cuando se producen los cambios en el sistema).

Este proceso iterativo finaliza una vez no existan personas infectadas, pues no habrá ningún foco de infección y tampoco habrá alguien que deba superar la enfermedad.

Dado que, en este algoritmo se está considerando que todos los individuos están conectados entre sí, podemos comparar estos resultados con los obtenidos al implementar

el método de Runge Kutta para resolución de EDOs (Anexo A.1). Ambos modelos tienen la misma hipótesis: una población perfectamente mezclada.

3.3. Algoritmo de Gillespie para la resolución del modelo SIR sobre una red

Como se mencionó en la sección 3.2, de acuerdo al tipo de proceso dado en el sistema se definen los parámetros $h_1 = SI$ y $h_2 = I$ para resolver el sistema de ecuaciones diferenciales dadas por el modelo SIR. Sin embargo, esto solo es verdad cuando todos los individuos están en contacto con todos. ¿Qué pasaría si no es así?

En una red de contactos es importante la topología que la caracteriza. Por ejemplo, una población mezclada de manera homogénea vendría representada por una red completa, pues todos estarían conectados mutuamente. Por otro lado, la red de Barabási-Albert presenta una estructura diferente. En consecuencia, en una red con una cierta topología dada, definiremos los valores de h_μ para este algoritmo de la siguiente manera:

$$h_1 = ne \quad y \quad h_2 = I \quad (3.3)$$

donde ne es el número de contactos (o enlaces) por los cuales se transmitirá la enfermedad.

De las ecuaciones (3.3) se tiene que h_1 ahora depende del medio de propagación de la enfermedad, y no de los valores de S e I . Debido a que la única manera de que un infectado contagie a un susceptible será si ambos están conectados, entonces el único 'reactivo molecular' será el enlace de propagación. En este sentido, si se considera una red completa $ne \approx SI$, por lo que se tendría la hipótesis inicial de la sección anterior.

Por otro lado, que una persona se recupere es independiente de las conexiones que tenga y solo dependerá del número de infectados I en el sistema. Con lo cual, el valor

de $h_2 = I$.

Los valores de c_μ permanecerán iguales debido a que seguimos manejando las ecuaciones del modelo epidemiológico SIR.

Dada una red cualesquiera de N nodos se propone el siguiente algoritmo para obtener las curvas de contagio sobre ellas:

1. Se almacenan los datos de la red en una matriz \mathcal{R}_{ij} con $i = (1, \dots, ct)$ y $j = (1, 2)$. Donde ct representa el número total de enlaces. Cada fila i de esta matriz representa un enlace dirigido del nodo \mathcal{R}_{i1} al nodo \mathcal{R}_{i2} . Recordemos que no utilizamos la matriz adyacente \mathcal{M} debido a la capacidad de almacenamiento del computador.
2. Se calcula el grado de cada nodo k de la red, que denotaremos por G_k .
3. Se almacenan los valores de los parámetros de infección y recuperación: $c_1 = \beta$ y $c_2 = \gamma$.
4. Para identificar el estado de cada nodo se crea un vector de estado E_n con $n = (1, \dots, N)$. Se designará el valor de: 0 para susceptibles, 1 para infectados y 2 para recuperados. Se inicializa $E_n = 0$.
5. Para identificar los ejes de propagación de la enfermedad se crea un vector de estado E_{c_e} con $e = (1, \dots, ct)$. Este vector tendrá como entrada 1 únicamente para ejes que conecten un nodo infectado con uno susceptible (enlaces de propagación). Se inicializa $E_{c_e} = 0$.
6. Se toma un nodo de la red de manera aleatoria, i . Será el primer infectado. Por lo tanto, el estado de este nodo pasará a ser $E_i = 1$.
7. Se dan los valores iniciales de $I_0 = 1$, $S_0 = N - I_0$ y $R_0 = 0$.
8. El número de enlaces de propagación (que conecten infectados con susceptibles) iniciales será igual al grado del único nodo infectado, pues solo está conectado con personas susceptibles. Entonces $ne = G_i$. Dado que estos enlaces serán el

camino por el cuál se propagará la enfermedad, el estado de estos ejes también debe actualizarse según el punto 5. Entonces para todos los ejes p donde i esté conectado con un susceptible $Ec_p = 1$.

9. Se inicia un proceso iterativo para la evolución temporal del sistema. Se inicializa con $t_0 = 0$.
10. Para cada iteración temporal j se calcula $h_1, h_2, a_1 = c_1 * h_1, a_2 = c_2 * h_2$ y $a = \sum_i a_i$
11. Se elige aleatoriamente un número aleatorio r_1 entre 0 y 1. Y se calcula el siguiente intervalo temporal en el cual se dará un cambio en el sistema utilizando la ecuación (2.19)
12. Se elige aleatoriamente un número aleatorio r_2 entre 0 y 1. Este número r_2 permitirá conocer que tipo de cambio se dará bajo las condiciones de la ecuación 2.20:

- Si $r_2 * a \leq h_1 c_1$:

Se toma un enlace de propagación aleatoriamente, p_j . Se infecta al nodo susceptible que se encuentra al final del enlace y se cambia el estado del nodo $E_{susceptible} = 1$.

Debido a que p_j ahora conectará a dos nodos infectados ya no es un enlace de propagación, por lo cual $Ec_{p_j} = 0$. Por otro lado, todos los enlaces z que conecten al nuevo nodo infectado con un susceptible ahora serán de propagación y por tanto $Ec_z = 1$. Con esto se actualiza el número de enlaces de transmisión, ne .

Además:

$$S_{j+1} = S_j - 1 \quad y \quad I_{j+1} = I_j + 1$$

En esta reacción $R_{j+1} = R_j$, se decir, permanece constante el número de personas recuperadas.

- Si $r_2 * a > h_1 c_1$ y $r_2 * a \leq a$:

Se escoge un nodo infectado aleatoriamente, x_j . Este nodo se recupera y se actualiza su estado $E_{x_j} = 2$.

Los enlaces z que conectaban x_j con personas susceptibles anteriormente eran de propagación. Por lo cual, al recuperarse x_j , estos enlaces ya no son de interés y por tanto $E_{c_k} = 0$ y se actualiza el número de enlaces de transmisión, ne .

Además:

$$I_{j+1} = I_j - 1 \quad y \quad R_{j+1} = R_j + 1$$

En esta reacción $S_{j+1} = S_j$, es decir, permanece constante el número de personas susceptibles.

El proceso iterativo culmina cuando no existan nodos infectados, pues habrá terminado de propagarse la enfermedad.

Las redes sobre las cuales se estudió la propagación de una enfermedad fueron tres diferentes: red completa, red aleatoria y red Barbási-Albert. El algoritmo presentado en esta sección fue implementado en FORTRAN 90 y el código se puede encontrar en el Anexo B.5.

Capítulo 4

Resultados y Análisis

En este capítulo se mostrarán los resultados de las simulaciones computacionales utilizadas en este proyecto. En la sección 4.2, se observa los resultados obtenidos para las ecuaciones diferenciales del modelo SIR con el método Gillespie y Runge Kutta para comprobar que ambos métodos nos permiten encontrar las soluciones de EDOs. En la sección 4.3, se muestran las curvas de contagio $S(t)$, $I(t)$ y $R(t)$ conseguidas sobre las redes completas, aleatorias y de libre escala creadas.

4.1. Degree distribution de las redes creadas

En primera instancia se verificó que las redes creadas cumplan con las propiedades topológicas teóricas mencionadas en el capítulo 2. Es decir, que el degree distribution coincida con la expresión teórica conocida para estas redes.

- *Degree distribution de una red Aleatoria:*

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

- *Degree distribution de una red Barabasi-Albert:*

$$P(k) = \frac{2m(m+1)}{k(k+1)(k+2)}$$

En la figura 4.1-(a) se puede apreciar la gráfica de $P(k)$ obtenido para una red aleatoria de 5 mil nodos y $p = 0,40$ (líneas verdes). Como predice la teoría, la gráfica sigue una distribución binomial (línea negra) con un pico alrededor del número $\langle k \rangle = 2 * N_{enlaces} / N = 2000$.

Por otro lado, la figura 4.1-(b) muestra el degree distribution de la red de Barabási-Albert. Se observa que la ecuación (2.5) (línea de color negra) coincide con los valores experimentales medidos para la red Barabasi-Albert (los puntos verdes de la misma gráfica).

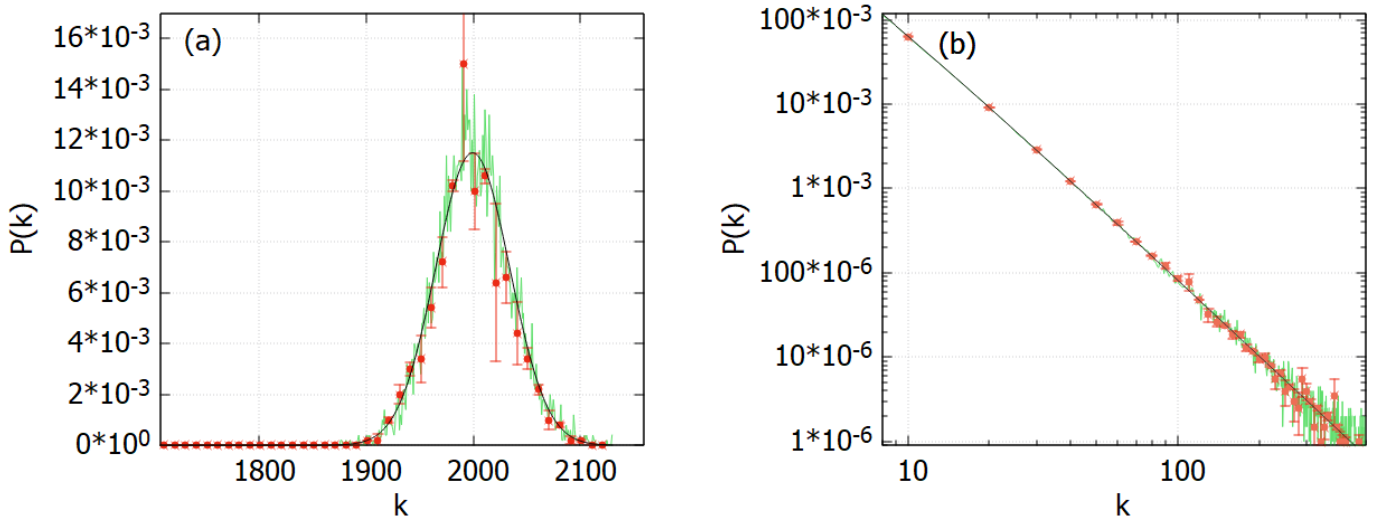


Figura 4.1: (a)-Degree distribution de una red aleatoria de 5 mil nodos. (b)-Degree distribution de una red de libre escala de $2,011 \times 10^6$ millones de nodos en escala logarítmica. (Los puntos rojos son valores experimentales, tomados cada 10 iteraciones, con sus respectivas barras de error)

Estas gráficas demuestran que las redes creadas cumplen con el $P(K)$ teórico correspondiente y validan, hasta cierto punto, los algoritmos construidos para generar estas redes.

4.2. Gillespie vs Runge Kutta

Con el fin de confirmar que ambos métodos: Gillespie (M.G.) y Runge Kutta (M.R.K), permiten obtener las soluciones numéricas a las ecuaciones del modelo SIR, se obtuvie-

ron las curvas de contagio utilizando los mismos parámetros para ambas metodologías. En el Anexo B.1 y B.2, se presenta la implementación que realizamos para el método Runge Kutta y el algoritmo de Gillespie respectivamente.

En la figura 4.2, se puede observar las curvas de $S(t)$, $I(t)$, $R(t)$ para una población con $N = 50 \times 10^3$ personas y con parámetros $\beta = 1,02 \times 10^{-5}$ y $\gamma = 0,04$. Dado que el propósito de este trabajo fue estudiar las variaciones entre las curvas de contagio cuando se consideran los medios de propagación, los valores de los parámetros β y α no son primordiales y se los tomó de manera aleatoria.

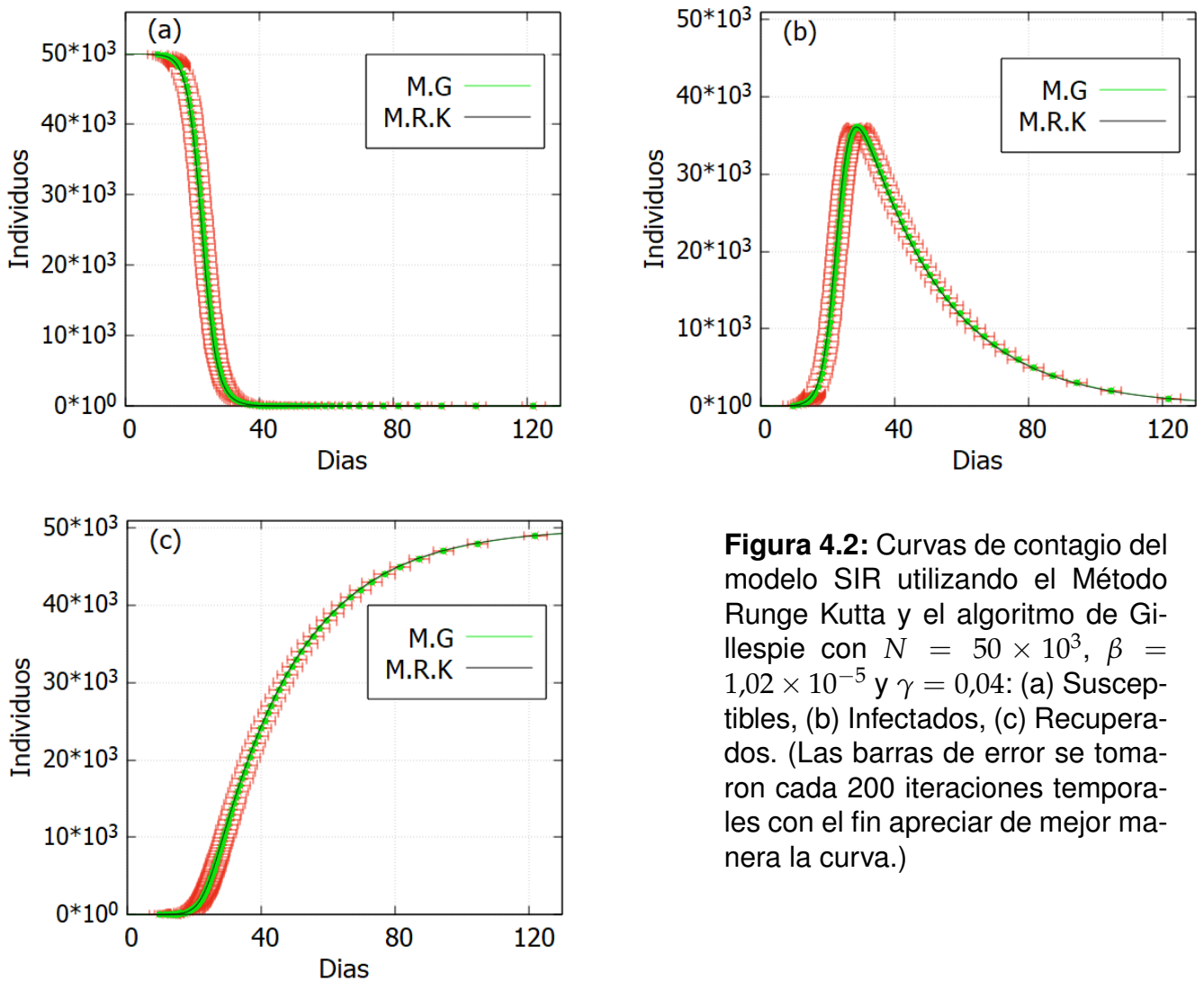


Figura 4.2: Curvas de contagio del modelo SIR utilizando el Método Runge Kutta y el algoritmo de Gillespie con $N = 50 \times 10^3$, $\beta = 1,02 \times 10^{-5}$ y $\gamma = 0,04$: (a) Susceptibles, (b) Infectados, (c) Recuperados. (Las barras de error se tomaron cada 200 iteraciones temporales con el fin apreciar de mejor manera la curva.)

Dado que el algoritmo de Gillespie es estocástico presenta una variabilidad alrededor de la media (representadas con las barras de error) y además permite tener datos

más realistas. Las barras de error se encuentran graficadas en x y y . Sin embargo, la variación en y es muy pequeña con respecto a la escala de la gráfica y no se logra apreciar. En toda esta sección, y las posteriores, todas las gráficas cumplen con esta característica.

En las curvas de contagio presentadas se puede observar como a medida que el número de infectados comienza a crecer, las personas susceptibles disminuyen. Cuando la curva de $I(t)$ empieza a disminuir, el número de individuos recuperados empieza a aumentar. Además, para las condiciones dadas, el número de infectados máximo se alcanza alrededor de los 30 días (de los 125 días que duró la propagación de la enfermedad). Esto indica que los individuos se contagian más rápido de lo que se recuperan.

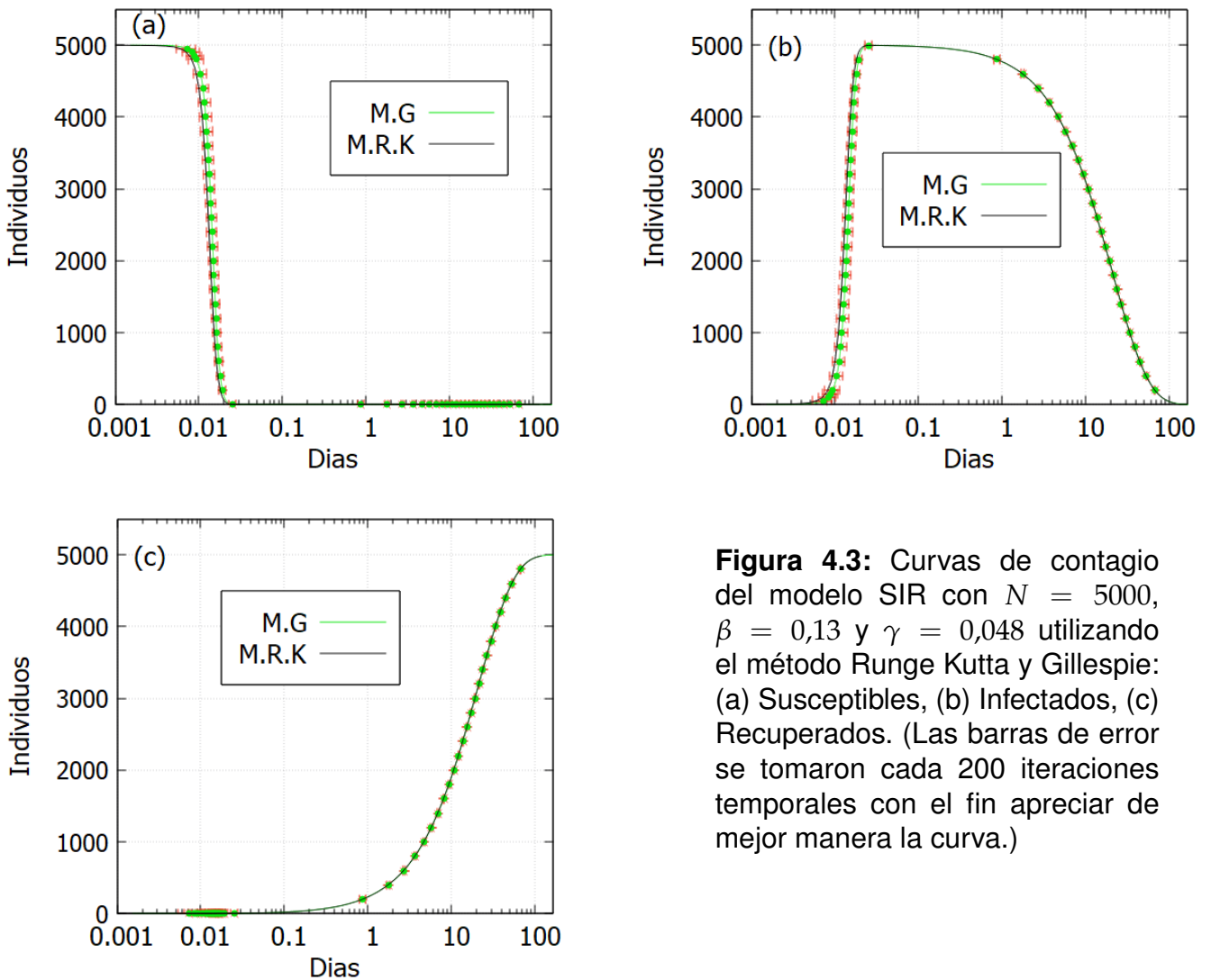


Figura 4.3: Curvas de contagio del modelo SIR con $N = 5000$, $\beta = 0,13$ y $\gamma = 0,048$ utilizando el método Runge Kutta y Gillespie: (a) Susceptibles, (b) Infectados, (c) Recuperados. (Las barras de error se tomaron cada 200 iteraciones temporales con el fin apreciar de mejor manera la curva.)

En la figura 4.3 se utilizaron los parámetros encontrados para el covid-19 en Texas, USA [48]: $\beta = 0,13$ y $\gamma = 0,048$ en una población de $N = 5000$. El tiempo se graficó en escala logarítmica para poder apreciar las curvas porque para estas condiciones las curvas de infectados y susceptibles cambian rápidamente.

Con los resultados obtenidos en esta sección se puede concluir que la modificación que realizamos al algoritmo de Gillespie para poder resolver un sistema de ecuaciones diferenciales es equivalente al método Runge-Kutta, con la ventaja de que Gillespie es capaz de captar la estocasticidad intrínseca del sistema. En este sentido, la implementación de nuestro algoritmo modificado (Anexo B.2) es correcta.

4.3. Curvas de contagio del modelo SIR sobre una red compleja

Las gráficas obtenidas en la sección anterior son una gran guía al estudiar propagación de enfermedades. Sin embargo, en este proyecto nos hemos propuesto estudiar el efecto de algunos patrones de contactos entre los individuos dentro de las ecuaciones del modelo SIR.

El algoritmo, detallado en el capítulo previo, se aplicó para tres clases de redes: completa, aleatorias y de libre escala.

4.3.1. Para la red completa

Como ya se ha comentado, que una población esté perfectamente mezclada significa que todas las personas están en contacto permanente en sí. En este sentido, se puede representar a una población con estas características como una red completa.

La figura 4.4 muestra las curvas que dan solución al modelo SIR sobre una red completa de $N = 5000$ nodos, $\beta = 1,02 \times 10^{-4}$ y $\gamma = 0,04$. Como se puede apreciar, la media de las curvas obtenidas coinciden con las curvas que se obtiene de Runge Kutta. Re-

cordemos que este método considera que la población se encuentra homogéneamente mezclada y comparar ambos métodos permite saber que efectivamente esta asociación con la red completa es acertada.

Debido a la capacidad del computador donde se realizaron las simulaciones no se logró obtener curvas para redes completas con más nodos, puesto que esto implicaría una cantidad gran cantidad de enlaces y, por tanto, un mayor almacenamiento computacional.

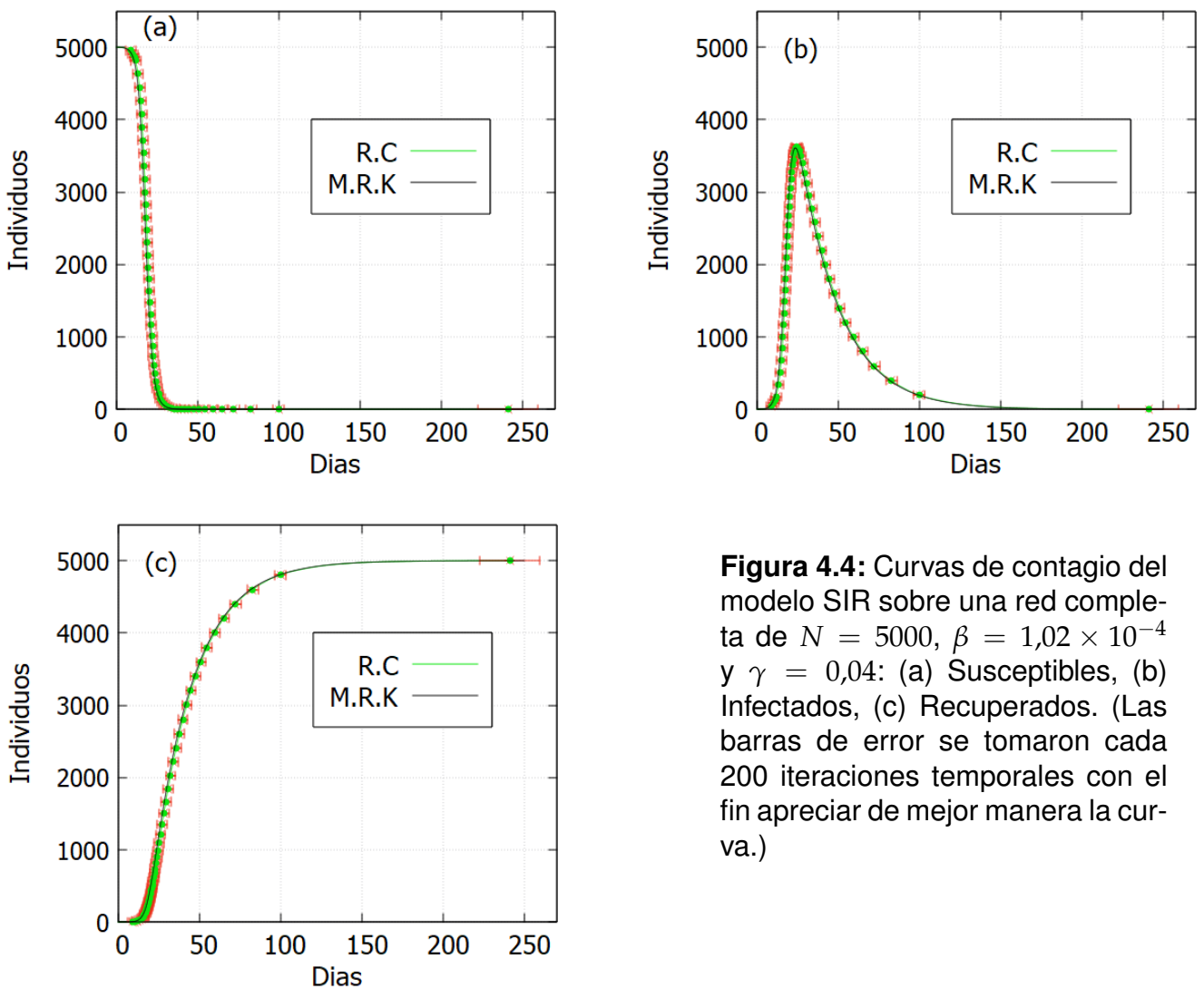


Figura 4.4: Curvas de contagio del modelo SIR sobre una red completa de $N = 5000$, $\beta = 1,02 \times 10^{-4}$ y $\gamma = 0,04$: (a) Susceptibles, (b) Infectados, (c) Recuperados. (Las barras de error se tomaron cada 200 iteraciones temporales con el fin apreciar de mejor manera la curva.)

4.3.2. Para la red aleatoria

Se crearon tres redes de $N = 5000$ nodos con diferente propabilidad de conexión: $p = 0,8$ (R.A.1), $p = 0,6$ (R.A.2.) y $p = 0,4$ (R.A.3). Además, también se utilizó la red completa, $p = 1$, que corresponde, como acabamos de ver, al modelo clásico con EDOs, para comparar como afecta la falta de contacto entre los individuos en la propagación de la enfermedad.

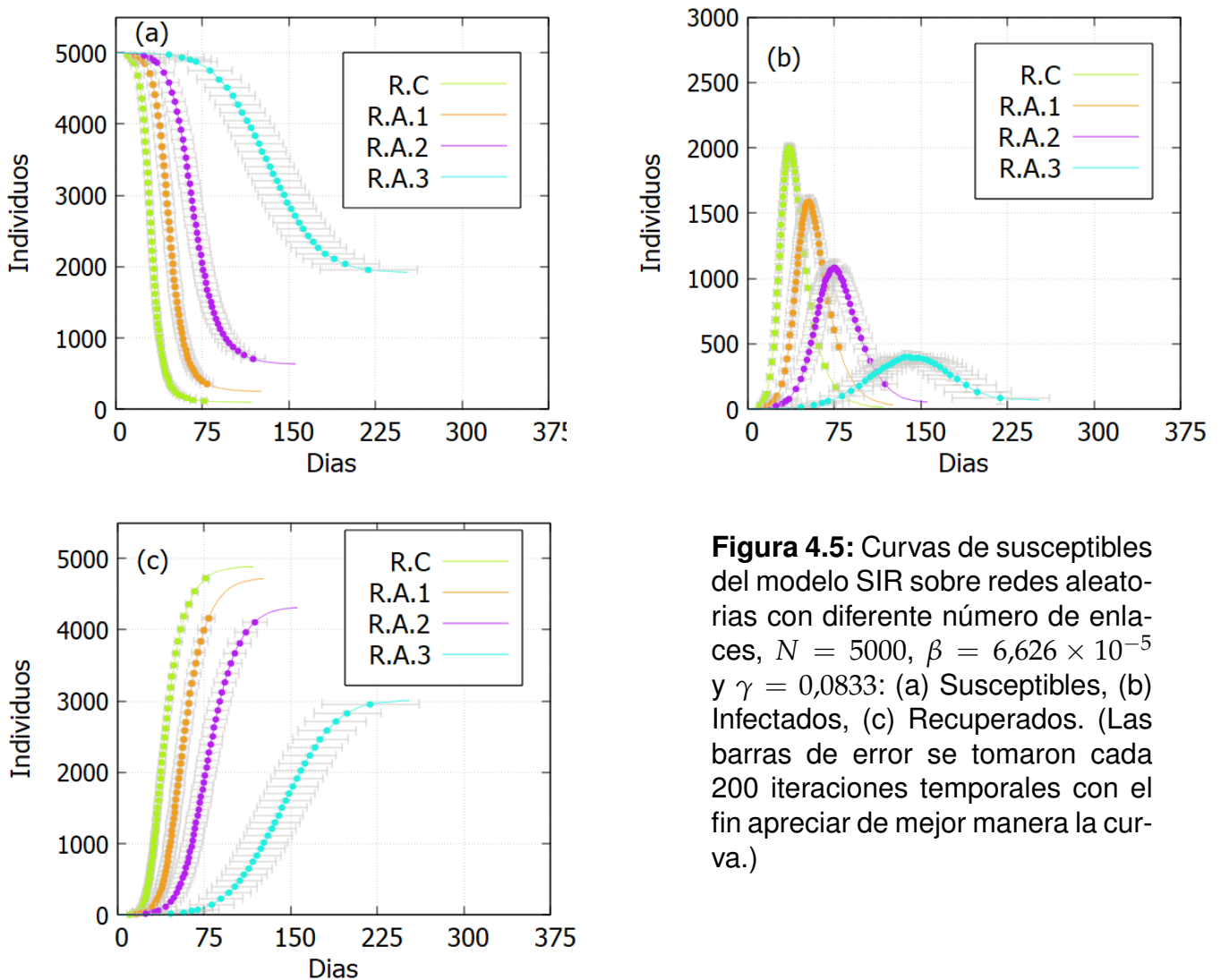


Figura 4.5: Curvas de susceptibles del modelo SIR sobre redes aleatorias con diferente número de enlaces, $N = 5000$, $\beta = 6,626 \times 10^{-5}$ y $\gamma = 0,0833$: (a) Susceptibles, (b) Infectados, (c) Recuperados. (Las barras de error se tomaron cada 200 iteraciones temporales con el fin apreciar de mejor manera la curva.)

La figura 4.5 muestra las curvas de contagio obtenidas para estas redes planteadas con los parámetros $\beta = 6,626 \times 10^{-5}$ y $\gamma = 0,0833$. Debido a que las simulaciones terminan cuando ya no hay individuos infectados, las curvas presentan diferentes termi-

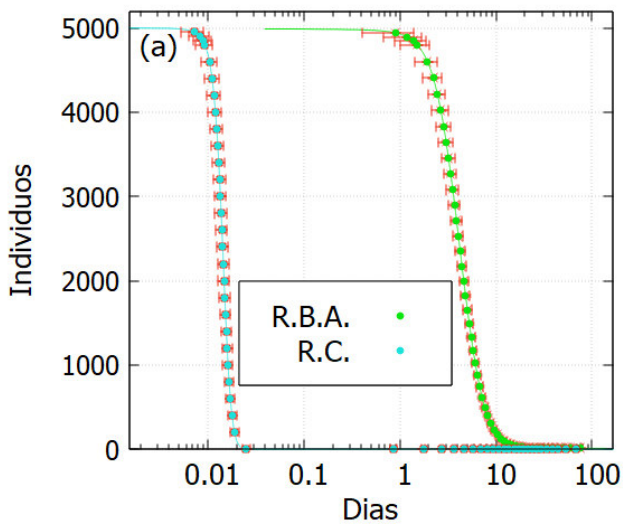
naciones, implicando que cuando la curva se acaba la propagación de la enfermedad también. Estas gráficas muestran como a medida que las redes tienen menos enlaces, las curvas se van desplazando hacia la derecha, indicando lo siguiente:

- A medida que la red posee más enlaces, el número de infectados máximo también aumenta. En otras palabras, cuantos más 'camino' de propagación de la enfermedad existan mayor será el número máximo de personas infectadas.
- Si la red tiene mayor número de enlaces, el número de infectados máximos se alcanzará en un período de tiempo más corto. Esto nos dice que, cuantos más contactos tengan las personas, más rápidamente aumenta el número de infectados. En otras palabras, la propagación del patógeno es más rápida.

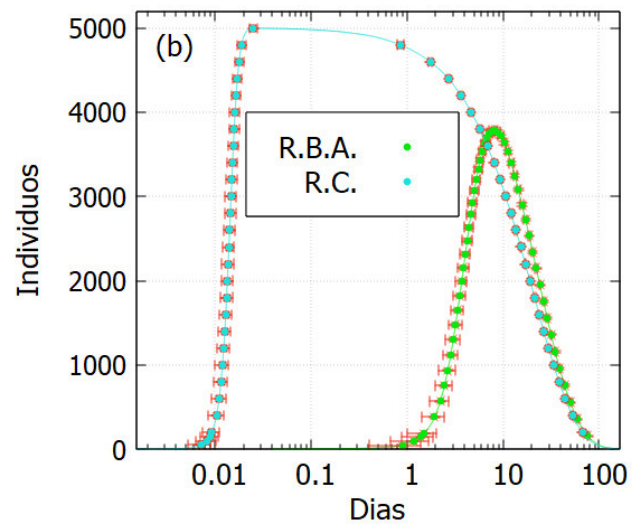
En consecuencia, se ha logrado ver que existe una dependencia del sistema al patrón de contactos específico que se tenga entre los individuos.

4.3.3. Para la red Barabási-Albert

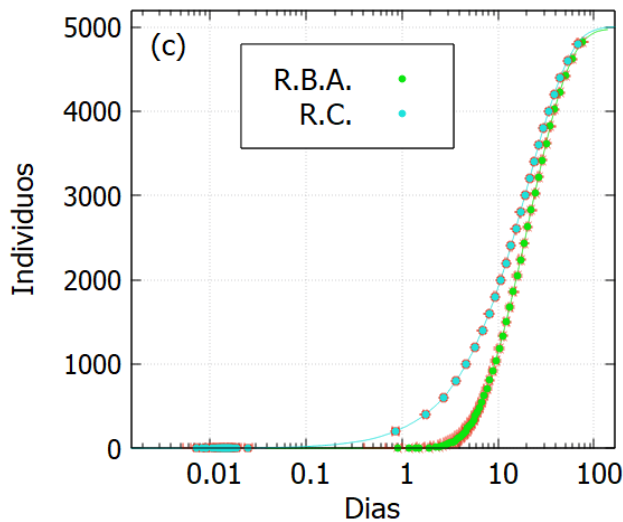
En la figura 4.6, se presentan las curvas $I(t)$, $S(t)$ y $R(t)$ para uno de los parámetros del coronavirus de Corea del Sur [48]: $\beta = 0,40$ y $\gamma = 0,035$ en una población de $N = 5000$ personas. El tiempo se graficó en escala logarítmica para poder apreciar el cambio existente en los resultados obtenidos para la red completa (R.C.), equivalente al modelo clásico con EDOs, y la red de Barabási-Albert (R.B.A.). Se observa como la curva de R.B.A. se desplaza hacia la derecha indicando la 'lentitud' de la propagación, debido a que en esta red existen menos 'camino' de propagación para la enfermedad (menos enlaces) y también de la particular estructura de la red de caminos (como están colocados los enlaces). Además, como podemos ver en la figura el número máximo de personas infectadas es menor en esta red. La mayor diferencia se logra ver en las curvas de $S(t)$ e $I(t)$, donde, para la red completa, las personas susceptibles se contagian en un período de tiempo mucho más corto y los individuos infectados aumentan rápidamente.



(a) Susceptibles



(b) Infectados



(c) Recuperados

Figura 4.6: Curvas de contagio del modelo SIR sobre una Red Barabasi con $N = 5000$, $\beta = 0,40$ y $\gamma = 0,035$: (a) Susceptibles, (b) Infectados, (c) Recuperados. (Las barras de error se tomaron cada 200 iteraciones temporales con el fin apreciar de mejor manera la curva.)

Al igual que las curvas obtenidas para la red aleatoria en la sección anterior, considerar una red de contactos entre las personas (en contraste a la hipótesis de "todos en contacto con todos") implica un cambio en el comportamiento de propagación de enfermedades y, por lo tanto, generan curvas de contagio diferentes al método con ecuaciones diferenciales clásico.

Como un dato curioso, la figura 4.7 muestra las curvas de contagio generadas sobre una red Barabási-Albert de $N = 2,011 \times 10^6$ nodos, población de Quito al año 2020, y con parámetros del covid-19 [48]: $\beta = 0,40$ y $\gamma = 0,035$. Se puede observar como,

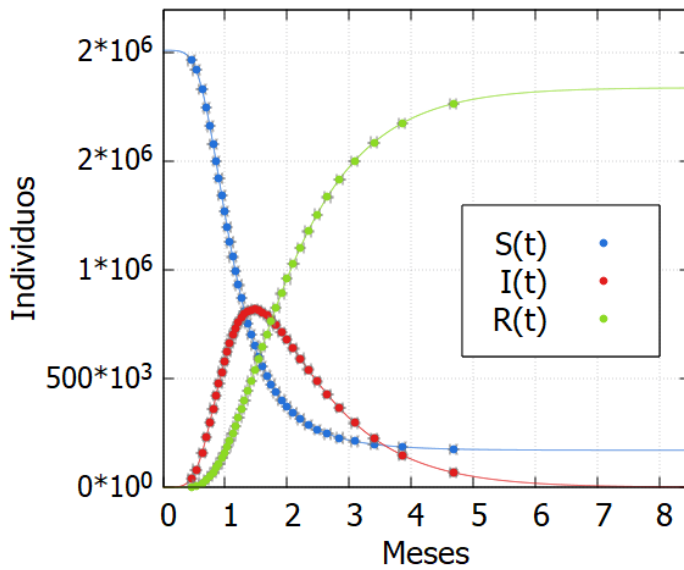


Figura 4.7: Curvas de contagio del modelo SIR sobre una Red Barabasi con $N = 2,011 \times 10^6$, $\beta = 0,40$ y $\gamma = 0,035$. (Las barras de error se tomaron cada 50×10^3 iteraciones temporales con el fin apreciar de mejor manera la curva.)

para estas condiciones, la pandemia alcanza el número máximo de infectados antes del segundo mes y finaliza alrededor del octavo. Incluso, existe una cierta cantidad de personas que no llega a contagiarse durante toda la propagación.

En esta gráfica no se presentan las curvas obtenidas con el método runge kutta (equivalente a una red completa) debido a que, como los resultados convergen al instante, se requiere mayor capacidad computacional.

Se intentó graficar las curvas de $I(t)$, $S(t)$ y $R(t)$ para las condiciones presentadas en la sección anterior para las redes aleatorias. Sin embargo, en la red Barabási-Albert la enfermedad no llega a propagarse bajo esas suposiciones. Esto implicaría que, para que una epidemia ocurra ya no dependería únicamente de β y γ , si no también de la interacción entre los individuos.

Como conclusión de todo este capítulo, nuestro estudio muestra la existencia de un cambio sustancial en las curvas de contagio generadas por el modelo epidemiológico SIR cuando consideramos una población perfectamente mezclada y cuando no lo está. Es decir, que para estudiar la propagación de enfermedades de forma más realista deberíamos tener en cuenta los 'caminos' de transmisión o la 'red de contactos' de los individuos.

Capítulo 5

Conclusiones

La hipótesis que se abordó en este trabajo de titulación fue saber si, al considerar las interacciones entre los individuos en la propagación de una enfermedad, se obtienen cambios sustanciales en las curvas de contagio obtenidas por los modelos clásicos donde la población que se considera se encuentra perfectamente mezclada.

A partir del algoritmo modificado de Gillespie se logró encontrar las curvas de contagio sobre una red completa, equivalente a una población homogénea, y sobre dos redes no completas: Aleatoria y Barabasi-Albert.

Logramos comprobar que las curvas de contagio sobre una red completa son equivalentes a las obtenidas por el método clásico de resolución de ecuaciones diferenciales, confirmando que el algoritmo Gillespie modificado para redes complejas funciona correctamente. Además, contestamos la pregunta planteada en el capítulo introductorio: ¿existen cambios sustanciales en las curvas de contagio si se consideran distintos patrones de contacto entre los individuos? La respuesta es sí.

Tanto para las redes aleatorias como para las redes de libre escala se comprobó que las curvas de contagio varían de manera significativa con respecto a los resultados obtenidos sobre redes completas. En el caso de las redes aleatorias, se observó como a medida que la red posee más enlaces (más medios de transmisión de la enfermedad) el número de infectados máximo aumenta y la propagación se dará de manera más rápida. Por otro lado, sobre la red de Barabási-Albert, a más de observar los mismos cambios que para una red aleatoria, se confirmó que para que una epidemia ocurra no

solo hay que considerar los parámetros que gobiernan las ecuaciones diferenciales, en este caso β y γ , si no también de la interacción entre individuos. Esto permite afirmar que considerar los medios de propagación durante una epidemia influye directamente en la dinámica de la misma.

En conclusión, nuestro estudio muestra que si queremos estudiar la propagación de enfermedades de forma realista deberíamos tener en cuenta la 'red de contactos' y descartar la hipótesis simplista de una población homogéneamente mezclada.

Apéndice A

Algoritmo

En este apéndice se presenta el algoritmo relacionado a la teoría expuesta en el Capítulo 2 sobre el método Runge Kutta.

A.1. Resolución por el método de Runge Kutta (RK4)

En esta sección se propone el siguiente algoritmo siguiendo el método de Runge Kutta para resolver las ecuaciones diferenciales del modelo epidemiológico SIR.

1. Definir el intervalo temporal a trabajar $t_0 \leq t \leq t_f$
2. Definir un vector con dimensión 3: $u_t = (s_t, i_t, r_t)$, donde estos corresponden a los valores iniciales de las variables de las ecuaciones *Aquí poner referencia a las ecuaciones.*
3. Dar valores iniciales para s_0, i_0, r_0 y de los parámetros β y α .
4. Definir un valor inicial para h .
5. Crear un vector con dimensión 3: $du_t = (ds_t, di_t, dr_t)$ de la siguiente manera:

$$du_1 = -\beta u_1 u_2$$

$$du_2 = \beta u_1 u_2 - \alpha u_2$$

$$du_3 = \alpha u_2$$

6. Encontrar los valores de k_i de las ecuaciones 2.9 - 2.12
7. Considerando la ecuación 2.8, se calcula u_t .
8. Se avanza temporal de la forma $t = t_0 + h * i$, donde i es la iteración temporal dada.
9. Repetir los pasos 5-8 hasta que $t = t_f$

Apéndice B

Códigos

B.1. Runge Kutta

```
subroutine deriv(du,u,x,n)!in: u y x , out:du
  implicit none
  integer::n
  real*8,dimension(n)::du,u! los elementos u_i son por ejemplo x,y,z y du
    contiene a las derivadas ex: dx/dt
  real*8::x
  real*8,parameter::g = 0.04d0, c = 0.51d0/5000.
  !aquí agregas las ecuaciones , por ejempl: dx/dt =x equivale a du(1)=u
    (1)

  du(1)=-c*u(1)*u(2)
  du(2)=c*u(1)*u(2)-g*u(2)
  du(3)=g*u(2)

end subroutine deriv
!****-sin paso adaptativo ****
subroutine rk4(y,dydx,n,x,h,yout)
  implicit none
  integer::n
  real*8::h,x
  real*8,dimension(n)::dydx,y,yout,dym,dyt,yt

  integer::i
  real*8::h6,hh,xh
  hh=h*0.5D0
  h6=h/6D0
  xh=x+hh

  do i=1,n
    yt(i)=y(i)+hh*dydx(i)
  enddo
  call deriv(dyt,yt,xh,n)
  do i=1,n
    yt(i)=y(i)+hh*dyt(i)
  enddo
  call deriv(dym,yt,xh,n)
```

```

do i=1,n
  yt(i)=y(i)+h*dym(i)
  dym(i)=dyt(i)+dym(i)
enddo
call deriv(dyt,yt,x+h,n)
do i=1,n
  yout(i)=y(i)+h6*(dydx(i)+dyt(i)+ 2D0*dym(i))
enddo
end subroutine rk4

```

B.2. Gillespie para resolución de EDOs

```

program gillespie
use,intrinsic :: ISO_Fortran_env
implicit none

integer*8 :: Nt, muu, i, j, condi, N
integer, allocatable, dimension(:) :: S, II, R
real*8, allocatable, dimension(:) :: T
integer*8, dimension(2) :: h, h_an
real*8, dimension(2) :: c
real*8 :: alpha, gama, dt, tau, mu, P, p0, P_ant, suma

Nt = 200000
N = 50000
allocate(S(Nt))
allocate(II(Nt))
allocate(R(Nt))
allocate(T(Nt))
S = 0
II = 0
R = 0

II(1) = 1
S(1) = N-II(1)
!R(1) = 0
c(2) = 0.04d0 !! parametro de recuperacion
c(1) = 0.51d0/dfloat(N) !! parametro de infeccion
condi = S(1)+II(1)+R(1)
T(1) = 0.
i = 1

! Inicializaci n

open(unit=10, file='SIR500mil_10.dat')

print*, N, c(1), c(2), condi

do while( i .lt. Nt .and. condi .eq. N)

  h(1) = S(i)*II(i)
  h(2) = II(i)

  suma = 0.

```

```

do j = 1, 2
    suma = suma + dfloat(h(j))*c(j)
enddo

call random_number(mu)

tau = (1./suma)*dlog(1./mu)

T(i+1) = T(i) + tau

if(i .eq. 1)print*, tau, mu

call random_number(mu)

if(i .eq. 1)print*, mu

if(mu*suma .le. dfloat(h(1))*c(1))then           !! Condicion para
    que alguien se infecte

    S(i+1) = S(i) - 1
    II(i+1) = II(i) + 1
    R(i+1) = R(i)
    write(10,*) T(i), S(i), II(i), R(i)
    i = i + 1

else

    if(mu*suma .gt. dfloat(h(1))*c(1) .and. mu*suma .le. suma)
    then

        S(i+1) = S(i)
        II(i+1) = II(i) - 1
        R(i+1) = R(i) + 1
        write(10,*) T(i), S(i), II(i), R(i)
        i = i + 1

    endif

endif

condi = S(i)+R(i)+II(i)

!print*, i, T(i), II(i)

if(II(i) .eq. 0)exit

end do

close(10)
close(99)

```

contains

B.3. Creación de la Red Aleatoria

```
program aleatoria
  implicit none

  integer :: n1, n2, Ne, N, i
  real :: x
  integer, dimension(:, :), allocatable :: aux, red
  integer, dimension(:), allocatable :: grado

  call initRandomSeed()

  N = 5000
  Ne = 50000*2
  i = 1

  allocate(aux(N,N))
  allocate(red(Ne,2))
  allocate(grado(N))

  aux(:, :) = 0

  p = 0.5

  do while (i .le. Ne)

    call random_number(x)
    n1 = 1 + floor(x*(N))
    call random_number(x)
    n2 = 1 + floor(x*(N))

    call random_number(r1)

    if(n1 .ne. n2 .and. (aux(n1,n2) .eq. 0 .and. aux(n2,n1) .eq. 0))
      then

        if( r1 .ge. p )then
          red(i,1) = n1
          red(i,2) = n2
          red(i+1,1) = n2
          red(i+1,2) = n1
          i = i + 2
          aux(n1,n2) = 1
          aux(n2,n1) = 1
        end if

      end if

    end do

  do i = 1, N

    grado(i) = sum(aux(i, :))
```

```

end do
end program

```

B.4. Creación de la Red Barabasi-Albert

```

program barabasi
  implicit none

  integer :: n0, c0, N, m, ct, i, j, ca, maxi, k
  real :: x
  integer, dimension(:, :), allocatable :: red
  integer, dimension(:), allocatable :: grado, aux
  real, dimension(:), allocatable :: proba, acum, degree, degreeth

  n0 = 10      !! nodos iniciales
  c0 = 90      !! enlaces iniciales
  N = 2011000  !! Nodos totales POBLACION DE QUITO 2020
  m = 4        !! numero de enlaces para los nodos a addos

  ct = c0 + 2*m*(N-n0)

  allocate(red(ct,2))    !! Le doy dimensi n a la red
  allocate(grado(N))     !! grado de cada nodo
  allocate(proba(N))     !! probabilidad para cada nodo
  allocate(acum(0:N))   ! Probabilidad acumulada
  allocate(aux(N))      !! Auxiliar para ver que no se repitan
                        enlaces

  open(unit=1, file='redcomplete10.dat')

  !! INICIALIZACION DE LA RED
  red = 0

  do i = 1, c0
      read(1,*) red(i,1), red(i,2)

  enddo

  close(1)

  print*, 'Lei la red'

  grado(:) = 0
  grado(1:10) = 9
  proba(:) = float(grado(:))/float(c0)    !! inicializacin
      probabilidad
  acum(:) = 0.
  ca = c0

  !! CRECIMIENTO DE LA RED

```

```

do i = n0+1, N

  print*, i
  aux(:) = 0
  acum(0) = 0.

  do j = 1, i-1

    acum(j) = acum(j-1) + proba(j)  !! Calculo de probabilidad
    acumulada

  end do

  do j = 1, m    !! Enlaces a a adir al nuevo nodo

    call random_number(x)

    do k = 1, i-1

      if((x .ge. acum(k-1) .and. x .lt. acum(k)) .and. aux(k) .
        eq. 0)then

        red(ca+1,1) = i
        red(ca+1,2) = k
        red(ca+2,1) = k
        red(ca+2,2) = i
        aux(k) = 1    !! nodo YA utilizado por el nuevo nodo
        ca = ca + 2
        grado(k) = grado(k) + 1    !! aumenta el grado del
        nodo viejo
        grado(i) = grado(i) + 1    !! aumenta el grado del
        nodo nuevo
        proba(:) = float(grado(:))/float(ca)    !!
        actualizaci n probabilidades

        exit

      end if

    end do

  end do

end do

!! Calculo DEGREE

maxi = grado(1)

do i = 1, N

  if(grado(i) .ge. maxi)then

    maxi = grado(i)

  endif

end do

```



```

        enddo

allocate (degree (maxi))
degree (:) = 0.

do i = 1, N

    degree (grado (i)) = degree (grado (i)) + 1.

end do

degree (:) = degree (:) / float (N)

allocate (degreeth (maxi))

do i = 1, maxi

    degreeth (i) = (2. * float (m) * float (m+1)) / (float (i) * float (i+1) * float (
        i+2))

end do

end program

```

B.5. Gillespie para resolución de EDOs sobre una red compleja

```

program gillered
use, intrinsic :: ISO_Fortran_env
use recablear

implicit none

integer :: i, j, k, auxi      ! contadores
integer :: n0, c0, N, m, ct, veci, nuevo, Nt      ! Variables
integer, allocatable, dimension (:, :) :: red, inter ! Red Barabasi-
    Albert, intervalo
integer, allocatable, dimension (:) :: grado2, degree2, S, II, R !
    ! vector que almacena el grado de cada nodo, probabilidades,
    degree sequense
real, allocatable, dimension (:) :: grado
real, allocatable, dimension (:) :: proba, proba_aux, degree,
    degree_th
real*8, allocatable, dimension (:) :: T ! Auxiliares
integer, allocatable, dimension (:) :: aux ! Auxiliares
real*8 :: pi, pr ! probabilidades de infectarse o recuperarse
real :: x, p, esperado, sumade, y      ! Auxiliares
integer :: ca, maxi, maxi2, u, auxii ! Auxiliares
real, allocatable, dimension (:, :) :: Eij, E

N = 5000      ! total de nodos
ct = 29984    ! RED BARABASI 5MIL

```

```

allocate(red(ct,2))    !! Le doy dimensi n a la red
allocate(grado(N))
allocate(grado2(N))
allocate(proba(N))
allocate(proba_aux(N))
allocate(aux(N))

open(unit=1, file='Barabasi5mil.dat')

red = 0

do i = 1, ct
    read(1,*) red(i,1), red(i,2)
enddo

close(1)

!! Calcular el grado de cada nodo
open(unit=11, file='grado5milbara.dat')

do i = 1, N
    read(11,*) grado(i)
enddo

!close(11)

allocate(inter(N,2))

call intervalo(N, grado, inter)

Nt = 100000

! CONDICIONES
!pi = 0.3313d0/dfloat(N)
!pr = 0.0833d0 !/dfloat(N)

allocate(S(Nt))
allocate(II(Nt))
allocate(R(Nt))
allocate(T(Nt))

call gillespie(N,ct,Nt,pi,pr,grado,red,inter, S,II,R,T)

end program albert

subroutine intervalo(M, grado, inter)
    implicit none

    integer, intent(in) :: M
    real, intent(in), dimension(M) :: grado
    integer, intent(out), dimension(M,2) :: inter

```

```

integer :: i

inter(1,1) = 1
inter(1,2) = grado(1)

do i = 2, M
    inter(i,1) = inter(i-1,2) + 1
    inter(i,2) = inter(i,1) + grado(i) - 1
enddo

end subroutine intervalo

subroutine gillespie(N,ct,Nt,beta,gama,degree, red, interv, S,I,R,T)
implicit none

integer, intent(in) :: N, ct, Nt !! Nt es las iteraciones temporales
real, dimension(N), intent(in) :: degree
integer, dimension(N,2) :: interv
real*8, intent(in) :: beta, gama !! parmetros
integer, dimension(ct,2), intent(in) :: red
integer, dimension(Nt), intent(out) :: S, I, R
real*8, dimension(Nt), intent(out) :: T

real*8 :: suma, tau, mu,y, pi, pr !! beta=pi*(numerodecontactos), gama
      =pr*I(t)
integer, dimension(N) :: estado ! 0 no infectado, 1 infectado, 2
      recuperado
integer, dimension(ct) :: estado_eje ! 0 y 1 de inter s
real*8, dimension(2) :: c
integer*8, dimension(2) :: h
integer, allocatable, dimension(:) :: nii
integer :: j, ne, condi, cti, ii, jj, k, ni, m, hh

! Esta subrutina me inicializa las semillas de forma que siempre sean
      distintas

S = 0
I = 0
R = 0

I(1) = 1
S(1) = N-I(1)

condi = S(1)+I(1)+R(1)
T(1) = 0.
j = 1

estado = 0
estado_eje = 0
ne = 0

allocate(nii(I(1))) ! nodos infectado

! ESTADO DE LOS PRIMEROS INFECTADOS Y LOS ENLACES DE INTER S

```

INFECTADO - SUSCEPTIBLE

```
do while ( j .le. I(1))

    call random_number(y)
    nii(j) = 1+floor(y*(N-1))

    do while(estado(nii(j)) .eq. 1)           !! Verifico que el nodo no
        est infectado
        call random_number(y)
        nii(j) = 1+floor(y*(N-1))
    end do

    estado(nii(j)) = 1

    do k = interv(nii(j),1), interv(nii(j),2)

        if(estado(red(k,2)) .eq. 0)then

            estado_eje(k) = 1

            ne = ne + 1

            do ii = interv(red(k,2),1), interv(red(k,2),2)

                if( red(ii,2) .eq. nii(j))then

                    estado_eje(ii) = 1

                end if

            end do

        else

            if(estado(red(k,2)) .eq. 1)then

                estado_eje(k) = 0    !! Si el NUEVO infectado est
                    conectado con otro infectado, hay que eliminar este
                    enlace
                ne = ne - 1           !! resto enlace

                do jj = interv(red(k,2),1), interv(red(k,2),2)

                    if(red(jj,2) .eq. nii(j))then

                        estado_eje(jj) = 0    !! enlace
                            bidireccional

                    end if

                end do

            end if

        end if

    end do

end if
```

```

        end if

    end do

    j = j + 1

enddo

! PRIMEROS ENLACES DE INTER S INFECTADO-SUSCEPTIBLE

j = 1

c(1) = beta
c(2) = gama

do while(j .lt. Nt .and. condi .eq. N)

    h(1) = ne
    h(2) = I(j)

    suma = 0.

    do k = 1, 2

        suma = suma + dfloat(h(k))*c(k)

    enddo

    call random_number(mu)
    tau = (1./suma)*dlog(1./mu)
    T(j+1) = T(j) + tau

    call random_number(mu)

    if(mu*suma .le. dfloat(h(1))*c(1))then !.and. S(j) .gt. 0)then
        !! Condicion para que alguien se infecte

        call random_number(y)
        cti = 1 + floor(y*(ct-1))

        do while(estado_eje(cti) .eq. 0)           !! Verifico que el
            enlace sea de inter s
            call random_number(y)
            cti = 1 + floor(y*(ct))
        end do

        S(j+1) = S(j) - 1
            I(j+1) = I(j) + 1
            R(j+1) = R(j)

            estado_eje(cti) = 0

            do k = interv(red(cti,2),1), interv(red(
                cti,2),2)

                if(red(k,2) .eq. red(cti,1))then

```

```

        estado_eje(k) = 0
    end if

end do

do k = 1, 2    !! Veo cual es el NODO que pasa de
              susceptible a infectado

    if(estado(red(cti,k)) .eq. 0)then

        estado(red(cti,k)) = 1    !! Cambio estado
                                   del nodo a infectado

        do ii = interv(red(cti,k),1), interv(red(cti,k),2)

            if(estado(red(ii,2)) .eq. 0)then

                estado_eje(ii) = 1    !! cambio es estado
                                       de los ejes que est n conectados con
                                       el nuevo nodo infectado
                ne = ne + 1            !! sumo este enlace

                do jj = interv(red(ii,2),1), interv(red(ii
                    ,2),2)

                    if(red(jj,2) .eq. red(cti,k))then

                        estado_eje(jj) = 1

                    end if

                end do

            else

                if(estado(red(ii,2)) .eq. 1)then

                    estado_eje(ii) = 0    !! Si el NUEVO
                                       infectado est conectado con otro
                                       infectado, hay que eliminar este
                                       enlace
                    ne = ne - 1            !! resto enlace

                    do jj = interv(red(ii,2),1), interv(
                        red(ii,2),2)

                        if(red(jj,2) .eq. red(cti,k))then

                            estado_eje(jj) = 0

                        end if

                    end do

                end if

            end if

        end if

    end if
end do

```

```

        end do

        end if

    end do

    j = j + 1

else

    if((mu*suma .gt. dfloat(h(1))*c(1)) .and. (mu*suma .le.
        suma))then

        call random_number(y)
        ni = 1 + floor(y*(N-1))

        do while(estado(ni) .eq. 0 .or. estado(ni) .eq. 2)  !!
            verifico que sea solo un nodo infectado
            call random_number(y)
            ni = 1 + floor(y*(N))
        end do

        S(j+1) = S(j)
        I(j+1) = I(j) - 1
        R(j+1) = R(j) + 1

        estado(ni) = 2          !! El nodo infectado se
            recupera

        do ii = interv(ni,1), interv(ni,2)

            if(estado(red(ii,2)) .eq. 0)then  !! Si estuvo
                en contacto con alg n susceptible se cambia el
                interes del eje

                ne = ne - 1  !! Quito un enlace
                estado_eje(ii) = 0  !! ya no es de interes

                do k = interv(red(ii,2),1), interv(red(ii,2)
                    ,2)  !! Quito tamb in bidireccionalidad

                    if(red(k,2) .eq. ni)then
                        estado_eje(k) = 0
                    end if

                end do

            end if

        end do

    end if

end do

j = j + 1  ! avance en el tiempo

```

```
                endif
            endif
            condi = S(j-1)+R(j-1)+I(j-1)
            if(I(j) .eq. 0) exit
        end do
    end subroutine
end module recablear
```


Bibliografía

- [1] María José Bianco, Pablo D Cruz, Alicia D Fraquelli, and Andrea L Gache. Modelo epidemiológico sir: Una aplicación de las ecuaciones diferenciales al sars-cov-2 (covid-19). *Revista de Investigación en Modelos Matemáticos aplicados a la Gestión y la Economía*, 7(1):16–38, 2020.
- [2] Antoni Virgili. La peste negra, la epidemia más mortífera. *National Geographic*, 2020.
- [3] Maria José Fresnadillo Martínez, Enrique García-Sánchez, Enrique García-Merino, Ángel Martín del Rey, and José Elías García-Sánchez. Modelización matemática de la propagación de enfermedades infecciosas: de dónde venimos y hacia dónde vamos. *Revista Española de Quimioterapia*, 26(2), 2013.
- [4] Ben R Craig, Tom Phelan, Jan-Peter Siedlarek, and Jared Steinberg. Improving epidemic modeling with networks. *Economic Commentary*, (2020-23), 2020.
- [5] Laura García Rovira et al. Modelos matemáticos compartimentales en epidemiología. 2017.
- [6] Mirjam Kretzschmar and Jacco Wallinga. Mathematical models in infectious disease epidemiology. In *Modern infectious disease epidemiology*, pages 209–221. Springer, 2009.
- [7] William Ogilvy Kermack and Anderson G McKendrick. A contribution to the mathematical theory of epidemics. *Proceedings of the royal society of london. Series A*,

- Containing papers of a mathematical and physical character*, 115(772):700–721, 1927.
- [8] Michael Y Li and James S Muldowney. Global stability for the seir model in epidemiology. *Mathematical biosciences*, 125(2):155–164, 1995.
- [9] Ricardo Almeida. Analysis of a fractional seir model with treatment. *Applied Mathematics Letters*, 84:56–62, 2018.
- [10] ROCIO MARIBEL AVILA AYALA. Inferencias de modelos epidemiológicos compartimentales en redes sociales. 2016.
- [11] Howard Howie Weiss. The sir model and the foundations of public health. *Materials matematics*, pages 0001–17, 2013.
- [12] Denia Morales Navarro. Riesgos y retos para los profesionales de las disciplinas estomatológicas ante la covid-19. *Revista Habanera de Ciencias Médicas*, 19(2), 2020.
- [13] Kavyu Mary Kamina, Samuel Mwalili, and Anthony Wanjoya. The modeling of a stochastic sir model for hiv/aids epidemic using gillespie’s algorithm. *Int. J. Data Sci. Anal*, 5(6):117, 2019.
- [14] Mark EJ Newman. The structure and function of complex networks. *SIAM review*, 45(2):167–256, 2003.
- [15] Stefano Boccaletti, Vito Latora, Yamir Moreno, Martin Chavez, and D-U Hwang. Complex networks: Structure and dynamics. *Physics reports*, 424(4-5):175–308, 2006.
- [16] Mark Ed Newman, Albert-László Ed Barabási, and Duncan J Watts. *The structure and dynamics of networks*. Princeton university press, 2006.
- [17] Maximino Aldana. Redes complejas. *Recuperado a partir de <http://www.fis.unam.mx/~max/English/notasredes.pdf>*, 2006.

- [18] Albert-László Barabási. Scale-free networks: a decade and beyond. *science*, 325(5939):412–413, 2009.
- [19] Albert-László Barabási, Erzsébet Ravasz, and Tamas Vicsek. Deterministic scale-free networks. *Physica A: Statistical Mechanics and its Applications*, 299(3-4):559–564, 2001.
- [20] Fabian Kuhn and Rotem Oshman. Dynamic networks: models and algorithms. *ACM SIGACT News*, 42(1):82–96, 2011.
- [21] Alexander S. Gillis. *Network Topology* .
- [22] Burcu Kantarci and Vincent Labatut. Classification of complex networks based on topological properties. In *2013 international conference on Cloud and Green Computing*, pages 297–304. IEEE, 2013.
- [23] Réka Albert and Albert-László Barabási. Statistical mechanics of complex networks. *Reviews of modern physics*, 74(1):47, 2002.
- [24] Albert-László Barabási. *NETWORK SCIENCE: RANDOM NETWORKS*. Capítulo 3. CC BY-NC-SA 2.0., 2003.
- [25] Paul Erdos, Alfréd Rényi, et al. On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci*, 5(1):17–60, 1960.
- [26] Paul Erdős and Alfréd Rényi. On the strength of connectedness of a random graph. *Acta Mathematica Hungarica*, 12(1):261–267, 1961.
- [27] Paul Erdos and Alfréd Rényi. Asymmetric graphs. *Acta Math. Acad. Sci. Hungar*, 14(295-315):3, 1963.
- [28] László Erdős, Friedrich Götze, and Alice Guionnet. Random matrices. *Oberwolfach Reports*, 16(4):3459–3527, 2020.
- [29] P Erdős and Alfréd Rényi. On the existence of a factor of degree one of a connected random graph. *Acta Math. Acad. Sci. Hungar*, 17:359–368, 1966.

- [30] P Erdős and A Rényi. On random matrices ii. *Studia Sci. Math. Hungar*, 3:459–464, 1968.
- [31] Albert-László Barabási. Network science. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 371(1987):20120375, 2013.
- [32] Edgar N Gilbert. Random graphs. *The Annals of Mathematical Statistics*, 30(4):1141–1144, 1959.
- [33] Albert-László Barabási. *NETWORK SCIENCE: THE BARABÁSI-ALBERT MODEL*. Capítulo 4. CC BY-NC-SA 2.0., 2013.
- [34] Béla Bollobás. Random graphs. In *Modern graph theory*, pages 215–252. Springer, 1998.
- [35] Albert-László Barabási. Scale-free networks: a decade and beyond. *science*, 325(5939):412–413, 2009.
- [36] Albert-László Barabási and Réka Albert. Emergence of scaling in random networks. *science*, 286(5439):509–512, 1999.
- [37] Dimiter Prodanov. Comments on some analytical and numerical aspects of the sir model. *Applied Mathematical Modelling*, 95:236–243, 2021.
- [38] Arieh Iserles. *A first course in the numerical analysis of differential equations*. Number 44. Cambridge university press, 2009.
- [39] Liancun Zheng and Xinxin Zhang. *Modeling and analysis of modern fluid problems*. Academic Press, 2017.
- [40] John C Butcher. On the implementation of implicit runge-kutta methods. *BIT Numerical Mathematics*, 16(3):237–240, 1976.
- [41] F VADILLO. Métodos de un paso. métodos runge-kutta.

- [42] Daniel T Gillespie. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *Journal of computational physics*, 22(4):403–434, 1976.
- [43] Naoki Masuda and Christian L Vestergaard. Gillespie algorithms for stochastic multiagent dynamics in populations and network. *arXiv preprint arXiv:2112.05293*, 2021.
- [44] Michael A Gibson and Jehoshua Bruck. Efficient exact stochastic simulation of chemical systems with many species and many channels. *The journal of physical chemistry A*, 104(9):1876–1889, 2000.
- [45] Daniel T Gillespie. Exact stochastic simulation of coupled chemical reactions. *The journal of physical chemistry*, 81(25):2340–2361, 1977.
- [46] Daniel T Gillespie. Approximate accelerated stochastic simulation of chemically reacting systems. *The Journal of chemical physics*, 115(4):1716–1733, 2001.
- [47] Joseph L Doob. Topics in the theory of markoff chains. *Transactions of the American Mathematical Society*, 52(1):37–64, 1942.
- [48] Ian Cooper, Argha Mondal, and Chris G Antonopoulos. A sir model assumption for the spread of covid-19 in different communities. *Chaos, Solitons & Fractals*, 139:110057, 2020.