



UNIVERSIDAD NACIONAL AUTÓNOMA
DE MÉXICO

FACULTAD DE CIENCIAS

**MECANISMOS DE DESARROLLO PARA REDES
CONSIDERANDO PROCESOS ESTOCÁSTICOS DE UN
PASO**

T E S I S

QUE PARA OBTENER EL TÍTULO DE:

FÍSICO

P R E S E N T A:

FRANCISCO MENDIOLA JERIA



DIRECTOR DE TESIS:
DR. GERMINAL COCHO GIL
DR. JOSÉ MATÍAS ALVARADO MENTADO

Cd. Universitaria, D. F. 2014



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

1. Datos del alumno

Mendiola
Jeria
Francisco
(55) 54792278
Universidad Nacional Autónoma de
México
Facultad de Ciencias
Física
408018218

2. Datos del tutor

Dr
Germinal
Cocho
Gil
Dr
José Matías
Alvarado
Mentado

3. Datos del sinodal 1

Dr
Denis
Pierre
Boyer

4. Datos del sinodal 2

Dra
Mariana
Benítez
Keinrad

5. Datos del sinodal 3

Dr
Germinal
Cocho
Gil

6. Datos del sinodal 4

Dr
José Matías
Alvarado
Mentado

7. Datos del sinodal 5

Dr
Pedro Eduardo
Miramontes
Vidal

8. Datos del trabajo escrito

Mecanismos de desarrollo para redes considerando procesos estocásticos de un paso
38 p.
2014

Contenido

1.	EL PROBLEMA Y LOS OBJETIVOS.....	4
2.	ANTECEDENTES.....	5
2.1.	Conceptos básicos de redes(6).....	5
2.2.	Redes complejas emblemáticas	7
2.3.	La ecuación maestra para modelar el desarrollo de una red compleja.....	8
3.	PROCESO DE ENLACE Y DESCONEXIÓN.....	11
3.1.	Elección de enlaces.....	11
3.2.	Condiciones de modelado y red inicial	12
3.3.	Eliminación de ciclos neutros	13
3.4.	Comparación con el algoritmo Gillespie	14
4.	DESARROLLO COMPUTACIONAL DE LA RED COMPLEJA	16
4.1.	Obtención de valores aleatorios y manejo de la red	17
4.2.	Rutina principal	20
4.3.	Las funciones Aleatorio y AleatorioLocal	23
4.4.	Una rutina diferente para la elección del nodo m.....	26
5.	ANÁLISIS DE RESULTADOS.....	27
5.1	Una rutina diferente para la elección del nodo m	28
5.2	Mecanismo de Enlace y Desconexión.....	29
	Topología de la red.....	32
6	DISCUSIÓN.....	33
7	CONCLUSIONES	36
8	REFERENCIAS.....	37

1. EL PROBLEMA Y LOS OBJETIVOS

Una red es un arreglo de nodos enlazados entre sí. Muchos sistemas observados en la naturaleza pueden describirse como redes; a menudo son llamadas complejas por tener un número grande de nodos, y presentar una estructura parcialmente ordenada y cambiante. Se denomina grado al número de conexiones que tiene un nodo. La distribución del grado de los nodos en una red compleja es fundamental para su caracterización y estudio. Se han distinguido (1) algunos tipos de distribuciones con distintos grados de heterogeneidad entre las conexiones de los nodos. Destacan, por su abundancia y comportamiento, las distribuciones llamadas libres de escala (2) que decaen como una ley de potencias. Existe gran interés por determinar los procesos –como homofilia y vinculación preferente– que originan este tipo de redes.

Por otro lado se ha observado (3) (4) que la descripción con una ley de potencias no es válida para todo el intervalo de los grados observados en numerosos casos de redes. En particular, existen discrepancias en los extremos del dicho intervalo. Se ha propuesto la Distribución Beta Generalizada (4) como alternativa a la ley de potencias. Además se ha presentado un modelo (5) que, mediante una ecuación maestra que toma a cuenta el nacimiento y muerte de enlaces (EMNM), describe una dinámica de cambios tal que, con expresiones simples para las probabilidades de enlace y desconexión, tiene como solución estacionaria una Distribución Beta Generalizada.

El objetivo de la presente tesis es, en primer lugar, la implementación de herramientas computacionales para la simulación de un mecanismo de desarrollo en redes que consta de la inserción y remoción iterada de enlaces en una red siguiendo un criterio dado. En segundo lugar, la investigación, de las condiciones bajo las cuales el desarrollo simulado produce una red cuya distribución del grado se ajuste a la solución estacionaria de la EMNM.

La red es descrita a través de una matriz booleana. Puede existir un solo enlace entre un par de nodos y no se permite que un nodo establezca enlaces consigo mismo. La implementación computacional de la red incluye rutinas que permiten manejar los cambios por inserción y remoción de enlaces, el conteo y control del grado de los nodos, el crecimiento de la red inicial y su posterior desarrollo. Se han considerado dos formas distintas para determinar los enlaces que se insertan o retiran. Sea $e(n, m)$ un enlace que conecta a los nodos n y m . El nodo n se elige con una probabilidad que es función de su grado k_n . El nodo m se elige aleatoriamente en el primer caso y con una probabilidad que depende de su grado k_m en el segundo. Se observa que para el segundo caso, la distribución del grado de la red generada se asemeja a la solución estacionaria de la EMNM.

2. ANTECEDENTES

Una red consta de un grupo de nodos enlazados entre sí. Es una abstracción útil para estudiar sistemas compuestos de muchas unidades vinculadas. Estos sistemas pueden ser físicos como los caminos dentro de un medio poroso, biológicos como las reacciones metabólicas, sociales como redes de amistad, entre otros. Son de particular interés las redes que, además de involucrar un gran número de elementos, presentan una estructura irregular, compleja y cambiante (6).

La investigación de redes complejas se intensificó desde que, en la última década del siglo pasado, se identificaron un par de características comunes a una gran variedad de redes observadas en la naturaleza. Mundo pequeño y distribución libre de escala es como se han nombrado y han motivado a muchos investigadores a buscar un cuerpo de conocimiento propio de las redes complejas en general.

2.1. Conceptos básicos de redes(6)

Una red $G(N, E)$ consta de un conjunto N de nodos y otro E de enlaces. Los enlaces son conexiones entre nodos. Así, si un enlace conecta a los nodos n_a y n_b tal que $n_a, n_b \in N$ se escribe $e(n_a, n_b) \in E$. Se dice que una red es no dirigida cuando los enlaces no tienen una dirección establecida. Los enlaces pueden ser simples, múltiples o tipo ciclo. Un enlace simple es una conexión entre un par de nodos. Tenemos un multienlace si existen varios enlaces simples entre un mismo par de nodos y un enlace tipo ciclo cuando tenemos un enlace simple entre un nodo y sí mismo. En adelante trabajaremos con redes cuyos enlaces son todos simples y no dirigidos.

Una subred es una red formada con un conjunto de nodos y enlaces pertenecientes a otra red. Se dice que un par de nodos son adyacentes si existe un enlace entre ellos. Una subred de vecinos se define como la sub red que incluye a todos los nodos adyacentes a un nodo en particular y los enlaces que los unen con el mismo y entre sí. La subred de vecinos para el nodo i se denota como G_i .

Una red puede representarse por medio de una Matriz de Adyacencia A de $N \times N$ donde N es el número de nodos existentes y el elemento A_{ij} es igual a uno si existe un enlace entre los nodos i y j o cero en caso contrario.

El número de conexiones k que tiene un nodo se denomina el grado del nodo, considerando la matriz de adyacencia correspondiente, se define así:

$$k_i = \sum_{j=0}^N a_{ij} \quad [1]$$

La distribución del grado $P(k)$ describe la probabilidad de que un nodo en la red, seleccionado aleatoriamente, resulte ser de grado k .

Se dice que una red con N nodos y K enlaces $G_{N,K}$ es completa si $K = N(N - 1)/2$. Es decir, cuando cada nodo está enlazado con todos los demás.

Un camino (*path*) entre los nodos i y j está definido como una secuencia entre nodos adyacentes tal que inicia en el nodo i , termina en el nodo j y cada nodo en la secuencia se visita una sola vez.

De esa manera, si existe un camino entre dos nodos, éstos están conectados aunque no compartan un enlace.

Cuando se piensa en transporte y comunicación a través de una red, es natural prestar interés al camino más corto entre dos nodos. El camino más corto promedio (L) es una característica de la red en conjunto y se define así:

$$L = \frac{1}{N(N-1)} \sum_{i \neq j} d_{ij} \quad [2]$$

Donde d_{ij} representa el camino más corto entre los nodos i y j . El valor máximo de los posibles d_{ij} es conocido como diámetro de la red.

Otra propiedad importante de las redes es la Transitividad o "Clustering", indica qué tan probable es que un par de nodos enlazados con un tercero estén enlazados entre sí. El coeficiente de "clustering" para el nodo i se define así:

$$c_i = \frac{2e_i}{k_i(k_i - 1)} \quad [3]$$

Donde e_i es el número de enlaces en la subred de vecinos del nodo i (G_i). El coeficiente de "clustering" de toda la red es el promedio del coeficiente para cada nodo, es decir:

$$C = \langle c \rangle = \frac{1}{N} \sum_i c_i \quad [4]$$

También son relevantes los motivos y las comunidades. Un motivo M es un arreglo específico de nodos y enlaces que se puede manifestar en la red bajo estudio G como una subred G' contenida en G . Los motivos son relevantes cuando se repiten en la red con una frecuencia extraordinaria. Para juzgar la relevancia de un motivo M en la red G , se compara el número de ocurrencias de M en la red G con el número de ocurrencias correspondientes en la red G_R que es equivalente a G en cuanto a número de nodos, enlaces y distribución del grado, pero que se ha generado de manera aleatoria. La importancia Z_M de M se define como sigue:

$$Z_M = \frac{n_M - \langle n_M^{rand} \rangle}{\sigma_{nM}^{rand}} \quad [5]$$

n_M es el número de ocurrencias de M en G , $\langle n_M^{rand} \rangle$ y σ_{nM}^{rand} son, respectivamente, el promedio y la desviación estandar de las ocurrencias de M para un ensamble de redes aleatorias G_{Ri} .

Dada una red $G(N, E)$, una comunidad es una subred $G'(N', E')$ cuyos nodos están densamente conectados. Hay varias formas de definir comunidad para casos concretos. Por ejemplo, una

comunidad tipo *n-clique* es una subred tal que la distancia de camino entre cualesquiera de sus nodos es, a lo sumo, de n enlaces.

Redes complejas: mundo pequeño y distribución libre de escala

La tecnología informática ha permitido aglomerar información en grandes bases de datos y analizarla de manera automatizada haciendo posible el estudio de grandes sistemas reales. Un resultado significativo es que, bajo el paradigma de las redes, existen características comunes a una gran variedad de sistemas distintos; las más sobresalientes son las denominadas Mundo pequeño y distribución del grado libre de escala.

Cuando la longitud de camino entre cualesquiera dos nodos es pequeña en relación al tamaño de la red, se dice que ésta presenta características de mundo pequeño. Dicha cualidad es provocada por la existencia de atajos o puentes entre distintas secciones de la red. La longitud de camino promedio L , y el coeficiente de “clustering” C caracterizan la propiedad de mundo pequeño; para que ésta exista la longitud del camino promedio debe ser relativamente pequeña y el coeficiente de acumulación ha de ser relativamente grande. Las redes que cuentan con la propiedad de mundo pequeño son muy eficientes para el transporte a escala global y local.

En las redes generadas de manera aleatoria donde la selección de los nodos a enlazarse es equiprobable, se observa que la distribución del grado es binomial. Pero la mayoría de las redes presentes en la naturaleza no siguen esa distribución. Muchas presentan una distribución del grado que, bajo una primera aproximación, sigue una ley de potencias, es decir:

$$P(k) \sim k^{-\gamma} \quad [6]$$

Con exponente γ variando, por lo general, entre 2 y 3. El nombre “libre de escala” se debe a que las leyes de potencia tienen la misma forma a cualquier escala i.e. $P(ck) \propto P(k)$. La distribución del grado en las redes libres de escala es altamente no homogénea; la diferencia entre los grados máximos y mínimos es muy grande en comparación con otras distribuciones.

Se ha observado (3) que el ajuste de los datos reales a una ley de potencias falla con frecuencia en los extremos de la distribución. De ahí el interés por encontrar modelos que sustenten la utilización de otro tipo de distribuciones.

2.2. Redes complejas emblemáticas

A continuación se comentan algunos casos en los que se aplica el conocimiento sobre redes. Redes sociales, internet, metabolismo y redes neuronales.

El paradigma de las redes ha sido utilizado para estudiar relaciones sociales desde principios del siglo 20. Típicamente, los nodos representan actores sociales que pueden ser individuos o agrupaciones, y los enlaces representan relaciones entre los actores sociales. Estos estudios han encontrado aplicaciones en políticas públicas y mercadológicas. El estudio de redes sociales ha contribuido con muchos conceptos a la teoría de redes en general. Por ejemplo el coeficiente de “clustering” y la centralidad de los nodos.

Internet es una red en la que las computadoras funcionan como nodos y los cables y/o conexiones inalámbricas satelitales que las conectan, como enlaces. Ésta red da lugar a la “World Wide Web” donde las páginas de internet son nodos y los hipervínculos son enlaces. Ambas redes han sido y siguen siendo estudiadas con gran interés. Es una tarea muy difícil por el gran tamaño y la naturaleza cambiante de dichas redes. Tanto internet como la World Wide Web presentan propiedad de mundo pequeño y distribución del grado libre de escala (6). Su estudio tiene aplicaciones tecnológicas inmediatas como el desarrollo de algoritmos de búsqueda, estrategias de mercado, y el diseño de estrategias de seguridad o para la optimización de recursos, etc.

Uno de los retos de la biología es describir no solo la función de los componentes celulares sino como se interconectan a través de una red de interacciones. Ello es importante porque algunas funciones de la célula no pueden explicarse considerando elementos aislados. Se han hecho, por ejemplo (6), estudios de las reacciones metabólicas donde los nodos representan a los metabolitos y los enlaces representan reacciones bioquímicas. Se observa que dichas redes presentan distribución del grado libre de escala y propiedades de mundo pequeño. Se espera que, al usar el paradigma de las redes se pueda relacionar la estructura de la red con las funciones del fenómeno que se estudia.

El estudio del cerebro ofrece campos muy interesantes para la herramienta de las redes. Los ensamblajes de las neuronas dentro del cerebro forman una red con características estructurales especiales y cambiantes. Por ejemplo, existen estudios (6) que buscan una relación entre la conectividad y la habilidad de cómputo en redes neuronales, otros buscan entender la relación entre la estructura y la función de un sistema nervioso, otros estudian mecanismos de desarrollo que dan como resultado redes neuronales similares las que se observan en la naturaleza, etc.

Es notorio como el paradigma de las redes se presta para el estudio de diversos fenómenos. Además, de la comparación entre los sistemas estudiados, aparecen algunos rasgos comunes como la propiedad de mundo pequeño y la distribución libre de escala para el grado. El presente trabajo pretende contribuir a la solución de una pregunta que surge de manera natural considerando los antecedentes. ¿Por qué razón el grado de los nodos se distribuye de forma especial?

2.3. La ecuación maestra para modelar el desarrollo de una red compleja

Como ya se ha mencionado, la ley de potencias describe la distribución del grado en muchos casos pero no se ajusta perfectamente a muchos otros. Se han propuesto explicaciones para dicho desajuste. Algunas consideran el tamaño de la red, otras consideran envejecimiento y costo de enlaces (3).

Existe una función que ha dado buenos resultados al ajustarse a una gran cantidad de datos ordenados por rango y que surgen de fenómenos de tipo social y biológico (4). Dicha función es una versión discreta de la Función Beta Generalizada (FBG), es decir:

$$f(r) = \frac{A(N + 1 - r)^b}{r^a} \quad [7]$$

Donde N es máximo valor que puede tomar r y A es una constante de normalización.

En esta sección se describe una dinámica de desarrollo propuesta por Alvarez-Martinez et al (5) que, bajo ciertas condiciones, deriva en una distribución del grado con forma similar a la FBG. Dicha dinámica está descrita mediante una ecuación maestra que considera el nacimiento y muerte de nodos con cierto grado k . La muerte de un nodo de grado k no implica la desaparición del nodo sino la ganancia o pérdida de un enlace conectado al mismo.

Un proceso estocástico es una función que depende de una variable aleatoria y del tiempo. Un proceso de Markov es un proceso estocástico tal que para cualquier conjunto de n tiempos sucesivos se cumple:

$$P_{1|n-1}(y_n, t_n | y_1, t_1; \dots; y_{n-1}, t_{n-1}) = P_{1|1}(y_n, t_n | y_{n-1}, t_{n-1}) \quad [8]$$

Es decir, la probabilidad de los eventos que ocurren en tiempo t_n queda determinada con el valor y_{n-1} en el tiempo t_{n-1} y no depende de los valores y_i a tiempos anteriores. Un proceso estocástico de un paso es un proceso de Markov a tiempo continuo cuyo rango pertenece a los números enteros y cuya matriz de transición solo permite brincos entre sitios adyacentes.

En el planteamiento de la ecuación maestra se asumen procesos estocásticos de un paso tales que, en el intervalo de tiempo $[t, t + dt)$, la probabilidad de que se realice una conexión o desconexión sobre algún nodo de grado k es $b(k)dt$ o $d(k)dt$ respectivamente. Entonces, el desarrollo del número de nodos $N(k, t)$ con grado k al tiempo t queda descrito por la ecuación:

$$\begin{aligned} \frac{\partial}{\partial t} N(k, t) = \{ \xi(k) - F[\sum(t)] \} N(k, t) \\ + \{ d(k+1)N(k+1, t) + b(k-1)N(k-1, t) \} - [d(k) + b(k)]N(k, t) \end{aligned} \quad [9]$$

Donde:

$\xi(k)$	Representa la tasa de crecimiento de la red
$\sum(t)$	Es el número total de nodos: $\sum(t) = \sum_k N(k, t)$
$F[\sum(t)]$	Es la tasa de decrecimiento global.

El primer término de la suma de la derecha se refiere al número de nodos de grado k que se ganan o pierden por por adición o sustracción de nodos. El segundo término de la suma considera a los nodos que, por adquirir o perder enlaces, se convierten en nodos de grado k .

El último termino considera nodos que, por ganar o perder enlaces, dejan de ser nodos de grado k . En adelante llamaremos Ecuación Maestra de Nacimiento y Muerte (EMNM) a la ecuación [9].

La densidad de probabilidad $n(k, t)$ de encontrar un nodo de grado k a un tiempo t se define como sigue:

$$n(k, t) = \frac{N(k, t)}{\sum(t)} \quad [10]$$

En adelante consideramos una población fija de nodos tal que $\{ \xi(k) - F[\sum(t)] \} = 0$. Sustituyendo $N(k, t) = \sum(t)n(k, t)$ en la ecuación [9] y tomando en cuenta que $\partial \sum(t) / \partial t = 0$ por la condición estática de la población de nodos, se obtiene:

$$\begin{aligned} \frac{\partial}{\partial t} n(k, t) = \{ d(k+1)n(k+1, t) + b(k-1)n(k-1, t) \} \\ - [d(k) + b(k)]n(k, t) \end{aligned} \quad [11]$$

El primer término de la suma de la derecha se refiere a la proporción de nodos que adquieren el grado k y el segundo a la proporción de nodos que pierden el grado k .

La solución estacionaria de la ecuación [11] es:

$$n^s(k) = \frac{b(k-1)b(k-2) \dots b(1)b(0)}{d(k)d(k-1) \dots d(2)d(1)} n^s(0) \quad [12]$$

Con:

$$n^s(0) = \left[1 + \sum_{k=1}^N \frac{b(0)b(1) \dots b(k-1)}{d(1)d(2) \dots d(k)} \right]^{-1} \quad [13]$$

Se puede tomar una expresión sencilla pero significativa para $d(k)$ y $b(k)$, por ejemplo:

$$d(k) = \lambda_1(C_1 + k)(N_1 - k) \quad [14]$$

$$b(k) = \lambda_2(C_2 + k)(N_2 - k) \quad [15]$$

Las ecuaciones [14]y [15] describen la probabilidad por unidad de tiempo de que desaparezca o aparezca un nodo de grado k en la red. La constante C_2 puede interpretarse como un umbral de activación para los nodos que no tienen ningún enlace, N_1 y N_2 son valores de saturación que se ajustan por retroalimentación y λ_1 , λ_2 son constantes relacionadas con la tasa de crecimiento.

Al sustituir las ecuaciones [14]y [15] en la ecuación [12]. Se obtiene que, para valores de k que cumplen $\bar{C} \ll k \ll \bar{N}$:

$$n^s(k) \sim N \frac{(\bar{N} - k)^{N_1 - N_2 - 1}}{(\bar{C} + k)^{C_1 - C_2 + 1}} \left(\frac{\lambda_1}{\lambda_2} \right)^k \quad [16]$$

Donde N es una constante de normalización. \bar{N} , \bar{C} se definen como sigue:

$$\bar{N} \equiv \frac{N_1 + N_2}{2} \text{ y } \bar{C} \equiv \frac{C_1 + C_2}{2} \quad [17]$$

En este trabajo se describe un programa computacional que simula el desarrollo de una red mediante un mecanismo que retira e inserta enlaces sobre una red de manera iterada y los elige utilizando las probabilidades $d(k)\Delta t$ y $b(k)\Delta t$. A tiempos grandes, la distribución del grado en la red simulada es similar a la solución estacionaria ecuación [16]. El resultado de la simulación se compara con la predicción del modelo y se discuten las condiciones bajo las cuales existe concordancia; también se comentan las limitaciones del mecanismo simulado en relación con las posibilidades planteadas por la EMNM.

3. PROCESO DE ENLACE Y DESCONEXIÓN

A continuación se describe el mecanismo y las condiciones consideradas para la simulación. El mecanismo de enlace y desconexión (MED) considera los siguientes pasos: 1. Retira un enlace de la red 2. Inserta un enlace a la red y 3. Repite desde el paso número 1. En la sección 3.1 se describe cómo se eligen los enlaces que se retiran e insertan.

El MED se aplica sobre una población estática N de nodos y K de enlaces, en la sección 3.2 se describen los factores considerados para determinar dichas cantidades.

El resultado de la aplicación iterada del MED es una red llamada *red final*, la característica que se observa de dicha red final es la distribución del grado para el caso estacionario es decir cuando, después de muchas iteraciones, deja de haber cambios significativos.

En la sección 3.3 se describe el manejo de las probabilidades de transición y la eliminación del tiempo entre reacciones. En la sección 3.4 se compara el algoritmo para la simulación del MED con el algoritmo Gillespie que se usa para simular procesos estocásticos para aplicaciones químicas.

3.1. Elección de enlaces

Los enlaces se definen a través que los nodos que conectan, por ejemplo, el enlace $e(1,2)$ conecta a los nodos n_1 y n_2 . Entonces elegir un enlace implica elegir un par de nodos.

Al retirar un enlace $e(n_d, m_d)$ entre los nodos n_d y m_d : Se elige que grado k_{d1} será desconectado con probabilidad $P_k = a_d d(k)$ con a una constante de normalización. Posteriormente se elige un nodo n_d de manera equiprobable entre todos los nodos que tienen grado k_{d1} . Se utiliza el mismo procedimiento para elegir m_d el segundo nodo que recibe una desconexión pero verificando que exista un enlace entre n_d y m_d antes de hacer cambios.

Al insertar un enlace $e(n_c, m_c)$ se sigue el mismo procedimiento, se elige el grado k_{c1} que será conectado con probabilidad $P_k = a_b b(k)$ y posteriormente un nodo n_c de manera equiprobable entre todos los nodos de grado k_{c1} . Se utiliza el mismo procedimiento para m_c el segundo nodo que recibe una conexión.

Entonces, en un paso de tiempo, la probabilidad de que un nodo s con grado k_s reciba un enlace es:

$$p_b(s) = \frac{b(k_s)}{e(k_s, t)} \quad [18]$$

Con $e(k, t)$ el número de nodos con grado k que están presentes en la red al tiempo t . La probabilidad de que un nodo s con grado k_s pierda un enlace es:

$$p_d(s) = \frac{d(k_s)}{e(k_s, t)} \quad [19]$$

3.2. Condiciones de modelado y red inicial

El MED se aplica sobre una red, en este caso se ha aplicado sobre una red sin multienlaces ni enlaces de tipo ciclo; es decir, no se permite que exista más de un enlace entre dos nodos y tampoco que un nodo se enlace consigo mismo. La red se describe mediante una matriz de adyacencias. La población de nodos N se mantiene constante y también la población de enlaces K . La condición estática de K se debe a que el MED inserta y retira un enlace a la vez y no tiene la capacidad de cambiar la proporción de los mismos. El hecho de que K sea constante tiene efectos sobre las funciones [14] y [15] pues permite que N_1 y N_2 tomen valores constantes. Por su parte, N_2 marca un límite para el grado a partir del cual la probabilidad de conexión por unidad de tiempo es cero; es decir, podemos encontrar en la red nodos con grado desde cero hasta $N_2 - 1$.

Las condiciones descritas en el párrafo anterior se reflejan en el modelado mediante la EMNM de la siguiente manera:

$\{\xi(k) - F[\Sigma(t)]\} = 0, \Sigma(t) = \text{constante}$	Porque N se mantiene constante.
$K = c_1, N_1 = c_2, N_2 = c_3$ con c_1, c_2, c_3 constantes	Porque K se mantiene constante.
$\sum_k b(k) = \sum_k d(k)$	Porque el MED hace conexiones y desconexiones en la misma proporción.

El número de enlaces K en la red inicial se ajustó para aprovechar al máximo el número de nodos N que participa en la simulación. La carga computacional y consecuentemente el tiempo que toma la simulación dependen fuertemente de la cantidad de nodos N , por eso hay especial interés en aprovechar todos los nodos simulados. Aprovechar los nodos simulados significa introducir en la red inicial un número de enlaces K que reduzca el número de nodos desconectados pero sin agotar la posibilidad de que exista la cantidad esperada de nodos para cada grado. Debe notarse que el MED es incapaz de regular el número de enlaces y por eso su rango de acción se reduce en una red muy saturada. Para estimar el número conveniente de enlaces K_c se ha utilizado la siguiente expresión:

$$K_c = \frac{N \sum_{i=1}^{N_2-1} i n^s(i)}{2} \quad [20]$$

Es decir, se usa la ecuación [16] y el número de nodos N para estimar el número de enlaces en una situación en la que la distribución del grado en la red final asemeja la solución estacionaria de la EMNM.

También se puede relacionar la población conveniente de nodos N_c con el parámetro N_2 y el número α de nodos con grado $N_2 - 1$ que se espera.

$$N_c = \frac{\alpha}{n^s(N_2 - 1)} \quad [21]$$

Es importante notar que las ecuaciones [20] y [21] se utilizaron como estimaciones informales para el número de nodos N_c y el número de enlaces K_c utilizados en la simulación, se establecieron sin conocimiento previo de cuál sería la distribución estacionaria del grado y no son relaciones significativas para el modelado mediante la EMNM.

La red inicial utilizada en las simulaciones se generó mediante la inserción de $K_c = 5000$ enlaces en una red formada por 1500 nodos. Los enlaces se insertaron de modo que la probabilidad de conexión por unidad de tiempo para un nodo s con grado k fue:

$$p(s, k, t) = \frac{b(k)}{e(k, t)} + \frac{1}{\sum(t)} \quad [22]$$

Con $e(k, t)$ el número de nodos de grado k presentes en la red al tiempo t . La probabilidad [22] es consecuencia de las rutinas mediante las cuales se realizaron las conexiones, dichas rutinas se eligieron para aprovechar fragmentos de código que ya estaban escritos y no tienen relevancia en el resultado final; es decir, la distribución del grado en la *red inicial* puede ser cualquiera y no afecta la distribución del grado en la *red final* cuando ésta refleja la distribución estacionaria.

La función $b(k)$ toma valores desde $k = 0$ y hasta $k = N_2 - 1$. Para la función $d(k)$ se considera que $d(0) = d(k > N_2 - 1) = 0$ ya que no tiene sentido la desconexión de nodos sin enlaces ni la desconexión de nodos con un grado que no puede existir en la red (no puede existir porque no nacen nodos con dicho grado).

Para los experimentos descritos en este documento se utilizaron los siguientes parámetros:

$$\begin{aligned} \lambda_1 &= 1, C_1 = 0.75, N_1 = 30 \\ \lambda_2 &= 1, C_2 = 1.25, N_2 = 28 \end{aligned} \quad [23]$$

Se han elegido esos parámetros porque al sustituirlos en la solución estacionaria $n^s(k)$ [16], dan como resultado una forma familiar es decir, una forma que se ha observado empíricamente (4). La similitud con las observaciones empíricas se da en términos de los coeficientes de la solución estacionaria: $a = C_1 - C_2 + 1 = \frac{1}{2}$ y $b = N_1 - N_2 - 1 = 1$. Existen varias combinaciones de parámetros que dan el mismo valor para los coeficientes a y b .

3.3. Eliminación de ciclos neutros

Ya que nuestro interés es observar las características de la red generada después de tiempos muy grandes (estado estacionario), se omiten los ciclos neutros, es decir, ciclos en los que no ocurre ningún cambio. Para forzar a que cada ciclo de lugar a una conexión o una desconexión se seleccionan todas las conexiones o desconexiones que son posibles en el estado presente de la red y se elige una considerando su probabilidad asociada. Es decir, se aplica una normalización en cada ciclo de modo que la suma de los eventos posibles sea 1.

Para ello se definen las funciones $b'(k)$ y $d'(k)$ que se actualizan en cada ciclo. El dominio de $b'(k)$ es el conjunto $E_b = \{k_i | \exists n_s \text{ con grado } k_i \text{ capaz de recibir una conexión}\}$. El dominio de $d'(k)$ es el conjunto $E_d = \{k_i | \exists n_s \text{ con grado } k_i \text{ capaz de perder una conexión}\}$. Los criterios que determinan si un nodo n_s puede recibir o perder una conexión cambian dependiendo del extremo del enlace que se está manipulando:

Inserción de enlace $e(n_c, m_c)$ primer nodo	La única condición sobre n_c es que exista en la red.
Inserción de enlace $e(n_c, m_c)$ segundo nodo	El nodo m_c debe existir en la red, ser distinto de n_c y además no estar enlazado a n_c .
Remoción de enlace $e(n_d, m_d)$ primer nodo	n_d debe existir en la red y tener algún enlace.

Remoción de enlace $e(n_d, m_d)$ segundo nodo	m_d debe existir en la red y compartir algún enlace con n_d .
---	---

Se establece la correspondencia:

$$d'(k_m) = n_1 d(k_m) \text{ si } k_m \in E_d \quad [24]$$

$$d'(k_m) = 0 \text{ si } k_m \text{ no pertenece a } E_d$$

$$b'(k_m) = n_2 b(k_m) \text{ si } k_m \in E_b \quad [25]$$

$$b'(k_m) = 0 \text{ si } k_m \text{ no pertenece a } E_b$$

Donde

$$n_1 = \frac{1}{\sum_{E_d} d(i)} \text{ y } n_2 = \frac{1}{\sum_{E_b} b(i)} \quad [26]$$

Al forzar un evento para cada ciclo, se pierde la correspondencia exacta entre el tiempo y los ciclos. Esta condición no afecta los resultados que se relacionan con la condición estacionaria, es decir para tiempos grandes o muchos ciclos.

3.4. Comparación con el algoritmo Gillespie

El algoritmo Gillespie es un algoritmo de simulación (7) estocástica ampliamente utilizado para simular reacciones químicas. Algunos elementos del MED pueden expresarse en términos del algoritmo Gillespie. A continuación se describe una posible implementación del algoritmo Gillespie para simular algunos elementos del MED. Luego se discuten las diferencias y similitudes con el programa realizado.

Tenemos k_i especies con abundancia $e(k_i, t)$. El índice $i = \{0, 1, \dots, N_2 - 1\}$. Hay cuatro reacciones que afectan la abundancia de cada k_i :

1. $k_i \rightarrow k_i - 1$ con probabilidad $d(k_i)\Delta t$
2. $k_i \rightarrow k_i + 1$ con probabilidad $b(k_i)\Delta t$
3. $k_i - 1 \rightarrow k_i$ con probabilidad $b(k_i - 1)\Delta t$
4. $k_i + 1 \rightarrow k_i$ con probabilidad $d(k_i + 1)\Delta t$

La cantidad de reacciones posibles (con cuidado de no contar alguna reacción dos veces) es de, a lo sumo, $2(N_2 - 1)$ y se asocian con las probabilidades $b(k_i)$ y $d(k_i)$. Por ejemplo, la reacción x implica que algún nodo de grado $k=10$ recibe una conexión; entonces la reacción x se asocia a la probabilidad $b(10)\Delta t$.

Pasos del algoritmo Gillespie:

g1. Generar números aleatorios r_1, r_2 .

g2. Computar $a_0 = \sum_i b(k_i) + \sum_i d(k_i)$ suma de la probabilidad de transición de todos los eventos posibles.

g3. Calcular el tiempo para la próxima reacción: $\tau = \frac{1}{a_0} \ln \left[\frac{1}{r_1} \right]$

g4. Computar la abundancia de k_i al tiempo $t + \tau$.

La abundancia de k_i al tiempo $t + \tau$ se computa dividiendo el intervalo $[0,1)$ con segmentos que representan cada tipo de reacción posible y cuya magnitud es proporcional a la probabilidad asociada. Si r_2 cae en el intervalo d_i , entonces ocurre la reacción asociada y se ajustan las cantidades $N(k, t)$ correspondientes. Los intervalos d_i se calculan así:

$$d_1 = \left[0, \frac{b(k_0)}{a_0}\right), d_2 = \left[\frac{b(k_0)}{a_0}, \frac{b(k_0) + b(k_1)}{a_0}\right), \dots, d_{2n} = \left[\frac{\sum_i b(k_i) + \sum_i^{n-1} b(k_i)}{a_0}, 1\right) \quad [27]$$

Y se relacionan con cada reacción de la siguiente manera:

d_1	Ocorre reacción asociada a $b(k_0) = b(0)$	$N(1, t + \tau) = N(1, t) + 1$ $N(0, t + \tau) = N(0, t) - 1$
d_2	Ocorre reacción asociada a $b(k_1) = b(1)$	$N(2, t + \tau) = N(2, t) + 1$ $N(1, t + \tau) = N(1, t) - 1$
...
d_{n+3}	Ocorre reacción asociada a $d(k_2) = d(2)$	$N(2, t + \tau) = N(2, t) - 1$ $N(1, t + \tau) = N(1, t) + 1$
...
d_{2n}	Ocorre reacción asociada a $d(k_n) = d(n)$	$N(n, t + \tau) = N(n, t) - 1$ $N(n - 1, t + \tau) = N(n - 1, t) + 1$

Previo a la ejecución del algoritmo se suministran las condiciones iniciales $N(i, 0)$.

El programa realizado es equivalente al algoritmo Gillespie en la selección de la secuencia de reacciones (pasos g2 y g4) pero no considera el tiempo entre cada reacción (paso g3). El algoritmo Gillespie computa en el paso g2 una constante de normalización a_0 que considera a las reacciones posibles y en el paso g4 la utiliza para dividir el intervalo $[0,1)$ en segmentos representativos para después elegir alguna reacción por medio de un número aleatorio. Ello es equivalente a la selección mediante las funciones $b'(k)$ y $d'(k)$ en el programa realizado. Sin embargo, en la simulación realizada no se considera el desarrollo temporal, es decir, no se calcula el tiempo τ para la siguiente reacción. Debe notarse que eso no altera la secuencia de las reacciones y por lo tanto la distribución a tiempos muy grandes es la misma se calcule o no el tiempo τ .

La virtud de la simulación realizada es que además de manejar la abundancia de cada especie k_i (de manera similar al algoritmo Gillespie), maneja la estructura de la red que resulta de las reacciones consideradas y gracias a ello, pueden medirse el diámetro, el coeficiente de clustering, el camino medio, etc.

4. DESARROLLO COMPUTACIONAL DE LA RED COMPLEJA

Se ha creado un programa en el lenguaje C++ que genera una red inicial mediante el procedimiento descrito en la sección 3.2 y posteriormente reproduce un *proceso de enlace y desconexión* en apego al *mecanismo* descrito en la sección 3.1.

Las entradas del programa son los parámetros descritos en [23] aunados a $N = 1500$ el número de nodos, $k = 5000$ el número de enlaces, y c el número de iteraciones del MED.

El parámetro N está asociado con el número de filas y columnas de la matriz de adyacencias, K está asociado con el número de iteraciones de la rutina que genera la red inicial, y c está asociado con el número de iteraciones de la rutina de enlace y desconexión.

En la sección 4.1 se describe el procedimiento que se aplicó para la obtención de valores aleatorios y también se comentan detalles sobre el manejo de la red en el entorno computacional.

Posteriormente, en las secciones 4.2 y 4.3, se describen los detalles del programa que implementa el MED y también un par de las subrutinas significativas. Finalmente, en la sección 4.4 se comenta la implementación de un mecanismo distinto al MED.

4.1. Obtención de valores aleatorios y manejo de la red

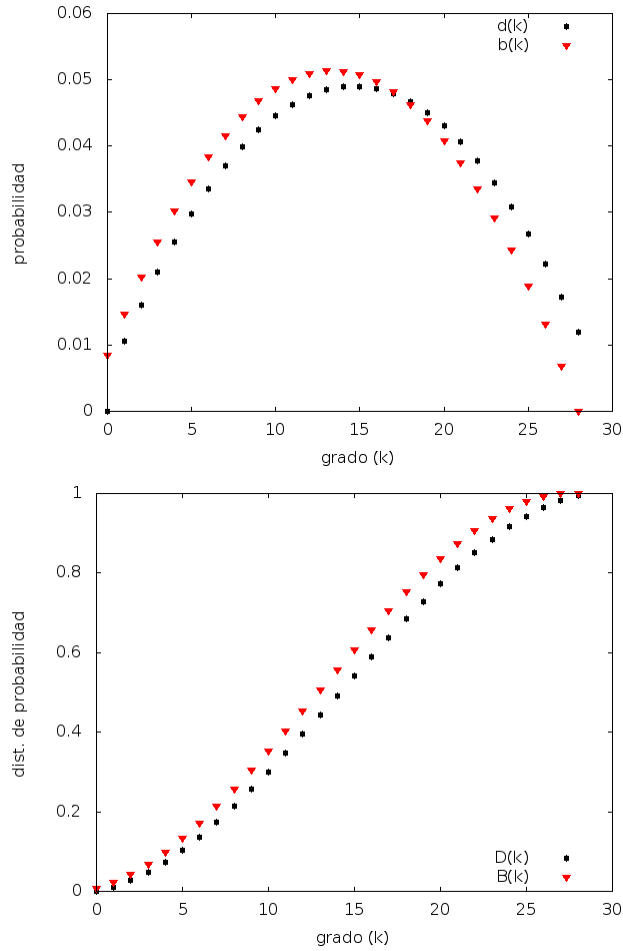


Figura 1: Arriba Las funciones de probabilidad $b(k)$ y $d(k)$. Abajo las funciones de distribución de probabilidad $B(k)$ y $D(k)$.

La “función de distribución acumulativa F ” de una variable aleatoria X es una función definida para cada número real x de la siguiente manera:

$$F(x) = \Pr(X \leq x) \quad [28]$$

Para una distribución discreta la probabilidad $\Pr(X \leq x)$ es la suma $\sum_i P(x_i)$ para $x_i \leq x$. Entonces, las funciones de distribución acumulativa $D(k)$ y $B(k)$ se definen así:

$$D(k) = \sum_{i=0}^k d(i) \text{ y } B(k) = \sum_{i=0}^k b(i) \quad [29]$$

El generador de números pseudo-aleatorios `rand()` produce una sucesión de números enteros que se distribuyen de manera uniforme en un intervalo entre cero y el valor `RAND_MAX`. El valor `RAND_MAX` depende de las librerías utilizadas por el compilador. Para los programas descritos

adelante, $RAND_MAX = 32\ 767$. Pueden generarse valores pseudo-aleatorios r en el intervalo $[0,1]$ definiendo:

$$r = \frac{rand()}{RAND_MAX} \quad [30]$$

Para generar valores particulares B^* y D^* de las variables aleatorias B y D , se utilizó un generador de números pseudo-aleatorios (función $rand$) y las funciones de distribución acumulativa $D(k)$ y $B(k)$.

En general, para generar valores aleatorios F^* a partir de una función de distribución acumulativa $F(x)$ con x que pertenece a un intervalo de los naturales, se puede utilizar la función AleatorioF descrita en la Figura 2.

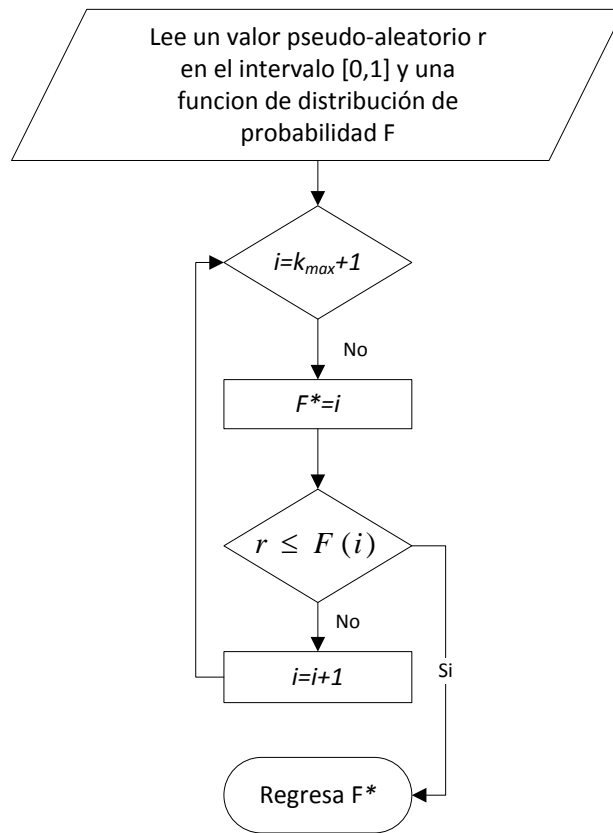


Figura 2: Función AleatorioF. Lee un número pseudo-aleatorio r en el intervalo $[0,1]$ y devuelve un valor aleatorio F^* .

En la Figura 3 se muestra una comparación entre la distribución de los valores aleatorios B^* generados por el programa y la distribución de probabilidad con la que fueron generados. Naturalmente, el ajuste entre la distribución y la función de probabilidad es mejor a mayor número de valores generados.

En el programa se usan expresiones similares a las indicadas en [29] pero que se restringen a los grados disponibles en cada ciclo. Si se ordenan los elementos de $E_b = \{e_0, \dots, e_n\}$ y los de E_d del mismo modo:

$$D'(k_i) = \sum_{j=e_0 \in E_a}^{e_i} d'(j) \text{ y } B'(k_i) = \sum_{j=e_0 \in E_b}^{e_i} b'(j) \quad [31]$$

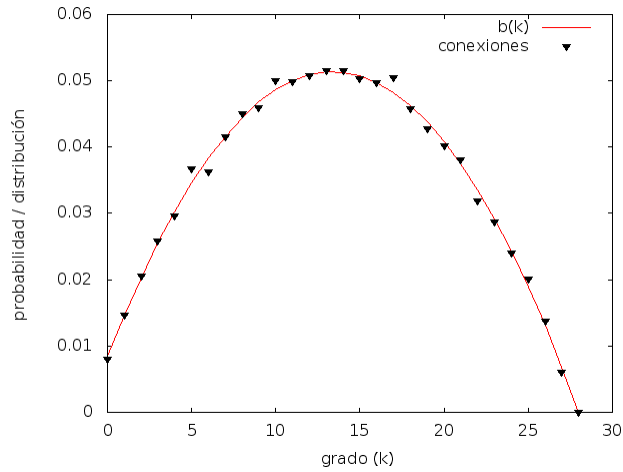


Figura 3: Comparación de la distribución 50 000 valores aleatorios B^* con la función $b(k)$.

Para seleccionar de manera equiprobable un elemento de una lista se utiliza la rutina descrita en la Figura 4. La posición p es un número entero entre 0 y $e - 1$ que puede usarse para señalar el número de pasos que ha de dar el iterador de la lista para acceder al elemento seleccionado.

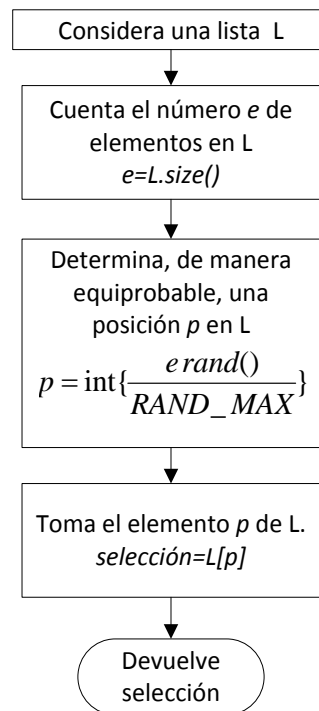


Figura 4: Rutina para seleccionar, de manera equiprobable, un elemento de una lista.

La forma más obvia de describir una red en el entorno computacional es utilizando una matriz de adyacencias. Puesto que no se permiten enlaces múltiples, dicha matriz puede ser de tipo booleano. Entonces, al inicio del programa se define una matriz A de $m \times m$ que describirá el estado de la red en todo momento.

Cada fila de la matriz representa a un nodo y los elementos no nulos en dicha fila representan sus enlaces. Por ejemplo, $A_{3,5} = 1$ indica que el nodo 3 tiene un enlace con el nodo 5. Entonces los nodos quedan etiquetados por el número de fila que les corresponde. En adelante denotamos a los nodos como n_i con $i = 1, 2, 3, \dots, N$.

Insertar un enlace entre los nodos n_a y n_b implica escribir un 1 en A_{n_a, n_b} y en A_{n_b, n_a} . Para eliminar un enlace entre los nodos n_a y n_b basta con hacer $A_{n_a, n_b}, A_{n_b, n_a} = 0$. Para agilizar el manejo es útil contar con un vector que lleve registro de los grados de cada nodo. Se define el vector de grados $G[n]$ tal que su n -ésimo elemento corresponde al grado del nodo n , es decir:

$$G[n] = \sum_{i=0}^m A_{n,i} \quad [32]$$

Para actualizar el vector $G[n]$ después de que el nodo n_a ha recibido un enlace basta con sumar +1 a la entrada correspondiente, es decir, $G[n_a] = G[n_a] + 1$. Si el nodo n_a ha perdido un enlace, el vector $G[n]$ se actualiza haciendo $G[n_a] = G[n_a] - 1$.

4.2. Rutina principal

En esta sección se describe la rutina principal del programa y los elementos que forman su contexto.

Al inicio del programa se inicializa la matriz de adyacencias A y su vector de grados $G[n]$. Posteriormente se genera una red inicial con K enlaces insertados mediante el procedimiento descrito en la sección 3.2. Luego se inicia la rutina principal que elige un enlace para eliminar y posteriormente inserta otro. La rutina principal se repite t veces.

El conjunto de grados disponibles E tal que $k_i \in E$ se determina cuatro veces en cada paso de tiempo y depende de los nodos disponibles para la conexión o desconexión en cada caso. La lista Lk guarda de manera ordenada y sin repeticiones a los grados disponibles, es decir $Lk = \{k \mid k \in E\}$.

Los pasos de la rutina principal, mediante la cual se retira un enlace $e(n_d, m_d)$ y se inserta otro $e(n_c, m_c)$, se muestran en la Figura 5.

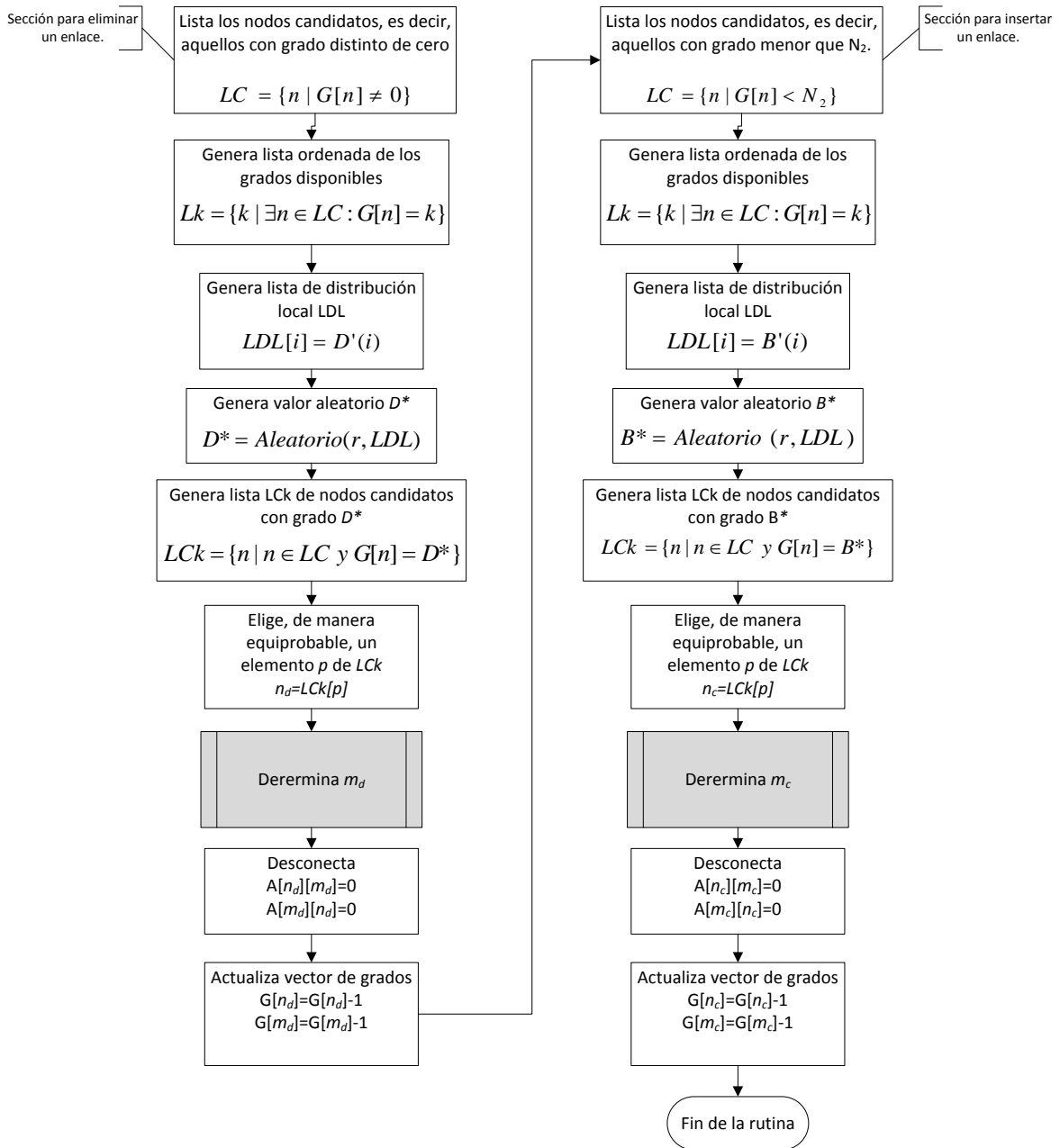


Figura 5: Rutina Principal

La lista LC almacena a los elementos del conjunto de nodos disponibles M . Para eliminar un enlace, la lista LC se llena con los nodos que tienen grado distinto de cero. Cuando se inserta un enlace, LC contiene a todos los nodos con grado menor que N_2 . Dicha restricción sobre el conjunto M podría parecer innecesaria considerando que, en una red generada por el procedimiento descrito en la sección 3.1, no pueden existir nodos con grado mayor que $N_2 - 1$. Sin embargo, en redes iniciales generadas de otra forma podrían existir nodos con grado mayor que $N_2 - 1$ y en ese caso la restricción sobre M sirve para trabajar solo con los nodos susceptibles a ser afectados por el proceso de enlace y desconexión. A partir de la lista LC se genera la lista de grados disponibles Lk .

La lista de distribución local LDL contiene los valores correspondientes a la función de distribución de probabilidad acumulada para cada paso de tiempo, es decir:

$$LDL[i] = B'(k_i) \text{ para conexiones y } LDL[i] = D'(k_i) \text{ para desconexiones} \quad [33]$$

Las listas de distribución local y, posteriormente, los valores aleatorios D^* y B^* , se generan mediante la función *AleatorioLocal* descrita en la sección 4.4.

En la sección de desconexión, el valor D^* indica que se tiene que eliminar un enlace a algún nodo de grado $k = D^*$. Para elegir a dicho nodo se crea la lista LCK que contiene a los nodos disponibles con grado D^* . Finalmente se elige de manera equiprobable un nodo n_d perteneciente a LCK .

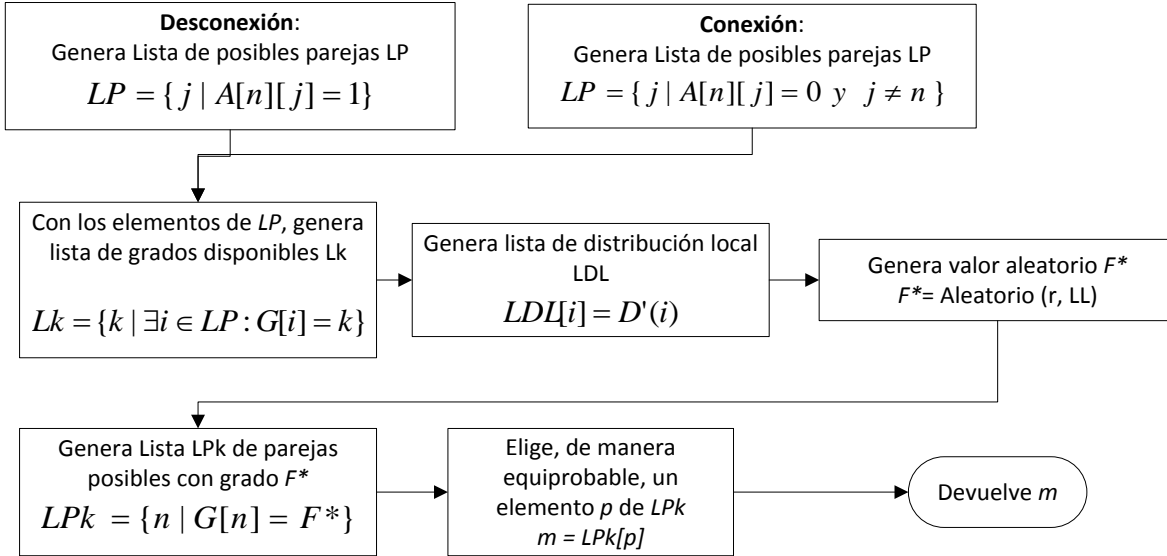
En la sección de conexión, el valor B^* indica que se tiene que insertar un enlace sobre algún nodo de grado $k = B^*$. Para elegir a dicho nodo se crea la lista LCK que contiene a los nodos disponibles con grado B^* . Finalmente se elige de manera equiprobable un nodo n_c perteneciente a LCK .

Para determinar m_d y m_c hay que considerar que ya se ha elegido n_d o n_c uno de los nodos al extremo del enlace que se elimina o inserta.

Para el caso de desconexión los nodos disponibles LP son todos aquellos enlazados a n_d . Para la conexión están disponibles todos los nodos que no están enlazados a n_c .

Con base en la lista de nodos disponibles LP se determina la lista de grados disponibles Lk . Posteriormente se genera una lista LDL con los valores $B'(k_i)$ o $D'(k_i)$ correspondientes. Haciendo uso de los números pseudo-aleatorios generados por la computadora (función $rand()$) y la lista LDL se determina el grado k que ha de tener el nodo seleccionado. Para generar la lista LDL y determinar un grado a partir de un valor pseudo-aleatorio se utilizan las funciones *Aleatorio* y *AleatorioLocal* que se describen a detalle en la siguiente sección.

Posteriormente se determina la lista LPk de nodos disponibles con grado k y se elige uno de sus elementos de manera equiprobable. El nodo seleccionado de esta manera corresponde a m_d o m_c según sea el caso.


 Figura 6: Rutina para la selección de m .

Una vez que se han seleccionado los nodos n_d y m_d , se retira el enlace $e(n_d, m_d)$. Para hacerlo basta con hacer $A_{n_d, m_d}, A_{m_d, n_d} = 0$. Para insertar un enlace $e(n_c, m_c)$, una vez que se han seleccionado los nodos n_c y m_c , basta con hacer $A_{n_c, m_c}, A_{m_c, n_c} = 1$.

4.3. Las funciones Aleatorio y AleatorioLocal

En esta sección se describen las funciones *Aleatorio* y *AleatorioLocal* que se utilizan en la rutina principal y en el mecanismo de enlace y desconexión que selecciona el nodo m con base en las funciones $b(k)$ y $d(k)$.

La función *AleatorioLocal*:

La función *AleatorioLocal* genera una lista de distribución local LDL y determina, usando la función *Aleatorio*, el grado k de un nodo a ser seleccionado con base en los nodos disponibles.

AleatorioLocal tiene dos argumentos de entrada, a saber: una función de probabilidad fp y una lista de nodos disponibles LD. La función de probabilidad fp se maneja como un arreglo que contiene los valores de $d(k)$ para la desconexión y $b(k)$ para la conexión, es decir:

$$fp[i] = b(i) \text{ o } fp[i] = d(i) \text{ con } i = 0, 1, 2, \dots, k_{max} \quad [34]$$

La lista de nodos disponibles LD es equivalente a LC en la rutina principal y a LP en los mecanismos de enlace y desconexión.

Primero se genera una lista LkD que guarda el grado de cada nodo listado en LD respetando el mismo orden. Posteriormente se determina la lista de grados disponibles Lk. La lista Lk contiene los mismos valores que LkD solo que ordenados y sin repeticiones. Luego comienza un ciclo que

incorpora elementos a la lista LDL. El número de elementos en LDL será igual al número de elementos en Lk y existe correspondencia en el orden de los mismos ya que el elemento en la posición i en LDL es el valor de la función de densidad de probabilidad (ecuación **¡Error! No se encuentra el origen de la referencia.**) para un grado k que es el elemento con posición i en la lista Lk.

El elemento i de LDL se obtiene sumando los valores correspondientes de la función de probabilidad para cada grado disponible k_i , es decir:

$$LDL[i] = \sum_{j=k_i}^{k_i} fp[j] \quad [35]$$

Al terminar el ciclo que genera los valores de la lista LDL se normaliza a cada uno de sus elementos dividiéndolos entre el último valor guardado en LDL. Se hace así porque LDL representa una función de densidad de probabilidad y el último de sus elementos debe ser igual a la unidad.

Se invoca a la función *Aleatorio* que devuelve un número que representa la posición p en Lk donde está guardado el valor del grado k del nodo que ha de seleccionarse. Finalmente se determina la lista LDk de los nodos disponibles con grado k y se selecciona uno de sus elementos de manera equiprobable. La función *AleatorioLocal* devuelve dicho elemento, es decir, un número que representa a uno de los nodos de la red.

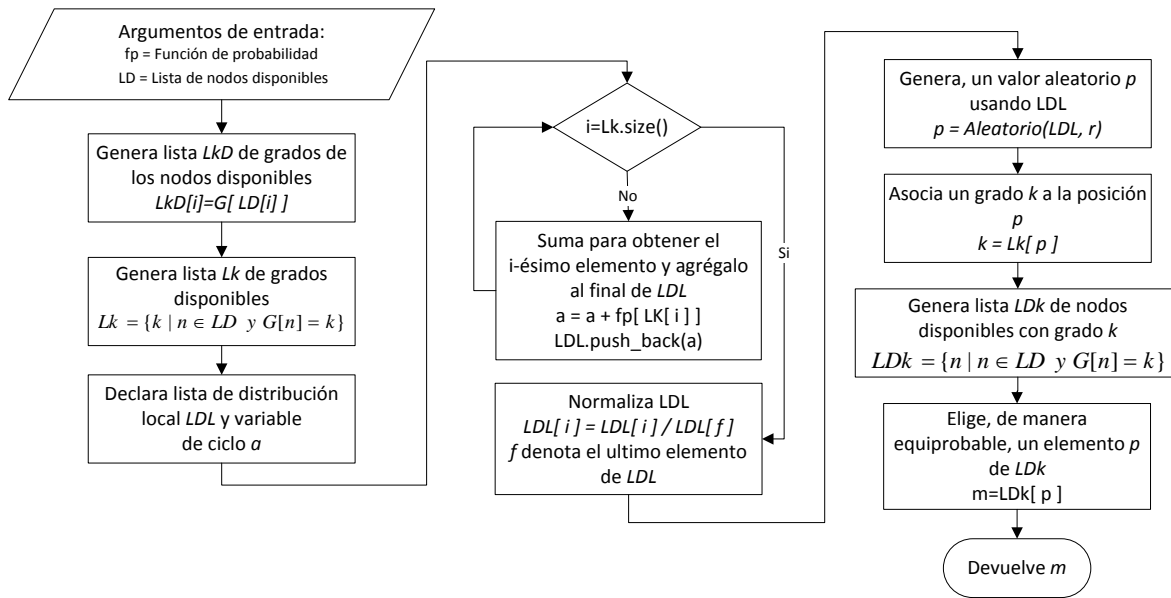


Figura 7: Diagrama para la función *AleatorioLocal*.

La función *Aleatorio*:

La función *Aleatorio* es la implementación de la función *AleatorioF* descrita en la sección 4.1. La función *AleatorioF* no se puede aplicar directamente por que en la práctica se presentan situaciones en las que el conjunto de grados disponibles es distinto al dominio de las funciones de distribución de probabilidad $D(k)$ y $B(k)$. Dicha situación se observa de manera muy clara cuando

se selecciona del nodo m_d con base en la función $d(k)$. En ese caso los nodos disponibles para la desconexión son aquellos que tienen un enlace con el nodo n_d , y los grados disponibles son solo los grados de dichos nodos.

La función Aleatorio tiene dos argumentos de entrada, un número r en el intervalo $[0,1]$ y una lista de distribución local LDL. El número r se genera previamente usando la función $\text{rand}()$ del lenguaje de programación como se muestra en la ecuación [30].

La lista de distribución local LDL se genera en la función *AleatorioLocal* y contiene los valores para una función de densidad de probabilidad como las descritas en la ecuación **¡Error! No se encuentra el origen de la referencia..**

Aleatorio compara, mediante un ciclo, el número r con cada valor listado en LDL. El ciclo se rompe si $r \leq LDL[i]$ para un elemento en la posición i de LDL. Se devuelve el valor de i que representa una posición en la lista de grados disponibles L_k que dio lugar a la lista de distribución LDL, es decir, $L_k[i] = k$ con k el grado del nodo que ha de elegirse.

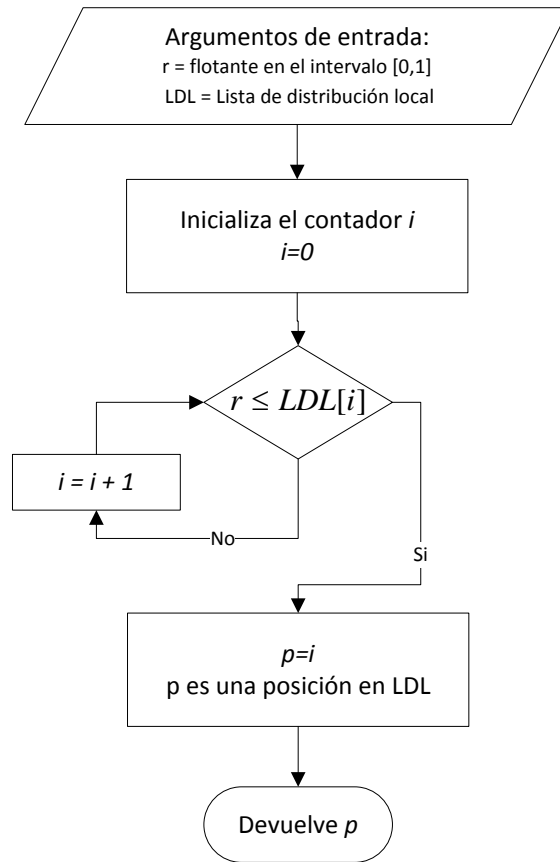


Figura 8: Diagrama para la función Aleatorio.

4.4 Una rutina diferente para la elección del nodo m

Originalmente se propuso un mecanismo tal que el segundo nodo de cada enlace, tanto para inserción como para remoción, se elige de manera equiprobable entre todos los nodos de la red. Aunque dicho mecanismo se descartó, se incluyen las rutinas y resultados relacionados con el solo objetivo de reseñar las experiencias de trabajo.

La rutina principal es la misma, la diferencia es la elección de los nodos m (recuadro gris en la figura 5). Primero se genera la lista LP de nodos disponibles. Cuando se pretende eliminar un enlace $e(n_d, m_d)$, los nodos disponibles para la elección de m_d son todos aquellos que tienen un enlace con n_d (n_d ya ha sido seleccionado en la rutina principal). Para insertar un enlace $e(n_c, m_c)$, los nodos disponibles para la elección de m_c son todos aquellos que no tienen un enlace con n_d pero sin contar a n_d . Se descarta a n_d pues no queremos que existan enlaces de tipo ciclo en la red.

Una vez que se ha generado la lista LP se elige, como m_d o m_c según corresponda, a uno de sus elementos de manera equiprobable.

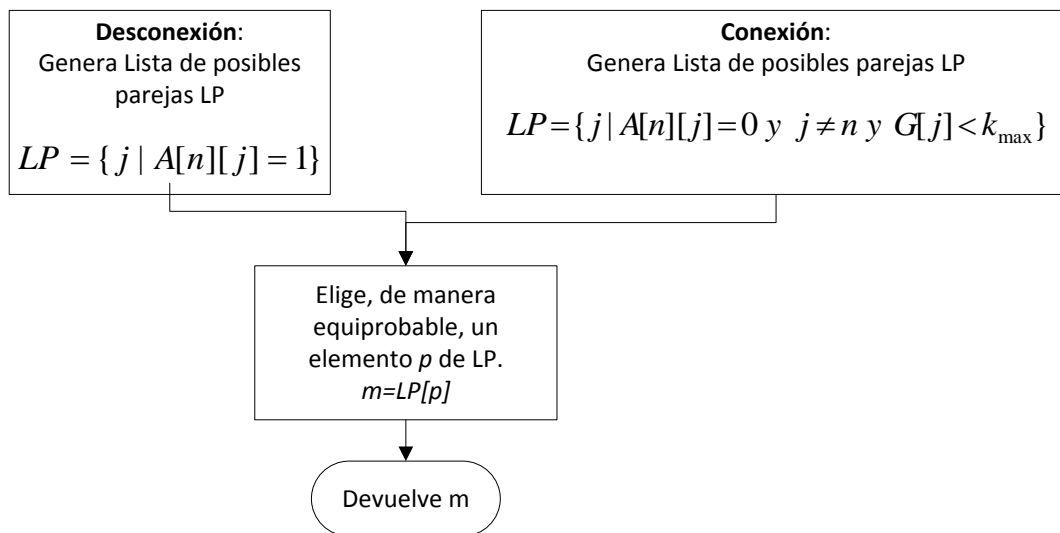


Figura 9: Rutina para la selección de m de manera equiprobable entre nodos.

5. ANÁLISIS DE RESULTADOS

La distribución del grado en la red cambia, partiendo desde la distribución propia de la *red inicial*, hasta alcanzar una distribución de equilibrio a raíz de la aplicación iterada del MED. El resultado final de la simulación es la descripción, a través de la matriz de adyacencias, de una red sobre la que se ha aplicado un proceso de enlace y desconexión un cierto número T de veces.

Se estudia la distribución del grado en la red final y el proceso mediante el cual se obtiene, en particular, el número total de ascensos del algún nodo con grado k a $k + 1$ y descensos de algún nodo con grado k a $k - 1$ para todos los grados a lo largo del proceso de iteraciones del MED.

Para obtener la distribución del grado de los nodos se cuenta el número de nodos con grado $i = 1, 2, \dots, k_{max}$ y posteriormente se normaliza. El conteo se realiza sobre el vector $G[n]$ que guarda el grado de cada nodo n , el resultado del conteo se normaliza dividiendo entre el número de nodos conectados, es decir, el número de nodos con grado distinto de cero.

Para obtener la distribución de ascensos a lo largo del proceso, se cuenta el número de veces que ocurre una conexión sobre un nodo de grado $i = 1, 2, \dots, N_2 - 1$ y se normaliza para cada valor de k . La distribución de las desconexiones se obtiene de la misma forma.

Los resultados que se muestran a continuación son producto de la aplicación del MED durante $T=100\ 000$ iteraciones sobre una red inicial con $N = 1500$ nodos y $K = 5000$ enlaces. Los parámetros de las funciones $b(k)$ y $d(k)$ son los mostrados en [23]. La red inicial se generó mediante el procedimiento descrito en la sección 3.2. La distribución del grado de una red inicial se muestra en la Figura 10 .

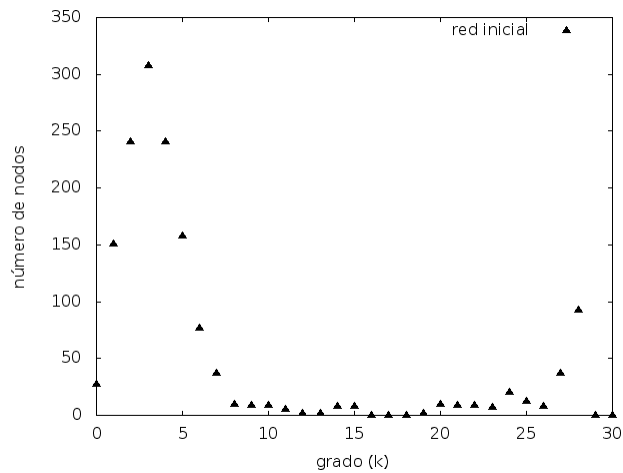


Figura 10: Distribución del grado en la red inicial.

5.1 Una rutina diferente para la elección del nodo m

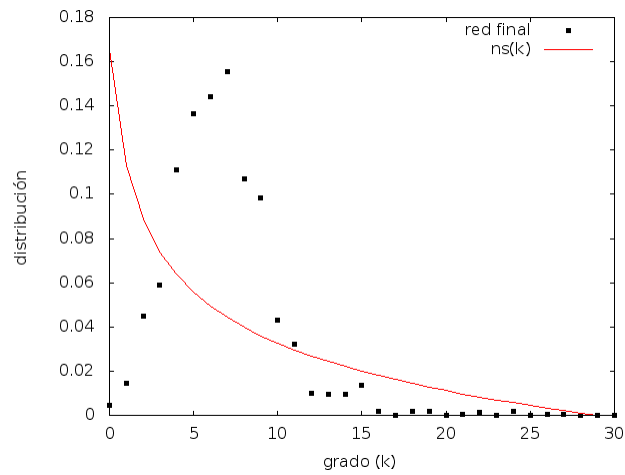


Figura 11: Distribución del grado en la red final después de aplicar un proceso de enlace y desconexión con elección de m equiprobable entre nodo. La solución a la ecuación maestra (ecuación [16]) se muestra en línea continua.

En la Figura 11 se muestra la distribución del grado después de aplicar el mecanismo de enlace y desconexión pero con selección de m equiprobable entre nodo. Se incluye, con línea continua, la ecuación [16] como referencia para comparar.

Se observa que la distribución del grado asemeja una distribución binomial en la parte central del intervalo. La distribución binomial es típica de (6) las redes aleatorias, es decir, en redes en las que se insertan enlaces sin algún criterio definido. Al igual que en las redes aleatorias, la sección que asemeja una distribución binomial es sensible al número de enlaces en la red y se desplaza hacia el máximo del intervalo del grado cuando se cuenta con un mayor número de enlaces.

5.2 Mecanismo de Enlace y Desconexión

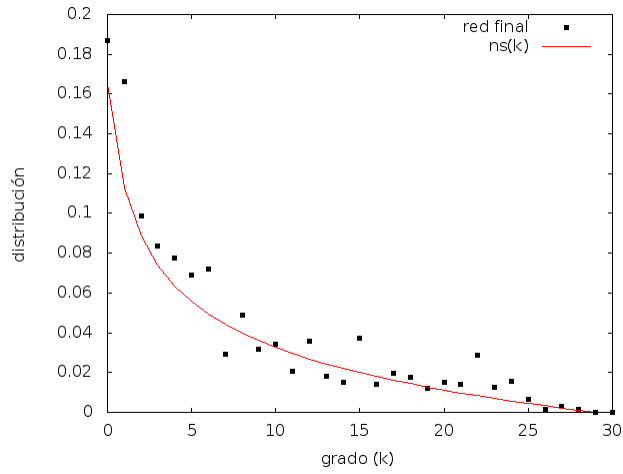


Figura 12: Distribución del grado en la red final después de aplicar un proceso de enlace y desconexión con elección de m usando $b(k)$ y $d(k)$. La solución a la ecuación maestra (ecuación [16]) se muestra en línea continua.

En la Figura 12 se observa que la distribución del grado en la red final se asemeja a la ecuación [16]. El parecido es más evidente cuando se observa (Figura 13) la distribución en escala logarítmica.

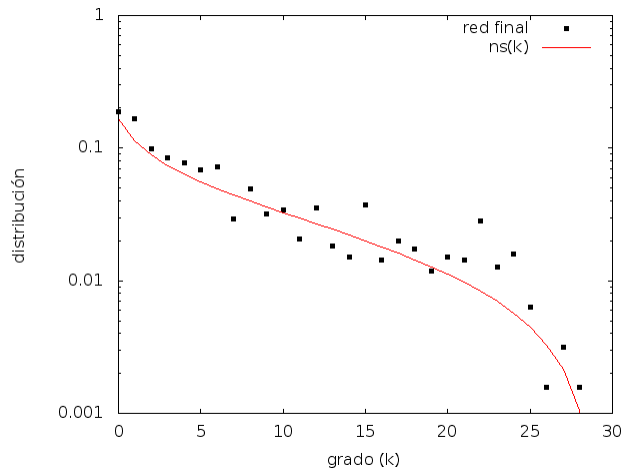


Figura 13: Distribución del grado en la red final en escala logarítmica después de aplicar un proceso de enlace y desconexión con elección de m usando $b(k)$ y $d(k)$. La solución a la ecuación maestra (ecuación [16]) se muestra en línea continua.

En la Figura 14 se observa que existe similitud entre la distribución de los ascensos y la función de probabilidad $b(k)$. También existe similitud (Figura 15) entre la distribución de los descensos y la función $d(k)$.

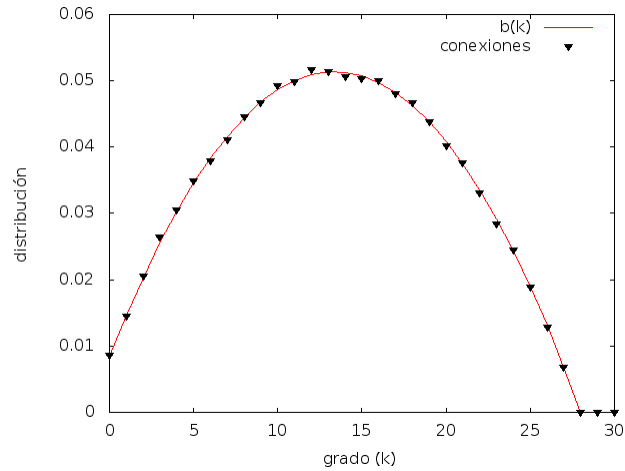


Figura 14: Distribución de conexiones a lo largo del proceso de enlace y desconexión. La función de probabilidad $b(k)$ se muestra en línea continua.

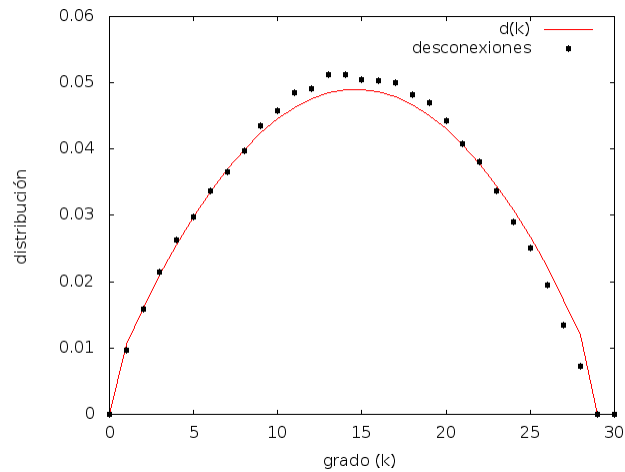


Figura 15: Distribución de desconexiones a lo largo del proceso de enlace y desconexión. La función de probabilidad $D(k)$ se muestra en línea continua.

La línea que representa la solución $n^S(k)$ en la Figura 12 y Figura 13 se obtiene sustituyendo en la solución de la ecuación maestra [16] los parámetros utilizados para la generación de los valores aleatorios [23] y normalizando los resultados.

Se hizo un promedio con veinte juegos de datos de la distribución del grado tomados cada mil iteraciones una vez alcanzada la distribución estacionaria; el resultado se muestra en la Figura 16.

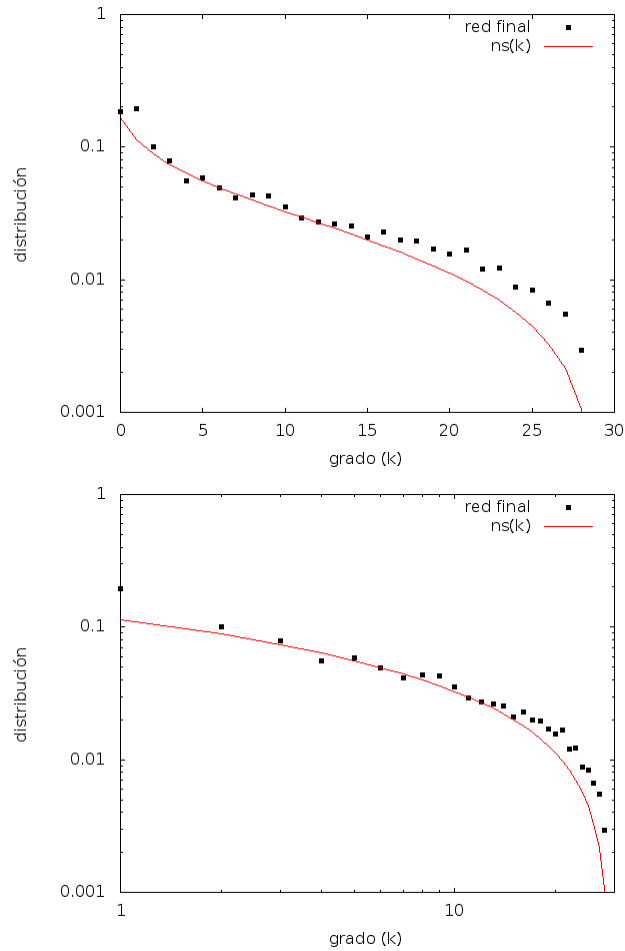


Figura 16: Promedio de los valores de 20 distribuciones del grado en la red final. Arriba en escala semi-log y abajo en escala log-log. La solución estacionaria $n^s(k)$ se muestra con línea continua.

Se observa que la caída, al final de la distribución, no es tan pronunciada como la solución estacionaria $n^s(k)$.

La expresión para la solución estacionaria descrita por la ecuación [16] es una aproximación para valores de k tal que $\bar{C} \ll k \ll \bar{N}$. Esa condición podría cumplirse mejor si k pudiera tomar valores en un intervalo mayor, sin embargo ello implica aumentar el número de nodos y por consiguiente el tiempo que tarda la simulación.

El MED tiene algunas características que escapan del modelo si se utiliza la EMNM. Primero: el MED impone un orden fijo de eventos, es decir, dos desconexiones para retirar un enlace y dos conexiones para insertar otro.

Segundo: La probabilidad de elección de m_d el segundo nodo recibe una influencia extra de la estructura de la red ya que está condicionado a ser un vecino del primer nodo para desconexión n_d .

Tercero: La selección de nodos realizada por el MED es comparable con la EMNM solo si k puede tomar un número grande de valores (en la simulación realizada, k puede tomar $N_2 - 1$ valores), es

decir, si $N_2 - 1 \rightarrow \infty$. Sea $p(s, k, t)$ la probabilidad de que el nodo s tenga grado k al tiempo t . La situación para el nodo s al tiempo $t + dt$ se puede expresar así:

$$p(s, k, t + dt) = \frac{b(k-1)dt}{e(k-1, t)} p(s, k-1, t) + \frac{d(k+1)dt}{e(k+1, t)} p(s, k+1, t) + \left(1 - \frac{b(k) + d(k)}{e(k, t)} dt\right) p(s, k, t) \quad [36]$$

La distribución del grado para toda la red es

$$P(k, t) = \frac{1}{N} \sum_{s=1}^N p(s, k, t) \quad [37]$$

Sumando desde $s = 1$ hasta N a ambos lados de la ecuación [36] se obtiene:

$$\begin{aligned} N \frac{P(k, t + dt) - P(k, t)}{dt} &= \frac{Nb(k-1)}{e(k-1, t)} P(k-1, t) + \frac{Nd(k+1)}{e(k+1, t)} P(k+1, t) \\ &+ \frac{N[b(k) + d(k)]}{e(k, t)} P(k, t) \end{aligned} \quad [38]$$

Si la distribución del grado abarca un numero grande de valores k i.e. si $N_2 - 1 \rightarrow \infty$ y ha alcanzado un estado estacionario, entonces:

$$e(k-1, t) \approx e(k, t) \approx e(k+1, t) \approx \text{constante} \quad [39]$$

Bajo esas condiciones las probabilidades de transición en la ecuación [38] son afectadas por un mismo factor constante y se puede establecer una equivalencia aproximada $P(k, t) \propto n(k, t)$ con la EMNM (ecuación [11]). En la simulación realizada las condiciones [39] se cumplen de manera aproximada para $k < 15$ pero no se cumplen en el resto del intervalo.

Topología de la red

Se han exportado los datos de la red generada a un programa para análisis de redes llamado Gephi (8). Se han medido algunas de sus características resultando lo siguiente:

Grado máximo	30
Grado promedio \bar{k}	8.12
Diámetro	8
Clustering C	0.02
Longitud de camino medio L	3.57

Los resultados mostrados corresponden a una red de 1500 nodos disponibles y 5000 enlaces. Los parámetros de las funciones de nacimiento y muerte son los descritos en [23]. El MED se repitió durante 100 000 ciclos. Al final del proceso 1250 nodos quedaron conectados.

Al comparar los resultados de longitud de camino medio y clustering con los valores esperados para una red aleatoria $L_{random} \approx \ln(N^c)/\ln(\bar{k}) = 3.4$ y $C_{random} \approx \bar{k}/N^c = 0.006$, podemos concluir que la red resultante es un mundo pequeño según la definición de Watts y Strogatz (9) puesto que $L \approx L_{random}$ y $C > C_{random}$.

La longitud de camino medio L se ha obtenido mediante el algoritmo de Ulrik Brandes(10) y el coeficiente de clustering C mediante el de Matthieu Latapy(11).

6 DISCUSIÓN

Se ha mostrado, mediante simulaciones computacionales, que el MED puede ser modelado mediante la Ecuación Maestra [9] pero solo bajo ciertas condiciones, a saber, tiempos muy largos y un intervalo grande para los posibles valores de k , es decir para un número muy grande de nodos. Los resultados de las simulaciones muestran que la distribución del grado en la red final tiene un comportamiento que, a grandes rasgos, asemeja al esperado, pero que no se ajusta de manera exacta a la solución estacionaria [16] de la EMNM.

Puesto que, para valores centrales del grado, la solución estacionaria $n^s(k)$ es muy similar a una ley de potencias, es posible que comparta algunos de los comportamientos sobresalientes de las redes con una distribución libre de escala para el grado. Por ejemplo, se ha observado que, en una red de osciladores con una distribución del grado que sigue una ley de potencias, la sincronización(12) se logra con mayor facilidad y el control(13) se consigue con gran eficiencia. En ambos casos, el efecto se sigue de la influencia que tienen los nodos con alto grado sobre toda la red. Otra característica famosa de las redes con distribución libre de escala es que son resistentes a errores y vulnerables a ataques (14). Se considera como error la remoción de nodos de manera aleatoria y como ataque la remoción de nodos altamente conectados. Como en el ejemplo anterior, ésta característica se sigue de la existencia de algunos nodos altamente conectados y la abundancia de nodos escasamente conectados: una red con distribución tipo $n^s(k)$ cumple con ambas características.

Nos ha parecido de gran interés un ejercicio que relaciona la teoría evolutiva de juegos y las redes con distribución libre de escala (15). En dicho ejercicio se ha considerado que cada nodo es un individuo que juega al dilema del prisionero con sus vecinos y, según el resultado de los juegos, decide qué estrategia le conviene. El dilema del prisionero es un tópico recurrente de la Teoría de Juegos, donde las estrategias posibles son *cooperar* o *defraudar* (16); la cooperación entre jugadores implica la recompensa óptima, pero la adopción de tal estrategia implica un riesgo individual. En el contexto de ciertas redes se ha observado que, después de un número grande de juegos, los nodos cooperadores desaparecen del arreglo cuando la estructura de la red es totalmente aleatoria –con distribución binomial del grado– o totalmente ordenada –con distribución uniforme del grado–. Por otro lado, los nodos cooperadores no desaparecen cuando la red tiene una distribución del grado libre de escala, en esos casos, los nodos cooperadores forman una comunidad que contiene a los nodos altamente conectados (17) y, debido a ello, el balance que resulta de los juegos entre los nodos miembros de dicha comunidad es favorable a la cooperación. Más aún, cuando existen interconexiones entre los nodos altamente conectados –llamadas *correlaciones de antigüedad* en las redes generadas por Vinculación Preferente– se observa que los cooperadores no solo se mantienen en la red sino que dominan a los defraudadores(15).

Parece ser que el requisito para la permanencia de cooperadores en la red es la existencia y proporción de nodos altamente conectados, ambas, consecuencias de una distribución del grado como ley de potencias. Un fenómeno similar podría manifestarse, al menos para la existencia de cooperadores, en las redes con distribución del grado como $n^s(k)$ por que dichas redes también

contienen nodos altamente conectados. Sería interesante indagar como dotar a una red con una característica similar a las *correlaciones de antigüedad*. Trabajos como éste podrían contribuir al diseño de redes con ciertas características a partir de condiciones sobre nodos individuales.

Existen modelos que describen cómo crece una red que en su estado final presenta

una distribución del grado que sigue una ley de potencias. Algunos son deterministas como, por ejemplo, las redes de Apolonio (18) que se construyen de acuerdo a un método predefinido, y otros son estocásticos y consideran mecanismos de conexión en los que intervienen sucesos aleatorios (2)(19).

El modelo estocástico más representativo(2) genera redes con distribución libre de escala mediante un mecanismo denominado Vinculación Preferente. El modelo considera una red inicial a la que progresivamente se agregan nodos que se vinculan de acuerdo a una distribución de probabilidad dada dando preferencia a los nodos con mayor grado. Se han presentado (3) adaptaciones al modelo de Vinculación Preferentes que consideran restricciones (envejecimiento y costo) sobre los enlaces que puede mantener un nodo. Dichas modificaciones alteran la distribución del grado de modo que ya no se sigue la ley de potencias para valores extremos del grado y de esa manera se consigue un mejor ajuste con algunos datos reales.

Es importante señalar que la mayoría de los modelos consideran redes en crecimiento, es decir, en una situación tal que la cantidad de nodos aumenta progresivamente. En un célebre artículo (2) Barabási critica los modelos que consideran un número fijo de nodos arguyendo que la redes en el mundo real están en constante crecimiento. Dicha idea se concentra en el crecimiento global de redes como la web, pero descarta situaciones que en ciertos contextos son relevantes. Por ejemplo, las páginas web pueden perder vínculos al grado de quedar desconectadas del resto de la red y si, en efecto, la web está creciendo, se debe a que el número de sitios que se incorporan es mayor al número de los que se separan. El enfoque planteado por la ecuación maestra [9] es novedoso por que plantea una dinámica de desarrollo que no necesariamente implica el crecimiento de la red y considera, además de la adición de enlaces, la remoción de los mismos como parte esencial de la dinámica. Debe notarse que, pese a que la simulación se hace con un número determinado de nodos, la propuesta planteada por la Ecuación Maestra [9] da lugar a la existencia de un número indeterminado de nodos desconectados y la cantidad de nodos conectados puede cambiar o no dependiendo de los parámetros considerados.

La Vinculación Preferente(2) se ha convertido en un paradigma para los mecanismos de desarrollo en redes. Con fines de comparación, puede interpretarse la función $b(k)$ en la forma propuesta por la ecuación [15] como una Vinculación Preferente atenuada. La Vinculación Preferente indica que la probabilidad de que un nodo reciba un enlace es proporcional a su grado. Puede ser que la conexión con nodos de alto grado se dificulte porque los nodos tiene una capacidad limitada para mantener enlaces – como se sugiere en (3)– o por inaccesibilidad de los nodos con alto grado como podría suceder si los nodos están sujetos a constricciones jerárquicas. En tal caso los enlaces se realizan preferentemente sobre nodos con grado alto pero que sean accesibles, es decir, nodos al centro del intervalo. Entonces la distribución de las conexiones adquiere una forma $b(k)$ como la ecuación[15]. Bajo el mismo orden de ideas, la probabilidad de desconexión $d(k)$ con la forma descrita en la ecuación [14] puede interpretarse como una Desvinculación Preferente atenuada, es decir, una situación en la que las desconexiones se realizan preferentemente sobre nodos de grado bajo pero existe algún factor que dificulta tales eventos. Se ha sugerido que existe una

capacidad limitada para mantener enlaces (3). De manera complementaria, se puede pensar que los nodos con pocas conexiones tienen una capacidad limitada para perder enlaces o, dicho de otra forma, son renuentes a perder sus pocas conexiones. Tales características, la vinculación - desvinculación preferente y atenuada, ocurren con frecuencia en el contexto de las redes sociales.

Se ha propuesto(20) que la similitud u homofilia es también un criterio determinante para la estructura de algunas redes. Las conexiones entre nodos con grado intermedio son frecuentes en redes sujetas a las condiciones de la Ecuación Maestra [9] y, a primera vista, podría pensarse que existe homofilia entre los nodos de grado intermedio. Sin embargo no se ha establecido ningún criterio de similitud entre nodos, y la distribución de las conexiones se sigue de la probabilidad de conexión $b(k)$ que considera todo el intervalo de los grados posibles en la red. Es decir, la dinámica de desarrollo planteada por la EMNM [9] tiene mayor similitud con la Vinculación Preferente que con la homofilia.

7 CONCLUSIONES

Se ha mostrado que el Mecanismo de Enlace y Desconexión (MED) produce una distribución del grado que asemeja una ley de potencias con una caída abrupta al final del intervalo. Eso es significativo porque las distribuciones del grado con esas características normalmente se asocian a poblaciones crecientes de nodos y enlaces. Como en el caso simulado la población de nodos y enlaces es fija, se sugiere que el crecimiento no es una característica indispensable para obtener dichas distribuciones del grado siempre que se considere la desaparición de enlaces.

Por otro lado, se ha observado que el MED puede ser modelado por la ecuación maestra solo bajo condiciones limitadas. Se sigue que dicho mecanismo no refleja de manera fiel la dinámica de reconexiones planteada por la EMNM. Sus limitaciones más evidentes son la rigidez en el orden de los eventos y la dependencia sobre el número de nodos para la probabilidad de transición. El MED es rígido por que impone la condición de que se hagan dos desconexiones para retirar un enlace seguidas de dos conexiones para insertar otro, además, la probabilidad de elección del segundo nodo a enlazarse es afectada por la selección del primero. Para elegir un nodo se determina primero el grado k que será afectado y después se toma cualquier nodo con grado k ; la elección de "cualquier nodo de grado k " es influenciada por la abundancia de nodos de grado k y esa condición establece una diferencia importante con el modelo de la EMNM para la mayoría de los casos.

Para reflejar de mejor manera la dinámica propuesta por la EMNM un mecanismo de reconexiones debería ser muy flexible en cuanto a la elección de nodos, es decir, permitir que las conexiones y desconexiones ocurran sin un orden previo para su secuencia y proporción. Eso se puede lograr permitiendo la existencia de medios enlaces o cabos sueltos en la red y así no forzar a que exista una correspondencia inmediata entre conexiones y desconexiones. En relación a las probabilidades de transición hay que considerar que las expresiones $b(k)$ y $d(k)$ son el resultado de la suma de las probabilidades de conexión y desconexión sobre todos los nodos de la red en un intervalo $[t, t + dt)$; es decir, se puede verificar la idoneidad de un mecanismo aplicando una suma similar a la de la ecuación [38].

El programa que simula el desarrollo de la red puede mejorar si se incorporan rutinas que relacionen el tiempo con las iteraciones tal como se hace en el paso $g3$ del algoritmo Gillespie. El programa realizado puede ser de utilidad para estudiar mecanismos susceptibles a ser modelados por la EMNM. También puede usarse para generar redes con una distribución del grado tipo función beta generalizada. Las redes generadas se pueden comparar con redes del mundo real (21)(22) y así identificar qué características de las redes reales son atribuibles a una dinámica de nacimiento y muerte de enlaces.

8 REFERENCIAS

1. *The structure and function of complex networks*. **Newman**. 2, s.l. : Siam Review, 2003, Vol. 45. PII S0036144503424804.
2. *Emergence of Scaling in Random Networks*. **Barabasi, Albert and Albert, Reka**. s.l. : Science, 1999, Vol. 286.
3. *Classes of small-world networks*. **L. A. N. Amaral, A. Scala, M. Barthé lémy, H. E. Stanley**. 21, s.l. : Applied Physical Sciences, 2000, Vol. 97. 11149–11152.
4. *Universality of Rank-Ordering Distributions in the Arts and Sciences*. **Martínez-Mekler G, Martínez RA, del Río MB, Mansilla R, Miramontes P, Cocho G**. 3, s.l. : PLoS ONE, 2009, Vol. 4. e4791.
5. **R. Alvarez-Martínez, G. Cocho, R. F. Rodríguez, G. Martínez-Mekler**. *Birth and Death Master Equation for the Evolution of Complex Networks*. s.l. : Elsevier, 2013. Preprint.
6. *Complex networks: Structure and dynamics*. **S. Boccaletti, V. Latora, Y. Moreno, M. Chavez , D.U. Hwang**. 424, s.l. : Elsevier, 2005, Vol. Physics Reports. 0370-1573.
7. **RADEK ERBAN, S. JONATHAN CHAPMAN, AND PHILIP K. MAINI**. *A PRACTICAL GUIDE TO STOCHASTIC SIMULATIONS OF REACTION-DIFFUSION PROCESSES*. Oxford : University of Oxford, Mathematical Institute, 2007. arXiv:0704.1908v1.
8. **Bastian, Mathieu**. Gephi. s.l. : gephi.org.
9. *Collective dynamics of small-world networks*. **Strogatz, Duncan J. Watts* & Steven H.** 440, s.l. : Nature, 1998, Vol. 393.
10. *A faster algorithm for Betweenness Centrality*. **Brandes, Ulrik**. s.l. : Journal of Mathematical Sociology, 2001, Vol. 25. 25(2):163-177.
11. *Main-memory Triangle Computations for very large graphs*. **Latapy, Matthieu**. 458, s.l. : Theoretical Computer Science, 2008, Vol. 407.
12. *Synchronization in Scale-Free Dynamical Networks*.: **Xiao Fan Wang**. 1, s.l. : IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS, 2002, Vol. 49. 1057–7122/02.
13. *Pinning control of scale-free dynamical networks*. **Xiao Fan Wang, Guanrong Chen**. 310, s.l. : Elsevier Science B.V., 2002, Vol. Physica A. 0378-4371/02.
14. *Efficiency of scale-free networks: error and attack tolerance*. **Paolo Crucitti, Vito Latora, Massimo Marchiori, Andrea Rapisarda**. 320, s.l. : Elsevier Science B.V., 2003, Vol. Physica A. 0378-4371/03.

15. *Scale-Free Networks Provide a Unifying Framework for the Emergence of Cooperation.* **F. C. Santos, J. M. Pacheco.** 098104, s.l. : PHYSICAL REVIEW LETTERS, 2005, Vol. 95. 0031-9007/05/95(9)/098104(4).
16. *Evolutionary games on graphs.* **György Szabó, Gábor Fáth.** 97 – 216, s.l. : Physics Reports, 2007, Vol. 446. 0370-1573.
17. *Dynamical Organization of Cooperation in Complex Topologies.* **J. Gómez-Gardeñes, M. Campillo, L. M. Floría, Y. Moreno.** 108103, s.l. : PHYSICAL REVIEW LETTERS, 2007, Vol. 98. 0031-9007/07/98(10)/108103(4).
18. *Apollonian Networks: Simultaneously Scale-Free, Small World, Euclidean, Space Filling,.* **José S. Andrade, Hans J. Herrmann, Roberto F. S. Andrade, Luciano R. da Silva.** 018702, s.l. : P H Y S I C A L R E V I E W L E T T E R S, 2005, Vol. 98. 0031-9007/05/94(1)/018702(4).
19. **S.N. Dorogovtsev, J.F.F.Mendes.** *Evolution of Networks.* s.l. : OXFORD, 2003.
20. *Popularity versus similarity in growing networks.* **Fragkiskos Papadopoulos, Maksim Kitsak, M. Ángeles Serrano, Marián Boguñá, Dmitri Krioukov.** 537, s.l. : Nature, 2012, Vol. 489. 10.1038/nature11459.
21. *Scale-free topology of e-mail networks.* **Holger Ebel, Lutz-Ingo Mielsch, Stefan Bornholdt.** 035103, s.l. : Phys. Rev., 2002, Vol. 66. 10.1103/PhysRevE.66.035103.
22. *The large-scale organization.* **H. Jeong, B. Tombor, R. Albert, Z. N. Oltvai, A.L. Barabaási.** 651, s.l. : Nature, 2000, Vol. 407.
23. **E.Gentle, James.** *Statistics ans Computing.* New York : Springer, 2009. 978-0-387-98143-7.
24. **DeGROOT, Morris H.** *Probabilidad y estadística.* México : Addison Wesley Iberoamericana, 1989. 0-201-64405-3.
25. *Random graphs with arbitrary degree distributions and their applications.* **M. E. J. Newman, S. H. Strogatz, D. J. Watts.** 026118, s.l. : PHYSICAL REVIEW E, 2001, Vol. 64. 1063-651X/2001/64(2)/026118(17).