

Análisis de Redes Sociales

Guillermo Jiménez Díaz (gjimenez@ucm.es)
Alberto Díaz (albertodiaz@fdi.ucm.es)

5 de diciembre de 2014

Prefacio

Estos son los apuntes de la asignatura Análisis de Redes Sociales, impartida en la Facultad de Informática de la Universidad Complutense de Madrid por los profesores Guillermo Jiménez Díaz y Alberto Díaz, del Departamento de Ingeniería del Software e Inteligencia Artificial.

Este material ha sido desarrollado a partir de distintas fuentes, destacando como referencia principal el libro *Network Science* de Laszlo Barabasi, el material de la asignatura *Social Network Analysis*, impartido por Lada Adamic a través de Coursera, y las transparencias de la asignatura Redes y Sistemas Complejos, creadas por Óscar Cordón García de la Universidad de Granada.

Para este capítulo se ha utilizado, adicionalmente, material de los libros *Analyzing the Social Web* de Jennifer Goldbeck (capítulo 10) y *Networks, Crowds, and Markets* de Easley y Kleinberg (capítulos 19 y 21).

Este obra está bajo una [licencia de Creative Commons Reconocimiento-NoComercial-CompartirIgual 4.0 Internacional](#).

Tema 8: Propagación y Difusión en redes

Las conexiones presentes en la redes permiten modelar la propagación de todo tipo de elementos entre sus nodos: enfermedades, vídeos virales, rumores, virus informáticos, productos, anuncios, información... En general, la mayoría de estos modelos de propagación son similares independientemente de lo que se pretenda propagar.

Existe desde hace muchos años un estudio intenso en la propagación de enfermedades. El conocimiento de cómo se propagan a través de una red de individuos nos puede servir para entender cómo se puede propagar cualquier otro tipo de información en dicha red. Por este motivo, en este tema vamos a hablar los

modelos fundamentales de propagación de enfermedades y vamos a estudiar como aplicarlos a las redes, analizando cómo la estructura de la misma afecta enormemente a estos fenómenos de propagación.

8.1 Modelos de contagio simple

La epidemiología es la ciencia que estudia la salud y control de enfermedades en una población, así como la predicción de expansión de dichas enfermedades. El modelo general epidemiológico se basa en dos hipótesis:

- **Modelo compartimental:** cada individuo puede estar en un determinado estado dependiendo de en qué fase se la enfermedad se encuentra. El modelo más simple, que será el que usemos, supone 3 estados:
 - **Susceptible (S):** El individuo está sano y puede ser infectado.
 - **Infectado (I):** El individuo está infectado y puede contagiar a otros individuos.
 - **Recuperado (R):** El individuo estuvo contagiado pero se ha recuperado y no puede volver a ser contagiado. También se utiliza para modelar los individuos que no han superado la enfermedad y que han muerto a causa de ella.
- **Mezcla homogénea:** Cualquier individuo tiene la misma probabilidad de estar en contacto con un individuo infectado. Esta hipótesis puede interpretarse como que la red de contactos está modelada mediante una red aleatoria aunque, realmente, elimina la necesidad de conocer los contactos de los individuos y se puede asumir que cualquiera puede infectar a cualquiera.

Los modelos generados a partir de estas hipótesis observan el comportamiento (los cambios de estado) de los individuos a lo largo del tiempo para predecir el alcance y la velocidad de propagación de la enfermedad, entre otras. A continuación vamos a estudiar la dinámica de los modelos de propagación clásicos, que combinan las letras del modelo compartimental.

8.1.1 Modelo SI

Es el modelo más sencillo, en el que un individuo susceptible puede quedar infectado pero, una vez infectado, no se puede recuperar. Un ejemplo de este tipo es el virus del VIH (o los zombies).

En este modelo suponemos que cada individuo tiene $\langle k \rangle$ contactos (enlaces) y que en cada instante de tiempo la enfermedad se propaga con una tasa de contagio δ , que representa la probabilidad de que un individuo infectado transmita la enfermedad a uno susceptible.

Para entender la dinámica del modelo vamos a definir los siguientes parámetros.

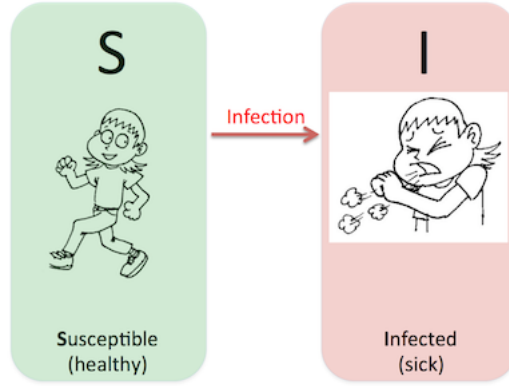


Figure 1: Esquema del Modelo SI

- N es el tamaño de la población y es $N = S(t) + I(t)$, donde $S(t)$ (número de individuos que están en el estado susceptible en el tiempo t) e $I(t)$ (número de individuos que están en el estado infectado en el tiempo t).
- En lugar de manejar valores absolutos vamos a manejar las proporciones o ratios de susceptibles e infectados. De este modo $s(t) = s = \frac{S(t)}{N}$ representa la proporción de individuos susceptibles de la población mientras que $i(t) = i = \frac{I(t)}{N}$.
- β es la tasa de transmisión que incluye el grado medio de cada individuo, esto es, $\beta = \lambda \cdot \langle k \rangle$.

La ecuación diferencial que modela la tasa a la que varía el número de infectados es:

$$\frac{di}{dt} = \beta \cdot i \cdot s = \beta \cdot i \cdot (1 - i)$$

Si resolvemos esta ecuación nos queda que:

$$i = \frac{i_0 \exp(\beta t)}{1 - i_0 + i_0 \exp(\beta t)}$$

donde i_0 representa a la tasa de infectados en el instante $t = 0$. De la representación gráfica de esta función extraemos las siguientes conclusiones:

- Inicialmente el número de infectados crece exponencialmente.
- A medida que el número de infectados se hace mayor, hay menos individuos susceptibles por lo que el crecimiento de infectados se ralentiza y la infección termina cuando todos están infectados ($i(t \rightarrow \infty) = 1$).

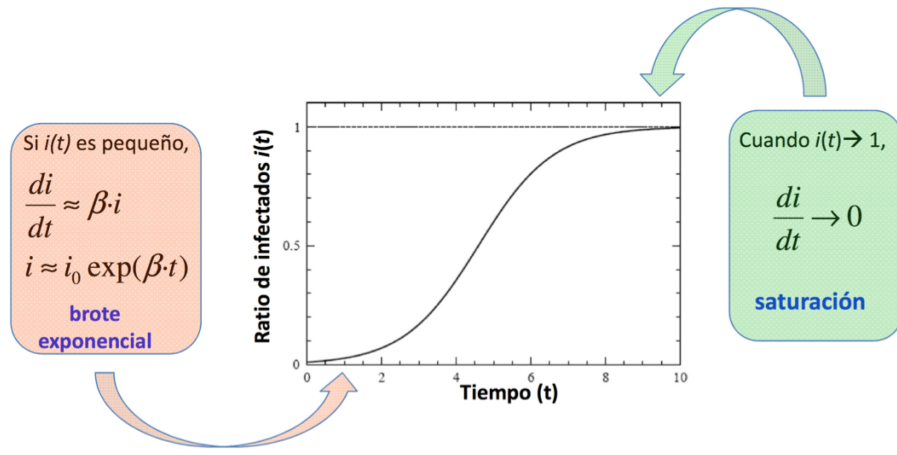


Figure 2: Representación de la proporción de infectados en el Modelo SI

8.1.2 Modelo SIS

Es similar al anterior salvo en que, en este modelo, los individuos infectados se pueden recuperar, volviendo al estado susceptible. Un ejemplo de este modelo es el resfriado común.

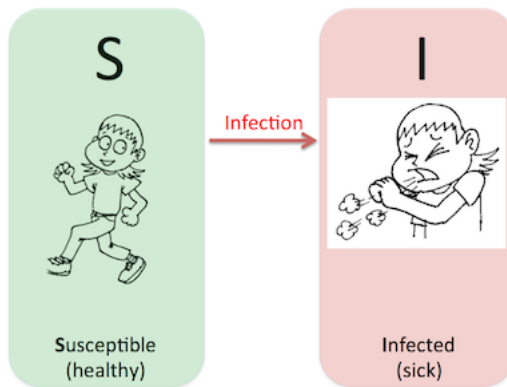


Figure 3: Esquema del Modelo SI

Para este modelo necesitamos, además de los parámetros del anterior, la tasa de recuperación μ , que representa la proporción de infectados que se recuperan y pasan al estado susceptible en cada instante de tiempo.

En este caso, la ecuación diferencial que modela la tasa a la que varía el número de infectados es:

$$\frac{di}{dt} = \beta \cdot i \cdot s - \mu \cdot i = \beta \cdot i \cdot (1 - i) - \mu \cdot i$$

La resolución de esta ecuación nos da el siguiente resultado:

$$i = \left(1 - \frac{\mu}{\beta}\right) \frac{C \cdot e^{(\beta-\mu)t}}{1 + C \cdot e^{(\beta-\mu)t}}$$

$$C = \frac{\beta \cdot i_0}{\beta - \mu - \beta \cdot i_0}$$

En este caso, las conclusiones que podemos extraer de la representación gráfica de la función son las siguientes:

- Como la recuperación es posible, el sistema alcanza un *estado endémico* en el que la tasa de infectados es constante:

$$i(\infty) = 1 - \frac{\mu}{\beta}$$

En este estado endémico, la proporción de infectados no varía con el tiempo y sólo se produce cuando la tasa de recuperación es inferior a la tasa de transmisión ($\mu < \beta$)

- En caso de que la tasa de recuperación sea mayor que la tasa de transmisión ($\mu > \beta$) entonces llegado a un determinado punto la proporción de infectados comienza a decrecer exponencialmente, alcanzado un estado libre de enfermedad, en la que todos los individuos se han recuperado y no hay infectados.

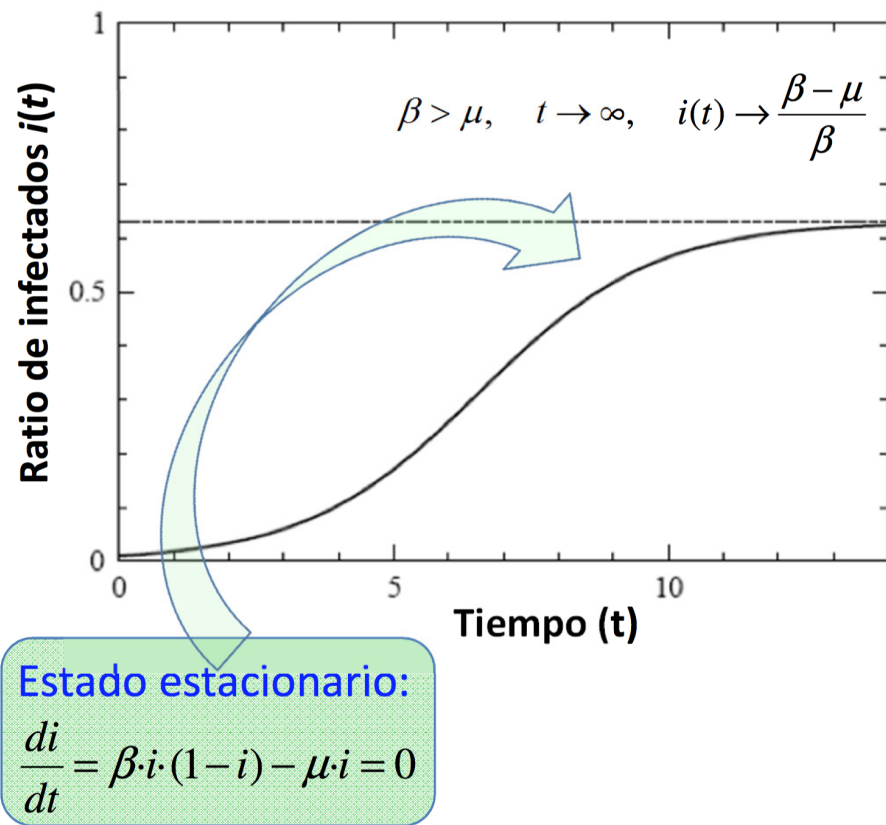


Figure 4: Representación de la proporción de infectados en el Modelo SIS