

# Simulação de Epidemias, 2017/1

Daniel Menasché e Vilc Rufino

July 25, 2017

## Abstract

O objetivo deste trabalho de simulação é introduzir conceitos básicos de um simulador de eventos discretos, e experimentar diferentes fatores que impactam na disseminação de epidemias.

## 1 Introdução

O estudo de epidemias iniciou-se há pelo menos 200 anos atrás. Recentemente, epidemias atraem atenção não só de biólogos, mas também de cientistas das computação, tendo em vista o entendimento da disseminação de malwares como o caso recente do Wanna Crypt<sup>1</sup>.

O objetivo deste trabalho é implementar um simulador de eventos discretos que capture o efeito de diferentes parâmetros, como taxa de cura e taxa de infecção, na disseminação de um vírus.

## 2 Conceitos Básicos

Consideramos uma população finita de  $N$  usuários, conectados em uma clique (todos os usuários estão conectados com todos os outros).

Cada usuário pode estar em um dentre dois estados:

1. **suscetível (S):** um nó no estado suscetível pode tornar-se infectado seja por conta de infecções exógenas (agente infectante fora da rede, e.g.: um pen drive) ou endógenas (agente infectante participante da rede, e.g.: um vizinho contaminado)
2. **infectado (I):** um nó infectado pode infectar seus vizinhos. Um nó infectado torna-se limpo depois de um tempo exponencial, com taxa de descontaminação  $\mu$  (tempo médio para descontaminação igual a  $1/\mu$ )

Assuma que infecções exógenas chegam a cada nó segundo um fluxo Poisson com taxa  $\lambda$ .

Assuma também que cada nó infectado infecta cada um de seus vizinhos segundo um fluxo Poisson com taxa  $\gamma$ .

A figura 1 apresenta um modelo de transição de estados para um nó:

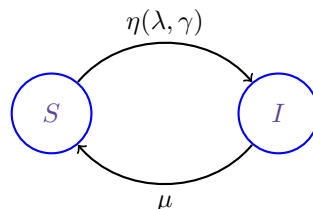


Figure 1: Diagrama de estados para o modelo de epidêmico SIS.

onde:

---

<sup>1</sup><https://blogs.technet.microsoft.com/nmpc/2017/05/12/wannacrypt-ransomware-worm-targets-out-of-date-systems/>

$\mu$	Taxa de descontaminação (limpeza)
$\eta$	Taxa de infecção em função das taxas exógena ( $\lambda$ ) e endógena ( $\gamma$ )
$S$	Estado suscetível
$I$	Estado infectado

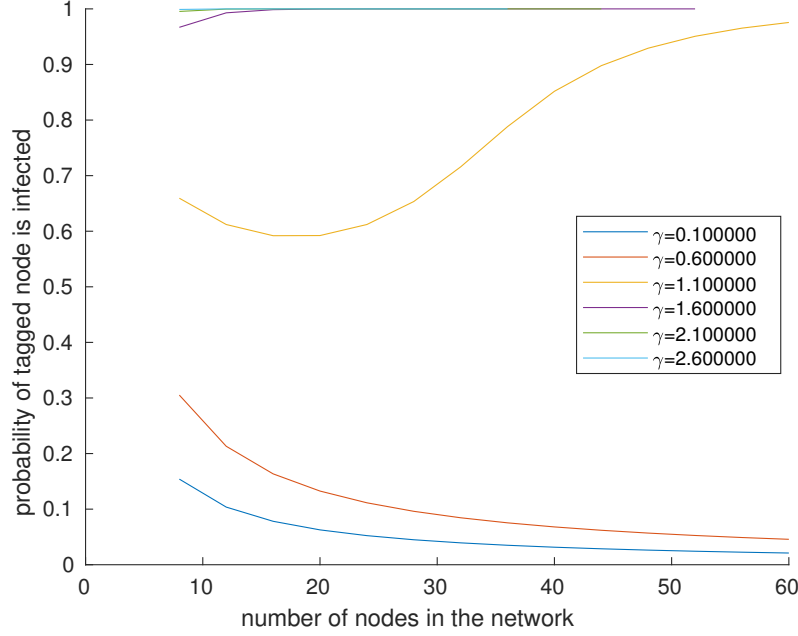


Figure 2: Probabilidade de um determinado nó estar infectado, em função do tamanho da população. A taxa de infecção endógena,  $\gamma$ , varia entre 0.1 e 2.6. Para valores pequenos de  $\gamma$ , a probabilidade de infecção diminui na medida em que a população aumenta. Para valores maiores de  $\gamma$ , na medida em que a população aumenta a probabilidade de infecção aumenta.

### 3 Contribuições Aditivas ou Multiplicativas

Suponha que um nó suscetível (S) tenha  $d$  vizinhos infectados (I). Assumimos que o tempo até infecção do nó suscetível é dado por uma distribuição exponencial. Mas qual a taxa com que o nó suscetível é infectado? Ou seja, qual a taxa da variável exponencial que caracteriza o tempo até infecção?

1. modelo multiplicativo: segundo o modelo multiplicativo, a taxa de infecção é igual a  $\lambda\gamma^d$ , ou seja, o tempo médio até infecção é igual a  $1/(\lambda\gamma^d)$
2. modelo aditivo: segundo o modelo aditivo, a taxa de infecção é igual a  $\lambda + d\gamma$ , ou seja, o tempo médio até infecção é igual a  $1/(\lambda + d\gamma)$

No restante deste trabalho, espera-se que você considere e implemente o modelo multiplicativo. Foi assim que a Figura 2 foi gerada. Embora seja obrigatória a implementação do modelo multiplicativo, se você quiser, pode também implementar o modelo aditivo, e comparar os dois modelos. Isso pode fazer parte, por exemplo, das suas contribuições conforme apontado na seção 5.2.

### 4 Exemplo Ilustrativo

A Figura 2 ilustra como a probabilidade de infecção varia em função do tamanho da população. O tamanho da população, por sua vez, está relacionado ao número de indivíduos que não aplicaram a vacina. Os indivíduos que tomaram a vacina estão fora de consideração, porque não podem nem se infectar, nem infectar outros nós.

Na Figura 2 assumimos que infecções exógenas chegam a cada nó segundo um fluxo Poisson com taxa  $\lambda = \frac{C}{N}$ , onde  $C$  é uma constante (definida como 1) e  $N$  é o tamanho da população (número

variável	descrição	valor padrão
$N$	tamanho da população (número de nós não vacinados)	-
$\mu$	taxa de desinfecção (cura)	1
$C$	taxa de infecção exógena agregada	1
$\lambda = C/N$	taxa de infecção exógena por nó	$1/N$
$\gamma$	taxa de infecção endógena	0.1:0.5:2.6

de nós que não se vacinaram). Note que na medida em que  $N$  aumenta, a probabilidade de um nó ser infectado por um agente externo diminui. Isso ocorre porque assumimos que o agente externo tem poder de ataque constante  $C$ , mas esse poder é dividido entre os  $N$  membros da população.

Seja  $\pi_I$  a probabilidade de um nó estar infectado.

**Regime exógeno predominante:** Se a taxa de contaminação endógena for pequena ( $0 \leq \gamma \leq 1$ ), observamos que  $\pi_I$  diminui na medida em que  $N$  aumenta. Isso ocorre porque nesse regime quanto maior o tamanho da população, menor a chance de um elemento externo selecionar um determinado indivíduo como alvo. Se a população for pequena, por outro lado, existe uma grande chance de aqueles poucos indivíduos suscetíveis serem alvos fáceis do poder de contaminação do agente exógeno (e.g.: scans que buscam por elementos vulneráveis, ZMap, que mapeia rapidamente em algumas horas vulnerabilidades em todo o espaço de endereços IPv4). Nesse regime, o sistema é dominado por efeitos exógenos.

**Regime endógeno predominante:** Por outro lado, quando a taxa de contaminação endógena é grande ( $\gamma > 1$ ), observamos que  $\pi_I$  aumenta na medida em que o tamanho da população aumenta. Isso ocorre porque nesse regime, dominado por fatores endógenos, quanto mais indivíduos suscetíveis, maior a circulação do vírus na rede, e maior a chance de um nó infectar-se. Lembre que um nó infectado pode infectar seus vizinhos.

Finalmente, note que quando a taxa de contaminação endógena assume valores intermediários ( $\lim_{\gamma \rightarrow 1}$ ), temos um comportamento híbrido (a curva de infecção primeiro diminui, e depois aumenta).

## 5 Questões

### 5.1 Parte 1: Primeiros Passos

O seu primeiro objetivo no trabalho de simulação é reproduzir a Figura 2. Para tal, simule o sistema para diferentes parâmetros, e procure regeerar a Figura 2 da forma mais fidedigna possível, incluindo intervalos de confiança.

### 5.2 Parte 2: Usando a Criatividade

O segundo objetivo do trabalho é fazer variações na topologia (por exemplo, considerar uma estrela ao invés de uma clique), taxas de infecção exógena e endógena, e taxa de cura, e verificar os efeitos dessas mudanças na probabilidade de um nó estar infectado.

Você pode usar-se de sua criatividade para variar o processo de epidemia, e ver os efeitos da variação em cima dos resultados. Uma sugestão, inclusive, foi dada na seção 3.

## 6 Trabalhos Relacionados

O principal trabalho relacionado a este é [ZM14]. É interessante que pelo menos um dos membros de cada grupo leia o trabalho relacionado, e traga ideias de como usar este trabalho relacionado para validar seu simulador, bem como para inspirar novas variantes da sua simulação que possam ser de interesse na análise.

## 7 Nota

A sua nota será proporcional a

1. qualidade dos seus resultados,

2. criatividade na variação de parâmetros e cenários,
3. profundidade da discussão de resultados,
4. detalhamento das abordagens usadas para verificação da corretude do seu simuladores, incluindo gráficos e números que comprovem que seu simulador está correto em vários casos base simples

## 8 Dúvidas

Dúvidas devem ser direcionadas ao Vilc (vilc@bol.com.br), de preferência com cópia para o restante da turma.

## 9 Anexo

### 9.1 Modelos de epidemias

Os métodos analíticos para modelar epidemias SIS em redes de heterogêneas por meio de suas topologias podem ser:

1. Descrição da topologia por meio de suas propriedades estatísticas como grau de distribuição e grau de correlações. Esses métodos consideram o processo de difusão no limite de grandes redes usando análise mean-field [BBV08], [DFH<sup>+</sup>11], [SM11], [HK11] e [PSV01];
2. Descrição exata da topologia por sua matriz de adjacência da rede (grafo) em análise, o qual é estático e de tamanho finito. A epidemia é tratada como um processo de Markov e a taxa de infecção é dependente do número de vizinhos infectados [DGM06], [GMT05], [CWW<sup>+</sup>08], [WCWF03], [ZM12];

#### 9.1.1 Representação multiplicativa

Uma forma de representar o processo de Markov na topologia exata é demonstrado por [ZM15] and [ZM14], denominado processo *scaled SIS*, onde o tempo que o nó está no estado suscetível,  $\hat{T}$ , é uma variável aleatória dada por uma distribuição exponencial, escalonada pela taxa de infecção exógena,  $\hat{T} \sim \exp(\lambda\gamma^d)$ , onde  $\lambda$  é a taxa de infecção exógena,  $\gamma^d$  é a taxa de infecção endógena e  $d$  é o número de vizinhos infectados.

Para o caso da representação *scaled SISI* o processo de Markov é reversível, e a equação de balanceamento é dada por [Kel79]:

$$\pi(\mathbf{x})q(\mathbf{x}, \mathbf{x}') = \pi(\mathbf{x}')q(\mathbf{x}', \mathbf{x}), \mathbf{x}, \mathbf{x}' \in \mathcal{Y}$$

Para o processo *scaled SIS*:

$$\pi(\mathbf{x}) = \frac{1}{Z} \left( \frac{\lambda}{\mu} \right)^{1^T \mathbf{x}} \gamma^{\mathbf{x}^T A \mathbf{x} / 2}, \mathbf{x} \in \mathcal{X}$$

Onde  $Z$  é a *função de partição* (normalização) definida como:

$$Z = \sum_{\mathbf{x} \in \mathcal{X}} \left( \frac{\lambda}{\mu} \right)^{1^T \mathbf{x}} \gamma^{\mathbf{x}^T A \mathbf{x} / 2}$$

$\mathbf{x}$  é o vetor binário que representa os agentes contaminados;

$\mathcal{X}$  é o conjunto de todos os valores possíveis de  $\mathbf{x}$

#### 9.1.2 Representação aditiva

A forma clássica de representação, que será denominada *non-scaled SIS*, considera o tempo que o nó está no estado suscetível,  $\hat{T}$ , como uma variável aleatória dada pelo menor valor obtido entre duas distribuições exponenciais  $\alpha \sim \exp(\lambda)$  e  $\beta \sim \exp(\gamma d)$ :

$$\hat{T} = \min(\alpha, \beta)$$

Portanto  $T \sim \exp(\lambda + \gamma d)$

## 9.2 Material disponível no dropbox

[https://www.dropbox.com/sh/x8zwkqu35m7uc21/AADZsQ\\_qPgLnzpZhOZiSZocSa?dl=0](https://www.dropbox.com/sh/x8zwkqu35m7uc21/AADZsQ_qPgLnzpZhOZiSZocSa?dl=0)

## References

- [BBV08] M Barthelemy, A Barrat, and A Vespignani. Dynamical processes on complex networks. 2008.
- [CWW<sup>+</sup>08] Deepayan Chakrabarti, Yang Wang, Chenxi Wang, Jurij Leskovec, and Christos Faloutsos. Epidemic thresholds in real networks. *ACM Trans. Inf. Syst. Secur.*, 10(4):1:1–1:26, January 2008.
- [DFH<sup>+</sup>11] Leon Danon, Ashley P Ford, Thomas House, Chris P Jewell, Matt J Keeling, Gareth O Roberts, Joshua V Ross, and Matthew C Vernon. Networks and the epidemiology of infectious disease. *Interdisciplinary perspectives on infectious diseases*, 2011, 2011.
- [DGM06] Moez Draief, Ayalvadi Ganesh, and Laurent Massoulié. Thresholds for virus spread on networks. In *Proceedings of the 1st international conference on Performance evaluation methodologies and tools*, page 51. ACM, 2006.
- [GMT05] Ayalvadi Ganesh, Laurent Massoulié, and Don Towsley. The effect of network topology on the spread of epidemics. In *INFOCOM 2005. 24th Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings IEEE*, volume 2, pages 1455–1466. IEEE, 2005.
- [HK11] Thomas House and Matt J Keeling. Insights from unifying modern approximations to infections on networks. *Journal of The Royal Society Interface*, 8(54):67–73, 2011.
- [Kel79] Frank P Kelly. Reversibility and stochastic networks. 1979.
- [PSV01] Romualdo Pastor-Satorras and Alessandro Vespignani. Epidemic spreading in scale-free networks. *Physical review letters*, 86(14):3200, 2001.
- [SM11] A. Santos and J. M. F. Moura. Emergent behavior in large scale networks. In *2011 50th IEEE Conference on Decision and Control and European Control Conference*, pages 4485–4490, Dec 2011.
- [WCWF03] Yang Wang, Deepayan Chakrabarti, Chenxi Wang, and Christos Faloutsos. Epidemic spreading in real networks: An eigenvalue viewpoint. In *Reliable Distributed Systems, 2003. Proceedings. 22nd International Symposium on*, pages 25–34. IEEE, 2003.
- [ZM12] June Zhang and José MF Moura. Accounting for topology in spreading contagion in non-complete networks. In *Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on*, pages 2681–2684. IEEE, 2012.
- [ZM14] June Zhang and José MF Moura. Diffusion in social networks as sis epidemics: Beyond full mixing and complete graphs. *IEEE Journal of Selected Topics in Signal Processing*, 8(4):537–551, 2014.
- [ZM15] June Zhang and José MF Moura. Contact process with exogenous infection and the scaled sis process. *arXiv preprint arXiv:1507.00396*, 2015.