

ESTIMANDO O “PACIENTE ZERO” DE UM SURTO EPIDÊMICO

OLAVO H. MENIN¹ & CHRIS T. BAUCH²

¹ Instituto Federal de Educação, Ciência e Tecnologia de São Paulo, Câmpus Sertãozinho, Brasil.
(olavohmenin@ifsp.edu.br)

² Department of Applied Mathematics, University of Waterloo, Canadá.

Área de conhecimento (Tabela CNPq): 1.01.04.02-0 Análise Numérica

Apresentado no
2º Congresso de Pós-Graduação do IFSP
06 a 09 de novembro de 2017 - Cubatão-SP, Brasil

RESUMO: O problema inverso de inferir a fonte de um surto epidêmico (paciente zero) é abordado utilizando-se o algoritmo *Simulated Annealing*. O modelo epidêmico suscetível-infectado-recuperado (SIR) numa rede regular é considerado. Simulações numéricas foram realizadas e os resultados mostram que o algoritmo é capaz de determinar a localização do paciente zero com boa precisão.

PALAVRAS-CHAVE: modelos epidêmicos; problema inverso; simulated annealing; otimização.

ABSTRACT: The inverse problem of inferring the source of an epidemic outbreak (patient zero) is addressed by using the Simulated Annealing algorithm. The epidemic model susceptible-infected-recovered (SIR) in a regular lattice is considered. Computational simulations were performed and results show that the algorithm is capable to retrieve the locate of the patient zero with good accuracy.

KEYWORDS: epidemic models; inverse problem; simulated annealing; optimization.

INTRODUÇÃO

Desde meados do século 20, modelos matemáticos e computacionais têm sido amplamente desenvolvidos e aplicados na área de epidemiologia, ajudando na elaboração e implementação de políticas públicas de controle e até erradicação de inúmeras doenças infecciosas (KEELING & ROHANI, 2008). Nesse contexto, uma questão bastante relevante refere-se em estimar a fonte de um surto epidêmico, conhecido como problema do “paciente zero”. Apesar do termo “paciente zero” ter sido cunhado erroneamente para identificar o primeiro paciente com HIV/AIDS nos EUA, no início da década de 1980 (AUERBACH, D. M. et al, 1984; WOROBEY, M. et al, 2016), ele continua sendo utilizado para se referir à origem de um processo de propagação, tanto de doenças infecciosas, como rumores ou vírus de computadores (ANTULOV-FANTULIN, N. et al, 2015; FARAJTABAR, M. et al, 2015).

De forma simplificada, o problema do “paciente zero” pode ser descrito da seguinte forma: conhecendo-se a distribuição espacial de indivíduos infectados (pessoas, cidades, computadores etc) de uma população em um dado momento do surto epidêmico (‘snapshot’), identificar o primeiro indivíduo que foi infectado e deu origem ao processo de propagação (LOKHOV, A. Y. M. et al, 2014; ANTULOV-FANTULIN, N. et al, 2015). Neste trabalho, propomos a aplicação do algoritmo *Simulated Annealing* (SA) para resolver o problema inverso do “paciente zero” (KIRKPATRICK, S. et al, 1983; MENIN, O. H. et al, 2013). Mais especificamente, consideramos o clássico modelo epidêmico suscetível-infectado-recuperado (SIR) em uma rede quadrada regular e analisamos a performance do algoritmo em relação à fração da população cujo estado epidêmico pode ser observado durante o ‘snapshot’.

MATERIAL E MÉTODOS

O algoritmo desenvolvido neste trabalho baseia-se na abordagem funcional. Essa técnica consiste em minimizar uma função objetivo que mede a discrepância entre dois conjuntos de dados: um obtido a partir do estado epidêmico da população real e outro a partir do estado epidêmico gerado por um um

modelo numérico, no qual considera-se uma possível solução candidata para a localização do “paciente zero”. Com isso, o processo de otimização consiste que encontrar a solução candidata que produza o mínimo global da função objetivo.

Nosso modelo numérico considera uma rede quadrada com $N \times N$ nós, (i, j) , com $i, j = 1, \dots, N$, todos inicialmente no estado suscetível (S) no tempo $t = 0$, com exceção do nó (i_0, j_0) , que se encontra no estado infeccioso (I) e é considerado como uma solução candidata para o “paciente zero”. A cada tempo discreto $t = 1, 2, \dots, t_s$ (em dias), onde t_s é o tempo no qual o ‘snapshot’ da rede é registrado, cada nó (i, j) no estado suscetível pode se tornar infectado ($S \rightarrow I$) com probabilidade $p^{(i,j)} = 1 - (1 - \beta)^{n_{i,j}}$, onde β é a probabilidade de contágio entre um nó infectado e um nó susceptível e $n_{i,j}$ é o número de vizinhos infectados do nó (i, j) . Além disso, cada indivíduo infectado pode passar para o estado recuperado ($I \rightarrow R$) com probabilidade γ , adquirindo imunidade permanente. No tempo t_s , os estados epidêmicos $Z^{(i,j)} = S, I$ ou R de cada nó (i, j) na população real e no modelo numérico são comparados por meio da função objetivo

$$f(x_0, y_0) = \frac{1}{N_\rho^2} \sum_{i,j=1}^{N_s} \phi^{(i,j)}, \quad \text{onde} \quad \phi_{i,j} = \begin{cases} 0, & \text{se } Z_{\text{real}}^{(i,j)} = Z_{\text{num}}^{(i,j)}, \\ 1, & \text{se } Z_{\text{real}}^{(i,j)} \neq Z_{\text{num}}^{(i,j)}, \end{cases} \quad (1)$$

onde N_ρ é o número de nós observado no ‘snapshot’, o que define a fração $\rho = N_\rho/N$.

Partindo de uma solução candidata inicial, X_{old} , o algoritmo SA gera, a cada iteração τ , uma nova solução X_{new} por meio do modelo de visitação

$$X_{\text{new}} = X_{\text{old}} + T_v N(0, 1), \quad (2)$$

onde $N(0, 1)$ é um número aleatório com distribuição normal padrão e T_v é a “temperatura” de visitação. A nova solução é então aceita com probabilidade

$$p(X_{\text{old}} \leftarrow X_{\text{new}}) = \begin{cases} 1, & \text{se } \Delta f \leq 0, \\ e^{-\Delta f/T_a}, & \text{se } \Delta f > 0, \end{cases} \quad (3)$$

onde $\Delta f = f(X_{\text{new}}) - f(X_{\text{old}})$ e T_a é a “temperatura” de aceitação. O controle da estocasticidade do processo de otimização é realizado por meio dos cronogramas de resfriamento

$$T_a^{(\tau)} = \alpha^\tau T_a^{(0)} \quad \text{e} \quad T_v^{(\tau)} = T_v^{(0)} \frac{\ln(2)}{\ln(\tau + 1)}, \quad (4)$$

onde $\alpha \in (0, 1)$ é a taxa de resfriamento.

RESULTADOS E DISCUSSÃO

Simulações computacionais foram realizadas considerando uma rede regular quadrada com 10000 nós ($N = 100$) e condições periódicas de contorno. A propagação da epidemia seguiu o modelo SIR com $\beta = 0,5$ e $\gamma = 0,1$. O estado epidêmico da população (rede) foi avaliado no tempo $t = 60$ dias (‘snapshot’). As “temperaturas” iniciais de aceitação e visitação foram definidas, respectivamente, como $T_a^{(0)} = 10$ e $T_v^{(0)} = 10000$, de forma que que no início do processo de otimização o algoritmo tivesse boas chances de escapar de mínimos locais e fosse capaz de realizar uma busca em todo o espaço de soluções. A taxa de resfriamento foi definida como $\alpha = 0,9$. Para avaliar a robustez do algoritmo, considerou-se $\rho = 1$, $\rho = 0,1$ e $\rho = 0,01$, correspondendo a 100%, 10% e 1% dos nós rede disponíveis para fornecerem seu estado epidêmico durante o ‘snapshot’. Para cada um desses valores, o algoritmo foi executado 100 vezes e os resultados são mostrados na Tabela 1 e na Figura 1. Pode-se notar que para todos os valores de ρ as soluções médias para a localização (x_0, y_0) do paciente zero obtidas pelo algoritmo foram satisfatoriamente próximas da localização real (20, 30). Contudo, a dispersão em torno da média (desvio padrão) é consideravelmente maior para $\rho = 0,01$ do que para $\rho = 1$ e $\rho = 0,1$.

CONCLUSÕES

O algoritmo apresentado, baseado no método de otimização *Simulated Annealing*, mostrou ser capaz de inferir a localização do “paciente zero” com satisfatória precisão, mesmo quando apenas 10% da população pode ter seu estado epidêmico averiguado. Estruturas de rede mais complexas, como redes mundo-pequeno (*small-world network*) e redes sem-escala (*scale-free network*) devem ser consideradas em estudos futuros.

Tabela 1: Média e desvio padrão das soluções para coordenadas x_0 e y_0 do paciente zero considerando 100 realizações e $\rho = 1$, $\rho = 0,1$ e $\rho = 0,01$. A localização real do paciente zero é (20, 30).

	$\rho = 1$		$\rho = 0,1$		$\rho = 0,01$	
	\hat{i}_0	\hat{j}_0	\hat{i}_0	\hat{j}_0	\hat{i}_0	\hat{j}_0
Média	19,1	31,5	20,0	31,6	19,9	33,5
Desvio Padrão	1,8	2,1	2,2	2,1	4,3	4,5

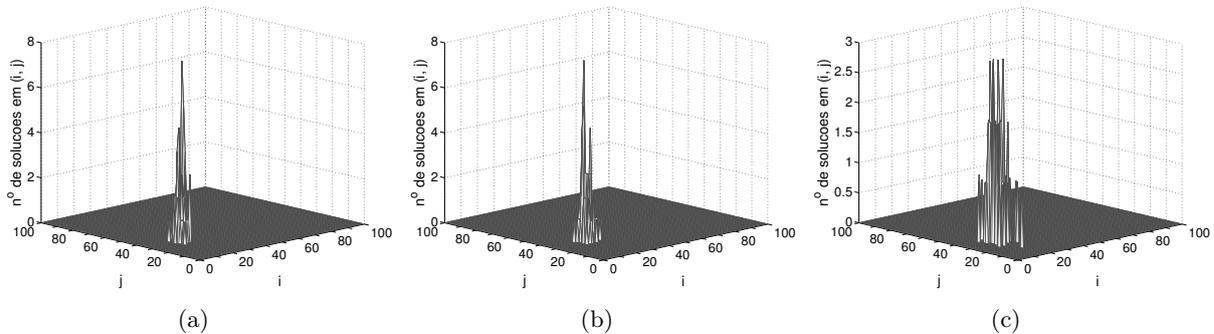


Figura 1: Distribuição espacial das soluções para a localização (i, j) do “paciente zero” considerando 100 realizações para (a) $\rho = 1$, (b) $\rho = 0,1$ e (c) $\rho = 0,01$. A localização real do “paciente zero” é (20, 30).

AGRADECIMENTOS

Os autores agradecem ao Instituto Federal de Educação, Ciência e Tecnologia de São Paulo e ao Departamento de Matemática Aplicada da Universidade de Waterloo.

REFERÊNCIAS

- ANTULOV-FANTULIN, N.; et al. Identification of patient zero in static and temporal networks: robustness and limitations. *Phys. Rev. Lett.* v. 114, p. 248701, 2015.
- AUERBACH, D. M.; et al. Cluster of cases of the acquired immune deficiency syndrome patients linked by sexual contact. *Am. J. Med.* v. 76, p. 487–492, 1984.
- FARAJTABAR, M.; et al. Back to the past: source identification in diffusion networks from partially observed cascades. *18th AISTATS*, v. 38, San Diego, USA, 2015.
- KEELING, M. J.; ROHANI, P. *Modeling Infectious Diseases in Humans and Animals*. Princeton University Press, Princeton 2008.
- KIRKPATRICK, S.; et al. Optimization by simulated annealing. *Science* v. 220, p. 671–680, 1983.
- LOKHOV, A. Y. M.; et al. Inferring the origin of an epidemic with a dynamic message-passing algorithm. *Phys. Rev. E* v. 90, p. 012801, 2014.
- MENIN, O. H.; et al. Boundary element method and simulated annealing algorithm applied to electrical impedance tomography image reconstruction. *Rev. Bras. Ens. Fís.* v. 35, p. 2304, 2013.
- WOROBAY, M.; et al. 1970s and ‘Patient 0’ HIV-1 genomes illuminate early HIV/AIDS history in North America. *Nature* v. 539, p. 98–101, 2016.