

## **EpiGTA/TerraView: Plug-in para Simulação de Enfermidades Infecciosas em Redes Sociais com Fluxo de Populações**

**Carolina S. Pena<sup>1</sup>, Marcelo A. Costa<sup>1</sup>, André de Souza Gomes<sup>2</sup>**

<sup>1</sup>Departamento de Estatística – Universidade Federal de Minas Gerais (UFMG)  
CEP 31.270 - 901 – Belo Horizonte, MG – Brasil

<sup>2</sup>Programa de Pós Graduação em Engenharia Elétrica – Universidade Federal de Minas Gerais (UFMG)

carolinasend@gmail.com, azevedo@est.ufmg.br, asgomes86@gmail.com

**Abstract.** *This project aims to develop a TerraView plug-in application that simulates spatial and temporal spread of highly contagious diseases based on social networks with flow of populations. Three discrete-time stochastic SIR (Susceptible-Infective-Recovered) models will be implemented. The main application of this project is the study of the spread of foot-and-mouth disease in the Brazilian territory using the electronic records of cattle movements, provided by the Ministry of Agriculture and Livestock.*

**Resumo.** *Este projeto tem como objetivo implementar uma ferramenta computacional, na forma de um plug-in do aplicativo TerraView, capaz de gerar cenários de disseminação espaço-temporal de doenças altamente contagiosas. Para isso é utilizado um grafo do fluxo de movimentos de populações em redes sociais. Três modelos estocásticos de tempo discreto do tipo SIR (Suscetível-Infetado-Recuperado) serão implementados. A aplicação inicial deste projeto é o estudo da disseminação da Febre Aftosa no território brasileiro a partir do registro da movimentação de bovinos, registradas nas Guias eletrônicas de Trânsito Animal (e-GTA).*

### **1. Introdução**

O desempenho do setor agropecuário brasileiro ao longo das últimas décadas contribuiu substancialmente para atenuar os efeitos locais da crise financeira internacional. Em 2008, o agronegócio brasileiro respondeu por um em cada três reais de renda gerada no país e contribuiu decisivamente para o superávit da balança comercial. O setor abrangeu 28% do PIB, 36% das exportações e 37% dos empregos (MAPA, 2009).

A evolução da condição sanitária têm sido um fator primordial para a manutenção e ampliação dos mercados internacionais. No entanto, alguns importantes mercados, como Estados Unidos, Japão e a União Européia ainda estabelecem restrições à carne bovina in natura, alegando problemas sanitários.

Por outro lado, a movimentação de animais no país é intensa, especialmente para as espécies bovina e suína. Se um surto de uma doença ocorrer, o trânsito de gado poderia ser suspenso temporariamente para ajudar a conter o avanço da doença. No entanto, parar todo esse transporte pode resultar em gastos muito altos para os

produtores, sendo que as indenizações pagas pelo governo pela intervenção e pelo sacrifício de animais não cobrem todos os custos.

Vindo de encontro às necessidades apontadas acima, este projeto propõe a implementação de uma ferramenta computacional, calibrada com os bancos de dados de trânsito de animais. Essa ferramenta, denominada *EpiGTA*, permitirá avaliar, mediante simulações, os possíveis cenários de disseminação inicial de doenças de animais altamente contagiosas no país, a partir do fluxo de movimentos de animais. O intuito é que o software possa auxiliar aos órgãos oficiais de defesa sanitária e sanidade animal no desenvolvimento e otimização de potenciais estratégias de prevenção e controle de tais doenças, visando impedir a sua incursão/re-introdução/disseminação no país.

## 2. Material e Métodos

A movimentação de bovinos no território brasileiro é atualmente documentada eletronicamente através da emissão de Guias de Trânsito Animal (GTAs). As informações utilizadas no desenvolvimento desse projeto foram retiradas dessas guias, são elas: municípios de origem e destino, finalidade (cria, engorda, reprodução, abate ou exposição), data da emissão da GTA e quantidade de animais transportados. Os dados analisados nesse artigo são referentes às GTAs emitidas em 2006 no estado do Acre.

Uma forma de analisar dados de trânsito de animais entre propriedades é através de uma rede de contatos direcionada, na qual as propriedades rurais são representadas por nós e os movimentos de gado entre propriedades são as arestas. A matriz de adjacência,  $G$ , é uma representação matemática adequada para descrever as conexões entre os nós de uma rede (Keeling & Eames, 2005). Considerando nós como indivíduos de uma população, define-se  $G_{ij}=1$ , se existe uma conexão que permita que a infecção passe do indivíduo  $i$  para o indivíduo  $j$ , e  $G_{ij}=0$ , em caso contrário. Nesse estudo considera-se que a infecção pode ocorrer somente em um sentido, dessa forma, a rede é representada por um grafo direcionado (Harary, 1969; Bollobás, 1979).

Representações que utilizam uma única matriz de adjacência para um período de  $n$  dias de simulação são chamadas de redes estáticas. Nesse trabalho, a representação dos movimentos será feita por meio de uma rede dinâmica. Na prática, esse tipo de rede consiste em um conjunto de inúmeras redes estáticas, uma para cada unidade de tempo. Dessa forma, se houver uma movimentação de animais da fazenda  $i$  para a fazenda  $j$  no dia  $t$ , então a rede para o dia  $t$  deve conter uma aresta  $i \rightarrow j$ . A rede dinâmica para um período de  $n$  dias contados a partir do dia  $t_0$  é dada por:

$$G(t_0, n) = \{G(t_0), \dots, G(t_{n-1})\},$$

As redes dinâmicas têm sido consideradas o padrão ouro de representação de movimentos em diversos estudos (Vernon & Keeling, 2008). Em termos de transmissão de doenças, quanto maior o número de conexões de uma rede, mais provável é a disseminação da doença.

Para realizar simulações, serão incorporados três modelos estocásticos de tempo discreto do tipo SIR (Suscetível-Infetado-Recuperado). A suposição básica destes modelos é que um indivíduo pode passar pelos estágios de suscetibilidade, infecção e recuperação, e que a imunidade é permanente após a recuperação. Existe ainda uma simplificação do modelo SIR, no qual o indivíduo infectado não se recupera, nesse caso

considera-se que o tempo de recuperação é infinito ( $R = \infty$ ), reduzindo o modelo SIR para o modelo SI (Suscetível-Infectado).

Em um dos modelos avaliados, a menor unidade considerada é a fazenda (Vernon & Keeling, 2008). Nas simulações desse modelo todas as fazendas começam no estado susceptível (S), exceto uma delas que é aleatoriamente escolhida para iniciar o processo no estado infectado (I). O modelo é sequencialmente atualizado utilizando passos diários. Durante cada dia de simulação a doença é transmitida, através do contato (aresta), de fazendas infectadas para fazendas susceptíveis (nos), com probabilidade  $p$ . As fazendas se mantêm no estado infectado (I) por um número inteiro de interações,  $\mu$ , e então passam para o estado recuperado (R). Nós que atingem o estado R se mantêm no mesmo até o final da simulação. O parâmetro  $p$  se mantém constante durante qualquer simulação específica e o parâmetro  $\mu$  pode ser substituído por uma variável aleatória como, por exemplo, uma distribuição exponencial com média  $\lambda$ .

Nesse tipo de modelo todas as fazendas são consideradas idênticas, de tal forma que nem mesmo o número de animais transportados afeta a dinâmica da transmissão da doença, conforme é descrito em (1) (Vernon & Keeling, 2008):

$$\begin{cases} p(\text{fazenda}(i,t+1) = I \mid \text{fazenda}(i,t) = S) = 1 - \prod_{\text{fazenda}(j,t)=I} 1 - pG_{j,i}(t) \\ p(\text{fazenda}(i,t+1) = R \mid \text{fazenda}(i,t - \mu_i) \neq S) = 1 \end{cases} \quad (1)$$

onde  $\text{fazenda}(i,t) \in \{SIR\}$  é o estado do nó (fazenda)  $i$  no tempo  $t$ ,  $\mu_i$  é o tempo de recuperação da fazenda  $i$  e  $G_{ji}(t) = 1$  se houve movimentação do nó  $j$  para o nó  $i$  no instante  $t$  e  $G_{ji}(t) = 0$ , caso contrário. A probabilidade de contaminação da fazenda  $i$  no tempo  $t+1$  (primeira linha da equação) depende a estrutura da rede, enquanto o processo de recuperação é independente, operando ao nível da propriedade (nó). Os dois outros modelos avaliados consideram o animal como a menor unidade de trabalho. Nesse caso, todos os animais começam no estado susceptível (S), exceto um deles que é aleatoriamente escolhido para iniciar o processo no estado infectado (I). A fazenda é considerada infectada (I) se possui pelo menos um animal infectado em seu rebanho. Durante cada dia de simulação, animais infectados espalham a doença dentro do rebanho do  $j$ -ésimo nó com probabilidade  $\pi_j(t)$ . A doença é transmitida para fazendas susceptíveis através do envio de animais contaminados, com probabilidade  $p_{ji}(t)$ . Os animais se mantêm no estado infectado (I) por um número inteiro de interações,  $\mu$ , e então passam para o estado recuperado (R). Animais que atingiram o estado R se mantêm no mesmo estado até o final da simulação. Nesse tipo de modelo considera-se ainda outra medida de controle da infecção, o abate de animais. Quando um movimento da GTA é classificado como abate, os animais retirados da fazenda de origem são sacrificados, impossibilitando que esse tipo de movimentação cause contaminação na fazenda de destino. Ao final de cada dia, verifica-se a existência de animais infectados em todas as fazendas presentes na simulação. Caso não haja animais infectados em seu rebanho, a fazenda é considerada susceptível (S).

Seja  $S_j^+(t)$  o número de animais susceptíveis na fazenda  $j$  no tempo  $t$  após todas as movimentações do dia  $t$ ;  $I_j^+(t)$  o número animais infectados na fazenda  $j$  no tempo  $t$

após todas as movimentações apuradas no dia  $t$ . Similarmente,  $I_j^-(t)$  representa o número animais infectados na fazenda  $j$  no tempo  $t$  no início do dia (ou seja, antes das movimentações) e  $R_j^+(t)$  o número animais recuperados na fazenda  $j$  no tempo  $t$  após todas as movimentações do dia  $t$ , então pode-se dizer que:

$$I_j^-(t+1) | I_j^+(t), S_j^+(t) \sim \text{Binomial}(S_j^+(t), \pi_j(t)),$$

onde  $\pi_j(t)$  é a probabilidade de contaminar um novo animal na fazenda  $j$  no tempo  $t$  dado que existem  $I_j^+(t)$  animais infectados na fazenda  $j$  no tempo  $t$ .

No modelo de Greenwood (Greenwood, 1931)  $\pi_j(t) = \pi$  é uma constante que não depende do número  $I_j^+(t)$  de infectados. No modelo Reed-Frost (Abbey, 1952)  $\pi_j(t) = 1 - (1 - \pi)^{I_j^+(t)}$ . Dessa forma, o modelo de Reed-Frost considera que a probabilidade de um novo animal se tornar infectado depende do número de animais infectados atualmente presentes no rebanho. Seja  $n_{ji}(t)$  o número de animais transportados da fazenda  $j$  para a fazenda  $i$  no tempo  $t$  e seja  $X_{ji}(t)$  o número de animais infectados em  $n_{ji}(t)$  animais transportados, então pode-se dizer que:

$$X_{ji}(t) \sim \text{Binomial}(n_{ji}(t), \pi_j(t)).$$

A probabilidade  $p_{ji}(t)$  de contaminação da fazenda  $i$  pela fazenda  $j$  no instante  $t$  é definida por:

$$p_{ji}(t) = P(X_{ji}(t) \geq 1) = 1 - (1 - \pi_j(t))^{n_{ji}(t)} \quad (2)$$

Dessa forma, nos modelos de Greenwood e Reed-Frost, a dinâmica de transmissão da doença é descrita por:

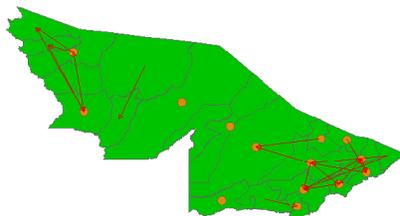
$$\begin{cases} p(\text{fazenda}(i, t+1) = I | \text{fazenda}(i, t) = S) = 1 - \prod_{\text{fazenda}(j, t)=I} (1 - p_{ji}(t)) G_{j,i}(t), \\ p(\text{animal}(u, t+1) = R | \text{animal}(u, t - \mu_u) \neq S) = 1, \end{cases} \quad (3)$$

onde  $\text{animal}(u, t) \in \{SIR\}$  é o estado do animal  $u$  no tempo  $t$ ,  $\text{fazenda}(i, t) \in \{SI\}$  é o estado da fazenda  $i$  no tempo  $t$  e  $p_{ji}(t)$  é descrito em (2).

Uma das inovações trazidas por esse trabalho foi retirar os modelos de Greenwood e Reed-Frost do contexto de populações confinadas (Tuckwell & Williams, 2007) expandindo a sua utilização para cenários em que populações, ou propriedades, estão interligadas por uma rede dinâmica de movimentos.

### 3. Descrição dos dados

Como exemplo, foi utilizada a rede de contatos entre as fazendas do estado do Acre construída a partir das GTAs emitidas no ano de 2006. A Figura 1 mostra a distribuição espacial da movimentação de bovinos no estado. O mapa foi gerado a partir do plug-in “Flow” do software *TerraView*.

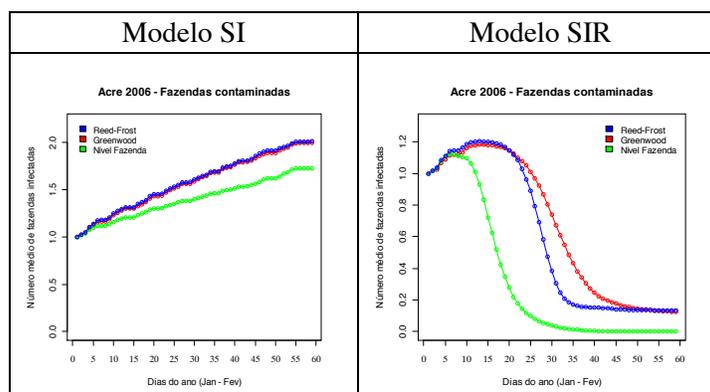


**Figura 1. Fluxo de movimentos entre os municípios do Acre no ano de 2006**

Os pontos vermelhos na Figura 1 revelam a presença de fluxo interno, ou seja, de GTAs emitidas entre propriedades localizadas no mesmo município. As setas indicam a direção da movimentação (origem → destino).

#### 4. Resultados

Os resultados preliminares mostrados nesse artigo são baseados nas movimentações do período de janeiro a fevereiro, totalizando 2220 GTAs; foram realizadas 10.000 simulações para cada dia do período estudado. As simulações foram direcionadas para o estudo do espalhamento da febre aftosa. Em consonância com características dessa doença, nas simulações dos modelos Greenwood e Reed-Frost, considerou-se que  $\pi = 0,4$  (ver equação (3)) e no modelo Nível Fazenda foi utilizado  $p = 0,4$  (ver equação (1)). Para o modelo SIR considerou-se ainda que o parâmetro  $\mu \sim Poisson(15)$ .



**Figura 2. Número médio de fazendas infectadas por dia de simulação**

Os resultados obtidos a partir do modelo SI mostraram que o número médio de fazendas contaminadas por dia nos modelos de Greenwood e Reed-Frost é praticamente o mesmo. Por outro lado, o modelo Nível Fazenda estima valores relativamente mais baixos, especialmente com o crescimento do tempo desde o início da epidemia.

No modelo SIR, observa-se que após 10 dias do início da contaminação, o modelo Nível Fazenda já atingiu o cenário mais crítico da doença. A partir desse ponto o número médio de fazendas contaminadas possui tendência decrescente, chegando à extinção dos focos da doença no decorrer do período simulado. Para os modelos Greenwood e Reed-Frost, a involução da doença começa a ocorrer após 15 dias do aparecimento do primeiro foco, sendo que o decaimento do número médio de fazendas infectadas no Reed-Frost é mais acentuado do que no modelo de Greenwood.

## 5. O aplicativo *EpiGTA*

O aplicativo *EpiGTA* permitirá ao usuário escolher um dos três modelos apresentados anteriormente. As simulações poderão ser realizadas utilizando apenas informações de GTAs, sem dados do rebanho. Caso esteja disponível o número de animais presentes nas propriedades, o programa será capaz de simular cenários mais reais. O aplicativo *EpiGTA* irá gerar resultados na forma de mapas e curvas epidemiológicas (ver também Coelho *et al.*, 2008). A interface sugerida é apresentada a seguir.

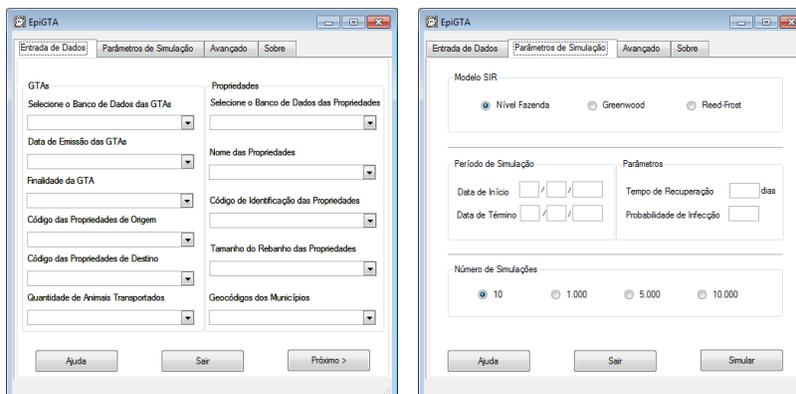


Figura 3. Interface do aplicativo *EpiGTA*

## Referências

- Abbey, H. (1952). An examination of the Reed-Frost theory of epidemics, *Hum. Biol.* 24, 201.
- Bollobás, B. (1979). Graph theory. New York: Springer
- Coelho, F.C., Cruz, O.G. & Codeço, C.T. (2008). Epigrass: a tool to study disease spread in complex networks. *Source Code for Biology and Medicine.* 3:3.
- Greenwood, M. (1931). On the statistical measure of infectiousness, *J. Hyg. Camb.* 31, 336.
- Harary, F. (1969). Graph theory. Reading, MA: Addison – Wesley.
- Keeling, M.J. & Eames, K.T.D. (2005). Networks and epidemic models. *Journal of the Royal Society Interface*, 2, 295-307.
- MAPA, Ministério da Agricultura Pecuária e Abastecimento (2009). Relatório de avaliação do plano plurianual 2008-2011, Ano Base 2008. Brasília, D.F, 91p. Disponível em: <<http://www.agricultura.gov.br>>. Acesso: Março 2010.
- TerraView – Aplicativo para visualização, consulta e análise de dados geográficos – Disponível em: <http://www.dpi.inpe.br/terraview>> Acesso em: Outubro 2010.
- Tuckwell, H.C. & Williams, R.J. (2007). Some properties of a simple stochastic epidemic model of SIR type. *Mathematical Biosciences.* 208, 76 – 97.
- Vernon, M.C. & Keeling, M.J. (2008). Representing the UK's cattle herd as static and dynamic networks. *Proceedings of the Royal Society B*, Published online, 1-8