

Simulador de Aceitabilidade para Produção Cinematográfica

André Albuquerque, Daniel Antunes, Luan Garrido, Juliana Vianna

¹Depto de Ciência da Computação, IM, UFRJ
2290-8091, Rio de Janeiro, RJ

{andrelfa, juvianna}@dcc.ufrj.br, {danoan2008, lbgarrido89}@gmail.com

Abstract. *The subjective character in movie analysis makes any prediction an unreliable one, even when made by experts. The purpose of this paper is to investigate how human relations contribute to the success of the cultural industry, namely, cinematographic productions. This work was based upon the Susceptible-Infected-Recovered (SIR) epidemiological model, the definition of the concept of influence coefficient - extracted from social networks - and the “Peer Pressure” phenomenon.*

Resumo. *O caráter subjetivo presente na avaliação de filmes torna qualquer previsão de sucesso incerta, mesmo quando feita por especialistas. A proposta deste trabalho é investigar como as relações humanas contribuem para o sucesso de produtos da indústria cultural, em especial, as produções cinematográficas. O trabalho foi baseado no modelo epidemiológico Suscetíveis-Infetados-Recuperados (SIR), na definição do conceito de coeficiente de influência - extraído da análise de redes sociais - e do fenômeno “Pressão por Grupo”.*

1. Introdução

O sucesso de um filme, embora muitas vezes possa parecer óbvio, não o é. São muitos os casos em que pequenas produções superam os grandes investimentos de Hollywood [Mlodinow 2009], sugerindo que, embora existam fatores mensuráveis de qualidade, o sucesso não é modelado por uma equação determinística.

A aparente aleatoriedade nos resultados de rentabilidade pode ser entendida analisando-se as características dos produtos da indústria cultural, a qual o cinema integra. Nessa indústria, as interações sociais constituem fator importante para a afirmação de um produto como um sucesso. Um comportamento típico nas relações humanas foi observado em um estudo sobre um mercado virtual de músicas. O fenômeno é denominado “Pressão por Grupo” e consiste no fato de que o grupo com maior número de membros possui, por si só, uma força que atrai novos membros para si [M. J. Salganik and Watts 2006]. No estudo, evidenciou-se que quando uma música obtém uma vantagem inicial (maior número de downloads ou melhor classificação), os próximos participantes são influenciados a ouvir a música mais baixada e avaliá-la de maneira positiva. Resultado semelhante foi observado recentemente no mundo animal [Iain D. Couzin 2011], onde foi constatado que indivíduos ignorantes tendem a seguir o grupo com maior número de membros.

O objetivo do trabalho consiste em analisar, a partir de um modelo de equações diferenciais baseado no modelo SIR, como as informações sobre as produções cinematográficas se difundem por meio de interações sociais e verificar se os fenômenos mencionados anteriormente contribuem direta ou indiretamente para essa difusão.

2. Modelagem

A modelagem SIR foi proposta em 1927 por Kermack e McKendrick [W. O. Kermack 1927] com o objetivo de modelar a propagação de uma epidemia em uma população. O modelo baseia-se na definição de três grupos: Suscetíveis (S), Infectados (I) e Recuperados (R). Neste trabalho, utilizaremos o modelo SIR a fim de estudar a propagação de uma informação e, portanto, foram realizadas algumas modificações no mesmo. Além disso, o modelo proposto possui as seguintes premissas:

1. Os consumidores, quando desconhecem a existência do filme, não influenciam outras pessoas a assisti-lo ou a não assisti-lo;
2. Os espectadores, uma vez que tenham assistido ao filme, formam uma opinião a respeito do produto e passam a influenciar seus amigos de maneira positiva ou negativa. Supomos que os espectadores não mudam de opinião;
3. A evolução do sistema se dá em função das relações ocorridas entre indivíduos que já assistiram ao filme e aqueles que ainda não o assistiram. Ao final, espera-se ter uma estimativa de quantas pessoas assistiram ao filme e de quantas deixaram de assistir.

Cada uma das equações que compõe o sistema proposto estão expostas abaixo. Considere que $s(t) = \frac{S(t)}{P}$ representa a proporção de pessoas suscetíveis num instante de tempo t , onde P representa o número total da população.

$$\frac{dS}{dt} = - \left(\frac{dV}{dt} + \frac{dN}{dt} \right) \quad (1)$$

O grupo V é dividido em outros dois: $L(t)$ e $D(t)$, que representam, respectivamente, as pessoas que gostaram e as que não gostaram do filme:

$$\frac{dV}{dt} = L(t)s(t)Q(t)I_p + eS(t) \quad (2)$$

$$\frac{dN}{dt} = D(t)s(t)Q(t)I_n + eS(t) \quad (3)$$

onde $Q(t)$ representa o número médio de indivíduos diferentes que um membro V_i de V tenha tido contato sem que nenhuma outra pessoa V_j de V (com $j \neq i$) tenha entrado em contato no momento t . Dessa forma, $Q(t)$ é limitado superiormente por $\frac{P}{V(t)}$ e sua expressão é dada por $Q(t) = \rho_1 \frac{P}{V(t)}$ onde ρ_1 é um número aleatório. I_p e I_n são os coeficientes médios de influência das pessoas, positivo e negativo, respectivamente. Finalmente e representa uma taxa referente a influências de fatores externos. No presente trabalho é considerado como influência externa somente o gênero do filme [Nash Information Services 2011]. A expressão de e é dada por $e = \rho_2 Ext(gen)$ onde Ext é uma taxa obtida a partir de estatísticas referentes a preferências de filmes por gênero e ρ_2 é um número aleatório. Para o grupo de pessoas que decidiram não ver o filme, utilizamos $(1 - \rho_2)$.

Para $dL(t)$ e $dD(t)$, temos:

$$\frac{dL}{dt} = M\left(\lambda, \frac{L}{V}\right) \frac{dV}{dt}, \quad \frac{dD}{dt} = M\left(1 - \lambda, \frac{D}{V}\right) \frac{dV}{dt} \quad (4)$$

onde $M(\lambda, rt)$ exerce o papel definido pelo fenômeno “Pressão por Grupo”. Neste trabalho, está sendo proposta uma função exponencial para modelar a função M , chamada grau de majoritariedade, expressa pela seguinte equação.

$$M(\lambda, rt) = \begin{cases} e^{Ax} + (\lambda - 1) & rt \geq 0.5 \\ 1 - M(1 - \lambda, 1 - rt) & rt < 0.5 \end{cases}$$

Sendo $A = \ln(\lambda(\sigma - 1) + 1)$, com $\sigma \in [1, \infty)$ e $x = \frac{(rt-0.5)}{0.5}$, onde σ representa o poder com que o fator aleatório λ será reforçado por M . Recomenda-se $1 \leq \sigma \leq 3$

$$\text{Observa-se que: } S(t) + V(t) + N(t) = P \Rightarrow \frac{dS}{dt} + \frac{dV}{dt} + \frac{dN}{dt} = 0.$$

3. Implementação

Primeiramente, considere um grupo inicial de pessoas que avaliaram o filme em um pré-lançamento, os críticos. Esse grupo inicial divide-se em outros dois: o das pessoas que gostaram do filme (L) e o das que não gostaram (D).

Para ambos os grupos são calculados os coeficientes de influência I_p (influência positiva) e I_n (influência negativa), cálculo esse feito a partir da análise e da definição de métricas da rede social Facebook, especificados de tal modo que é possível avaliar o grau de influência de um membro da rede.

Para cada uma dessas métricas é necessário atribuir um peso relativo a sua parcela de contribuição na determinação do coeficiente de influência. A determinação desses valores foi obtida por meio de uma pesquisa realizada entre usuários do Facebook. Os voluntários deveriam atribuir um valor de 1 a 5 para cada uma das métricas, sendo 5 o mais importante.

O resultado da pesquisa foi usado para o cálculo de médias ponderadas que, por fim, definiram os coeficientes e a função a seguir: 2,6226(taxa de “Curtir”), 3,2930(taxa de “Compartilhamento”), 2,0000(Marcação em fotos), 3,4339(Comentários) e 2,1698(Amigos)

$$\text{Influencia} = \frac{(2.6226)L + (3.2930)S + 2P + (3.4339)C + (2.1698)F}{13.5193} \quad (5)$$

As quatro primeiras são calculadas sempre em relação ao número de amigos, de maneira que haja justiça e uniformidade no cálculo de inúmeros perfis. A última métrica, porém, diz respeito a variável que é tomada como referência, isto é, o número de amigos. Para essa métrica é estipulado um valor limite, considerado como um valor satisfatório para que essa alcance a sua maior contribuição para a determinação de influência do perfil.

É válido ressaltar a existência de inúmeras métricas que poderiam ser usadas[Nani Karthik 2011]. Contudo, optou-se por essas por simplicidade.

A função é aplicada para cada um dos críticos. A média desses valores determina o coeficiente de influência positivo. De maneira análoga, define-se o coeficiente de influência negativo. Vale destacar que os dados da rede social poderiam ser extraídos de uma aplicação criada para o Facebook. Entretanto, neste trabalho os perfis foram gerados de maneira aleatória.

O fator externo revela-se como parte importante da implementação, na medida em que representa uma grande generalização que pode abranger elementos tais como: gênero do filme, elenco, campanha de marketing, entre outros. A expressão definida na implementação baseia-se somente na popularidade de um gênero.

4. Resultados

As simulações executadas no scilab corresponderam com as expectativas no que diz respeito a aleatoriedade dos resultados, respeitando, porém, uma tendência definida pelas condições iniciais.

As simulações abaixo foram iniciadas com os mesmos valores, indicando uma tendência positiva maior do que a negativa. Entretanto, mesmo tendo uma tendência positiva, pode acontecer de a negativa ganhar, o que de fato era o esperado devido aos fatores aleatórios presentes no programa.

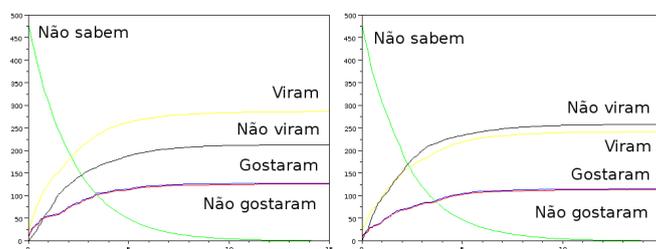


Figure 1. Pessoas que gostaram inicialmente = 11, Pessoas que não gostaram inicialmente = 9, Influência positiva = 0.7, Influência negativa = 0.5, Grau de majoritariedade = 1, População = 500

A simulação, embora não possa afirmar com precisão a real quantidade de pessoas que assistiriam ou não o filme, é capaz de identificar tendências, desde que seja executada diversas vezes.

References

- Iain D. Couzin, Christos C. Ioannou, G. D. T. G. C. J. T. A. H. L. C. S. A. L. N. E. L. (2011). *Uninformed Individuals Promote Democratic Consensus in Animal Groups*.
- M. J. Salganik, P. S. D. and Watts, D. J. (2006). *Experimental Study of Inequality and Unpredictability in an Artificial Cultural Market*. Number 5762. 311 edition.
- Mlodinow, L. (2009). *Peering Through the Eyepiece of Randomness*. Knopf Doubleday Publishing Group, 1st edition.
- Nash Information Services, L. (2011). *US Movie Market Summary 1995 to 2011*. Nash Information Services, LLC.
- W. O. Kermack, A. G. M. (1927). *A Contribution to the Mathematical Theory of Epidemics*.
- Mani Karthik, "How to Measure Facebook Influence.", Internet:<http://www.dailybloggr.com/2010/10/how-to-measure-your-facebook-influence/>, 14 October 2010 [14 December 2011].