

A Contribution of Dynamical Systems Theory and Epidemiological Modeling to a Viral Marketing Campaign

João N.C. Gonçalves^{1(✉)}, Helena Sofia Rodrigues^{2,3},
and M. Teresa T. Monteiro⁴

¹ Department of Production and Systems, University of Minho, Braga, Portugal
jncostagoncalves@gmail.com

² Center for Research and Development in Mathematics and Applications (CIDMA), Department of Mathematics, University of Aveiro, Aveiro, Portugal
sofiarodrigues@esce.ipvc.pt

³ School of Business Studies, Viana do Castelo Polytechnic Institute, Valença, Portugal

⁴ Algoritmi R&D Center, Department of Production and Systems, University of Minho, Braga, Portugal
tm@dps.uminho.pt

Abstract. Nowadays, the interest in analyze and study the behavior of uncontrollable nature phenomena related to the impact of marketing campaigns is an action of prime importance to prevent chaotic dynamics. In this paper we assess the influence of Dynamical Systems theory and Mathematical Epidemiology on a real viral marketing campaign: Dove Real Beauty Sketches, based on a SIR epidemiological model. Motivated by the overwhelming success of this campaign, we study the mathematical properties and dynamics of the campaign real data - from the parameters estimation and its sensitivity to the stability of the mathematical model, simulated in *Matlab*. Mathematically, we show not only that the campaign was a viral epidemic, but also that it can be leveraged and optimized by epidemiological and mathematical modeling, which offer important guidelines to maximize the impact of a viral message and minimize the uncertainty related to the conception and outcome of new marketing campaigns.

Keywords: Marketing campaign · SIR epidemiological model · Optimization · Stability and equilibrium

1 Introduction

Viral Marketing (VM) can be seen as an advertising strategy that benefits from the impact of network effects, mainly of word-of-mouth, within a specific set of

© Springer International Publishing AG 2017
A.M. Madureira et al. (eds.), *Intelligent Systems Design and Applications*,
Advances in Intelligent Systems and Computing 557,
DOI 10.1007/978-3-319-53480-0_96

sofiarodrigues@esce.ipvc.pt

population [1]. Due to the fact that, according recent studies, people tend to trust and be influenced by their social network (e.g. friends, coworkers or relatives) in detriment of standard media, VM is one of the most efficient strategies to face stochastic behavior related to general viral marketing campaigns [1]. Unlike most other marketing techniques, VM is a powerful marketing option to reach the masses with a low cost associated to it [2]. Whereas, in traditional marketing, communication strategies are deeply concentrated in the customer, VM creates mechanisms in order to avoid the implication of the original source, by increasing information exchange among customers and consumers [3]. Moreover, VM has a snowball effect by replicating the message among the susceptible population, increasing the number of infected individuals [4].

However, to be successful, this marketing method requires that the message involved in the process must be attractive to the target audience. Hence, in order to meet the challenge of increasing competition among marketing companies, researchers and marketing professionals have been trying to design marketing advertisements with robustness and resistance to fast followers [5]. In this regard, since the behavior of an epidemic is analogous to VM, the relationship between Epidemiology and Marketing, despite being recent, has been increasingly explored by researchers [6]. As the main motivation, to contribute to the development of this relationship, this paper studies mathematical properties of a specific viral campaign - Dove Real Beauty Sketches - modeled by a SIR epidemiological model under Dynamical Systems and Mathematical Epidemiology theory. As a result, we show some aspects which stand out how can VM campaigns be improved, during its planning phase. In a sense, the choice of this viral campaign is related to the fact that it was a perfect sample of an advertising that combines both solid marketing strategies and resistance to imitation.

Produced in 2013, Dove Real Beauty Sketches is a short movie within the marketing campaign for Real Beauty by Dove. As a viral marketing campaign, Dove Real Beauty Sketches focuses on state the definition of beautiful, promoting self-esteem among women from different age and make themselves safe and confident about their look [7].

The campaign was a huge success across the globe, considered the most viral video advertisement of all time [8], and represents an excellent example of how to build a successful viral marketing campaign. The article is organized as follows. In Sect. 2, supported by Dove Real Beauty Sketches marketing campaign real data, is formulated and discussed the mathematical model. After the model formulation, Sect. 3 presents the parameters estimation, simulated in *Matlab*. In Sect. 4 we study the dynamics of the model, under Dynamical Systems and Mathematical Epidemiology theory. In addition, also in this section, the mathematical behavior is sustained by marketing concepts related to the campaign. Conclusions are carried out in Sect. 5.

2 SIR Epidemiological Model

Taking into consideration that VM can be modeled under standard epidemic models [5,9], we consider a SIR epidemiological model to analyze, over time, the dynamics of the Dove Real Beauty Sketches marketing campaign.

This model describes the variation, in time, of the susceptible, infected and recovered population by the following system of ordinary differential equations, subject to initial conditions related to the campaign real data [8]:

$$\begin{cases} \frac{dS(t)}{dt} = -\beta \frac{S(t)I(t)}{N} \\ \frac{dI(t)}{dt} = \beta \frac{S(t)I(t)}{N} - \gamma I(t) \\ \frac{dR(t)}{dt} = \gamma I(t), \end{cases} \quad \begin{cases} S(0) = 10^9 - 30000 \\ I(0) = 30000 \\ R(0) = 0. \end{cases} \quad (1)$$

$S(t)$, $I(t)$ and $R(t)$ represent, respectively, the mutually-exclusive compartments of susceptible, infected and recovered population at time t . In terms of marketing, we define compartment $S(t)$ as the audience that marketers want to reach with the marketing message; $I(t)$ as the individuals that transmit the message or recommend a product to friends or relatives, motivated by monetary interests or just for the attractiveness of the intended message, and $R(t)$ as the portion of population who stop transmitting the message because, for instance, it is no longer appealing. Regarding to the parameters in terms of epidemiology, $\beta, \gamma > 0$ traduce, respectively, the infectivity and recovery rate. In the marketing context, β can be seen as the predisposition to deal with the message and be motivated to share it, and γ as the interruption of the message transmission by individuals [5]. Regarding to the construction of the model equations, the rate $\beta S(t)I(t)$ results from the probability of transmit the message, within the susceptible population, by an individual who had contact with a susceptible one $\left(\frac{S(t)}{N}\right)$, multiplied by the number of new transmission cases in time (βN) per infected individual ($I(t)$). Over time, individuals stop to share the marketing message according to the rate ($\gamma I(t)$).

The total host population N , estimated by 10^9 [10], is considered constant in time, i.e.,

$$\frac{dN}{dt} = 0 \Leftrightarrow \frac{d(S + I + R)}{dt} = 0 \Leftrightarrow N = S(t) + I(t) + R(t), \quad \forall t \in [0, \infty).$$

Another relevant aspect related to the dynamics of the SIR model is the asymptotic state of the susceptible population. In this sense, marketers should predict the expectable duration time of the success related to the campaign diffusion. Thus, by [11], the evolution of the number of susceptible individuals, over time, can be measured by the following expression

$$\frac{dS}{dR} = \frac{-\beta S}{\gamma N} \Rightarrow S(t) = S(0) \exp\left(\frac{-\beta R(t)}{\gamma N}\right), \quad \forall t \in [0, \infty).$$

Since $I_\infty = 0$, we can also deduce that, over time

$$\lim_{\mathbb{R}_0^+ \ni t \rightarrow \infty} S(t) = S_\infty = S(0) \exp\left(\frac{\beta(S_\infty - N)}{\gamma N}\right), \quad (2)$$

and, consequently,

$$R_\infty = N - S_\infty . \tag{3}$$

These expressions allow marketers to have a global view, in real time, about the evolution of the formulated system. As expected, according to (2), the number of individuals who have not been in contact with the viral message tends to decrease over time, since the number of infected individuals tends to grow up as long as the message continues to be appealing to the diffusion agents. Over time, by (3), when the message is no longer a novelty, the number of people who stop to share it increases. Noting that $R(t)$ derives from $S(t)$ and $I(t)$ and considering $\hat{S} = \frac{S}{N}$, $\hat{I} = \frac{I}{N}$, $\hat{R} = \frac{R}{N}$, the SIR model (1) and its initial conditions can be rewritten as:

$$\begin{cases} \frac{d\hat{S}(t)}{dt} = -\beta\hat{S}(t)\hat{I}(t) \\ \frac{d\hat{I}(t)}{dt} = \beta\hat{S}(t)\hat{I}(t) - \gamma\hat{I}(t) \\ \hat{R}(t) = 1 - \hat{S}(t) - \hat{I}(t) , \end{cases} \quad \begin{cases} \hat{S}(0) = \frac{S(0)}{N} \\ \hat{I}(0) = \frac{I(0)}{N} \\ \hat{R}(0) = 0 . \end{cases} \tag{4}$$

To the marketing message be widely diffused, it is clear that the number of infected individuals must increase, i.e., $\frac{d\hat{I}}{dt} > 0$. Considering $\hat{S} \approx 1$ in this inequality, resulting from the fact that in the beginning of an epidemic almost everyone is susceptible, we get $\frac{\beta}{\gamma} > 1$. The basic reproduction number, $\mathcal{R}_0 = \frac{\beta}{\gamma}$, is defined as the expected number of secondary infections produced by a single infected individual, in a susceptible population, and governs the behavior of the SIR model, hence its importance. In the marketing context this number represents secondary transmissions sent by a single individual to a specific target audience. Moreover, if $\mathcal{R}_0 < 1$ the message is not widespread. On the other hand, if $\mathcal{R}_0 > 1$, the message is broadly sent among the susceptible population [5]. The basic reproduction number is determinant on the system dynamics, as we shall see later.

Note that, by the side of marketers, an accurate estimation of the parameters β and γ deeply helps to monitor the evolution of the campaign among the social network environment. Hence, in the next section, we estimate β and γ in order to adjust the model to the campaign real data.

3 Parameters Estimation Using an Optimization Approach

The estimation of β and γ has a significant importance in the model dynamics and adjustment to the campaign real data. However, it is not trivial to find strategies or recipes to estimate model parameters. Moreover, in the marketing field, this difficult increases exponentially, due to its erratic behavior. Another important aspect is linked to the fact that, related to the word-of-mouth and

network effects processes, social behavior is not predictable with a desired accuracy. So, find a perfect method to estimate parameters related to stochastic patterns it is quite difficult, and marketers must take these critical aspects into account upon the design of future campaigns, in order to have more control on the evolution of it. Although, there are accurate estimation methods among the literature to fit the model to data from an outbreak. Based on [12], for this particular model, we considered Least Squares Parameter Estimation method and `ode45` routine in `Matlab`. Using these optimization techniques, were obtained $\beta = 20.4843$ and $\gamma = 19.7947$ for optimal parameters. This information allows to conclude that $\mathcal{R}_0 > 1$, which represents an epidemic related to the success of Dove Real Beauty Sketches. In addition, in case of the incorporation of control variables into the mathematical model, parameters estimation assumes a key role in order to understand, in mathematical terms, not only the real impact of an advertisement, but also the costs associated to its launch. In the appendix of [13] it is possible to find the sensitivity equations for the SIR model, which constitute an important indicator to study the effect of β and γ variation. In this context, we performed some simulations to assess and analyze the parameters variation (Fig. 1).

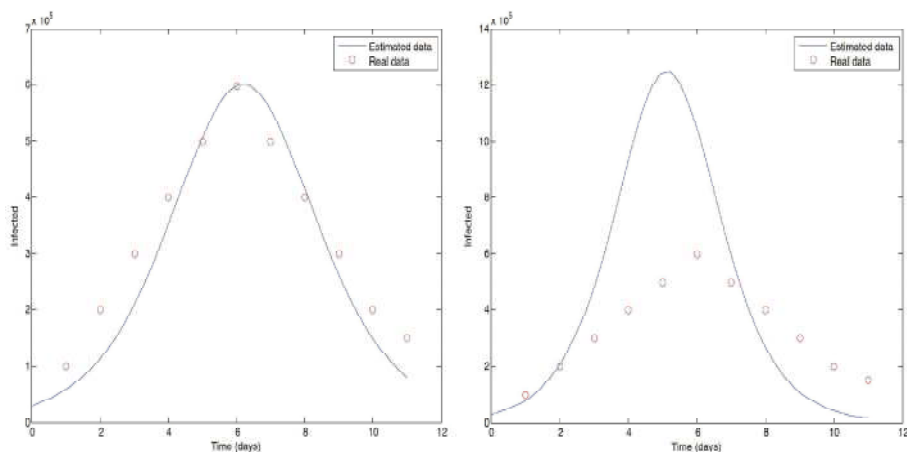


Fig. 1. $\beta = 20.4843$, $\gamma = 19.7947$ (left) and $\beta = 20.1843$, $\gamma = 19.1947$ (right). The effect of β and γ variation in the adjustment to the campaign real data.

Figure 1 (left) displays the estimated data with optimal parameters. Figure 1 (right) presents the estimated data with a smooth variation in β and γ . According to Fig. 1 (left), the optimal values obtained show an almost perfect fitting to campaign real data and it becomes clear that a slight variation on the optimal parameters leads to a misalignment between the estimated and real data (Fig. 1 (right)). So, we can characterize the estimation of β and γ as an ill-conditioned problem. At this point, some guidelines can be given to marketers. With a strong

applicability, the Least Squares Parameter Estimation method leads to accurate results, minimizing the error related to the parameters adjustment to model data. Hence, as well as others, it can be used to manage more efficiently the model dynamics related to future viral campaigns. In this regard, combining with prediction models, a more controlled model dynamic, resulting from an accurate parameters estimation, suggests less costs to the marketers - since the expected margin of error should be residual.

4 Stability and Dynamics of the Model

This section performs a stability and dynamical analysis of the model, by linking it to the Dove Real Beauty Sketches marketing campaign, based on Mathematical Epidemiology and Dynamical Systems results. Henceforth, as a starting point to analyze and discuss mathematical properties related to stability and dynamics of the model, it will be considered the system (4) with its initial conditions, for all $(\widehat{S}, \widehat{I}) \in [0, 1]^2$. Note that the system of ordinary differential equations is autonomous, since \widehat{S} and \widehat{I} do not depend explicitly on time t .

Let $f : [0, 1]^2 \rightarrow \mathbb{R}^2$ be a C^1 map. It is possible to write (4) in the matrix form

$$\frac{dx}{dt} = f(x(t)), \forall t \in [0, +\infty),$$

where

$$x(t) = \begin{bmatrix} \widehat{S} \\ \widehat{I} \end{bmatrix} \text{ and } f(x(t)) = \begin{bmatrix} -\beta\widehat{S}(t)\widehat{I}(t) \\ \beta\widehat{S}(t)\widehat{I}(t) - \gamma\widehat{I}(t) \end{bmatrix}.$$

By solving $f(x(t)) = 0_{\mathbb{R}^2}$ it is concluded that all the points in the form $(S^*, 0)$, where S^* can be any number on $[0, 1]$, are equilibrium solutions of the system. However, special focus is devoted to two disease-free equilibrium points: $E_1 = (1, 0)$, when the levels of infection are null, and $E_2 = (0, 0)$, related to the end of the epidemic, when $\widehat{R} = 1$. In analyzing equilibrium solutions it is pertinent to obtain the linearized matrix of the system (4), in a neighborhood of a critical point (S^*, I^*) , for which $I^* = 0$.

$$\begin{aligned}
 J_f(S^*, 0) &= \left[\begin{array}{cc} \frac{\partial (-\beta\widehat{S}(t)\widehat{I}(t))}{\partial \widehat{S}} & \frac{\partial (-\beta\widehat{S}(t)\widehat{I}(t))}{\partial \widehat{I}} \\ \frac{\partial (\beta\widehat{S}(t)\widehat{I}(t) - \gamma\widehat{I}(t))}{\partial \widehat{S}} & \frac{\partial (\beta\widehat{S}(t)\widehat{I}(t) - \gamma\widehat{I}(t))}{\partial \widehat{I}} \end{array} \right] \Bigg|_{(S^*, 0)} \quad (5) \\
 &= \begin{bmatrix} 0 & -\beta S^* \\ 0 & \beta S^* - \gamma \end{bmatrix}.
 \end{aligned}$$

Considering the characteristic equation $\det(J_f(S^*, 0) - \lambda I_2) = 0$, the eigenvalues of $J_f(S^*, 0)$ are $\lambda_1 = 0$ and $\lambda_2 = \beta S^* - \gamma$. In general, by evaluating the eigenvalues at the equilibrium solutions, we can infer its stability. However, for E_2 , linear stability analysis may not be sufficient. In this particular case are necessary more advanced mathematical techniques to study the dynamics of the system around E_2 . Firstly, considering the optimal parameters obtained in the previous section and the disease-free equilibrium solutions it follows, for E_1 and E_2 , respectively, that

$$\lambda_2 = \beta - \gamma > 0 = \lambda_1 \text{ and } \lambda_2 = -\gamma < 0 = \lambda_1 .$$

In relation to E_1 , since $\Re(\lambda_2) > 0$, according to basic stability properties of linear systems, E_1 is an unstable solution. Establishing a relationship with marketing, this solution corresponds to the case that the number of individuals who recommend the marketing campaign increases faster than it stops being shared (viral marketing epidemic). Considering the equilibrium E_2 , is not possible to conclude, from the linear stability analysis by itself, whether or not the equilibrium solution E_2 is stable. In this regard, the behavior of the autonomous systems as well as its equilibrium solutions can be described through a phase-space. Thus, in a two-dimensional space, the orbits of the system (4) can be described by the equation

$$\frac{d\hat{I}}{d\hat{S}} = \frac{\beta\hat{S}\hat{I} - \gamma\hat{I}}{-\beta\hat{S}\hat{I}} = -1 + \frac{1}{\mathcal{R}_0\hat{S}} \Leftrightarrow \hat{I}(\hat{S}) = -\hat{S} + \frac{1}{\mathcal{R}_0} \log(\hat{S}) + C, \quad C \in \mathbb{R} . \quad (6)$$

This expression enables marketers to predict the final state of the epidemic, from a specific initial state. In this context, through $\frac{d\hat{I}}{dt} = 0$ we can calculate the maximum number of infections, at a time t . This number is obtained when $\hat{S} = \frac{1}{\mathcal{R}_0}$. Thus, by substituting $\hat{S} = \frac{1}{\mathcal{R}_0}$ and $\hat{I} = \hat{I}_{\max}$, the maximum number of infections in time is given by

$$\hat{I}_{\max} = \hat{I}(0) + \hat{S}(0) + \frac{1}{\mathcal{R}_0} \left(\log \left(\frac{1}{\mathcal{R}_0} \right) - \log \left(\hat{S}(0) \right) - 1 \right) .$$

Generated by (6), Fig. 2 shows one orbit of the SIR system and traduces the impact, in time, of the Dove Real Beauty Sketches marketing campaign.

Following an assertive analysis of the Fig. 2, $\hat{S} = 1$ represents the launching of the marketing campaign, where almost nobody knows about it ($\hat{I} \approx 0$). Over time, individuals begin to spread the marketing campaign through network effects and the number of people who have contact with the marketing message reaches a maximum level (viral epidemic), that is

$$\lim_{\hat{S} \rightarrow \frac{1}{\mathcal{R}_0}^+} \hat{I}(\hat{S}) = \hat{I}_{\max} .$$

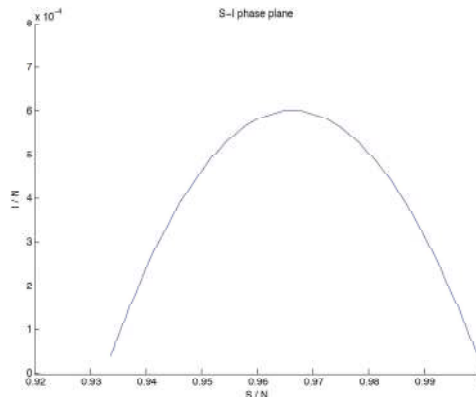


Fig. 2. Orbit of the nonlinear SIR model.

Thus, when a large portion of people have already contact with the campaign, the number of infected people tends to decay to zero, and the number of people who can have contact with the marketing message tends to stabilize around $\widehat{S} \approx 0.93$. It should be noted that the greater the sharing of the message, the greater the number of infected individuals. In the figure below, are illustrated the mathematical results previously analyzed.

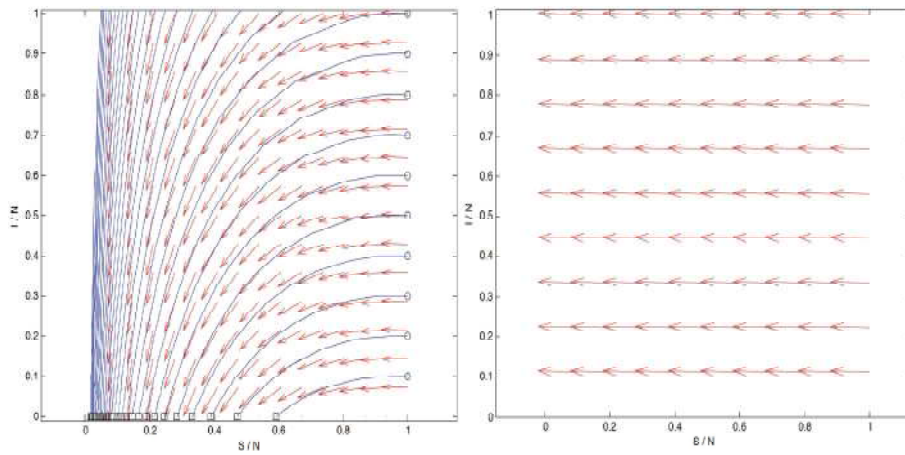


Fig. 3. Dynamics of the nonlinear SIR model (left) and linearized model at E_1 (right).

Figure 3 (left) refers to the nonlinear model (4) and, for each trajectory considered, the initial and final point are marked with a *circle* and a *square*, respectively. Figure 3 (right) displays the vector field of the linearized model, generated by (5) at the equilibrium E_1 . From the phase-plane analysis we infer that E_2

is stable, since all trajectories starting in an arbitrary $a \in B(E_2, \delta)$, for $\delta > 0$, move about a finite range of distance, converging to an ending point $b \in \mathbb{R}_+^2$ nearby E_2 . In other words, the solutions contract back in the direction of E_2 . In what concerns to the vector field showed in Fig. 3 (left), the vectors behavior is perfectly expectable due to the fact that, by the first equation of (4), $\frac{d\hat{S}}{dt} < 0$, since $\hat{S}, \hat{I} > 0$. At this point, $\hat{S} = \frac{1}{\mathcal{R}_0}$ is crucial to understand the dynamics of the system. For starting points $(a_0, b_0) \in \mathbb{R}_+^2$ in a neighborhood of $(S^*, 0)$, for which $S^* \ll \frac{1}{\mathcal{R}_0}$, the trajectories tend to return to a nearby solution $(S_f, 0) \in B(E_2, \delta)$, for $\delta > 0$. Relating to the marketing context, this means that if a small group of persons have contact with the marketing campaign, then it will not be widespread (no viral epidemic).

Concerning to E_1 , as calculated previously, this solution is unstable, since the trajectories grow away from the equilibrium, and computing simulations confirm this findings (Fig. 3 (left)). In other words, supposing that starting points $(a_0, b_0) \in \mathbb{R}_+^2$ are nearby $(S^*, 0)$, for which $S^* > \frac{1}{\mathcal{R}_0}$, the trajectories tend to diverge from the equilibrium, causing a viral epidemic. Hence, these equilibrium solutions are unstable. In terms of marketing, this means that the number of infected individuals - who recommend the campaign - tends to, gradually, grow up, until the maximum level of infection be reached. Then, over time, this number tends to fall down to nearby zero. According to Fig. 3 (right), the instability of E_1 is even more evident, through the vector field, which corroborates that Dove Real Beauty Sketches marketing campaign was a viral epidemic.

5 Conclusions

In this paper, we studied the impact of Dynamical Systems and Mathematical Epidemiology on the dynamics of a specific viral campaign. We considered a SIR epidemiological model and performed a parameters estimation. Some numerical simulations showed not only that Dove Real Beauty marketing campaign was a viral success in terms of marketing strategy, but also that mathematical analysis, applied to viral campaigns, can provide a key to differentiation on the marketing market. We also studied the stability of the two disease-free equilibrium solutions related to the SIR epidemic model. To the marketing business, this model creates an interesting application since it is a cheap and efficient way to measure the stochastic behavior of a marketing campaign. However, some drawbacks can be pointed out to the model: the parameters estimation is not a trivial task, due to the challenge in predict the behavior and dynamics of a viral campaign, and a slight variation on the input data can lead to huge variations on the output.

All in all, the use of Dynamical Systems and Mathematical Epidemiology theory has a very positive impact on the orientation and analysis of successful and general viral marketing campaigns.

Acknowledgments. This work was supported in part by the Portuguese Foundation for Science and Technology (FCT - Fundação para a Ciência e a Tecnologia), through CIDMA - Center for Research and Development in Mathematics and Applications, within project UID/MAT/04106/2013; and through Algoritmi R&D Center, under COMPETE: POCI-01-0145-FEDER-007043 within the Project Scope: UID/CEC/00319/2013.

References

1. Long, C., Wong, R.C.W.: Viral marketing for dedicated customers. *Inf. Syst.* **46**, 1–23 (2014)
2. Mochalova, A., Nanopoulos, A.: A targeted approach to viral marketing. *Electron. Commer. Res. Appl.* **13**(4), 283–294 (2014)
3. Woerndl, M., Papagiannidis, S., Boulakakis, M., Li, F.: Internet-induced marketing techniques: critical factors in viral marketing campaigns. *Int. J. Bus. Sci. Appl. Manage.* **3**(1), 33–45 (2008)
4. Haryani, S., Motwani, B.: Discriminant model for online viral marketing influencing consumers behavioural intention. *Pac. Sci. Rev. B: Humanit. Soc. Sci.* **1**(1), 49–56 (2015)
5. Rodrigues, H.S., Fonseca, M.J.: Can information be spread as a virus? Viral marketing as epidemiological model. *Math. Methods Appl. Sci.* **39**(16), 4780–4786 (2016)
6. Gardner, J., Sohn, K., Seo, J., Weaver, J.: A sensitivity analysis of an epidemiological model of viral marketing: when viral marketing efforts fall flat. *J. Mark. Dev. Competit.* **7**(4), 25–46 (2013)
7. Breen, C.: “Real Beauty” Campaigns in the Media: Dove Campaign for Real Beauty. <http://www.writing.ucsb.edu/sites/secure.lsit.ucsb.edu/writ.d7/files/sitefiles/news/2015%20Breen.pdf>. Accessed 28 Sept 2016
8. Stampler, L.: How Dove’s ‘Real Beauty Sketches’ Became The Most Viral Video Ad Of All Time. <http://static4.businessinsider.com/image/519cd1336bb3f76344000015-960/screen%20shot%202013-05-22%20at%209.48.32%20am.png>. Accessed 28 Sept 2016
9. Leskovec, J., Adamic, L.A., Huberman, B.A.: The dynamics of viral marketing. *ACM Trans. Web (TWEB)* **1**(1), 1–39 (2007)
10. Schroeder, S.: YouTube Now Has One Billion Monthly Users. <http://mashable.com/2013/03/21/youtube-one-billion/#Wv1EUzqZGkqx>. Accessed 28 Sept 2016
11. Britton, T.: Stochastic epidemic models: A survey. *Math. Biosci.* **225**(1), 24–35 (2010)
12. Lloyd, A.: Sensitivity Analysis and Least Squares Parameter Estimation for an Epidemic Model. <http://www4.ncsu.edu/~allloyd/RTG/RTG.Workshop.Parameter.Estimation.pdf>. Accessed 28 Sept 2016
13. Capaldi, A., Behrend, S., Berman, B., Smith, J., Wright, J., Lloyd, A.: Parameter estimation and uncertainty quantification for an epidemic model. *Math. Biosci. Eng.* **9**(3), 553–576 (2012)