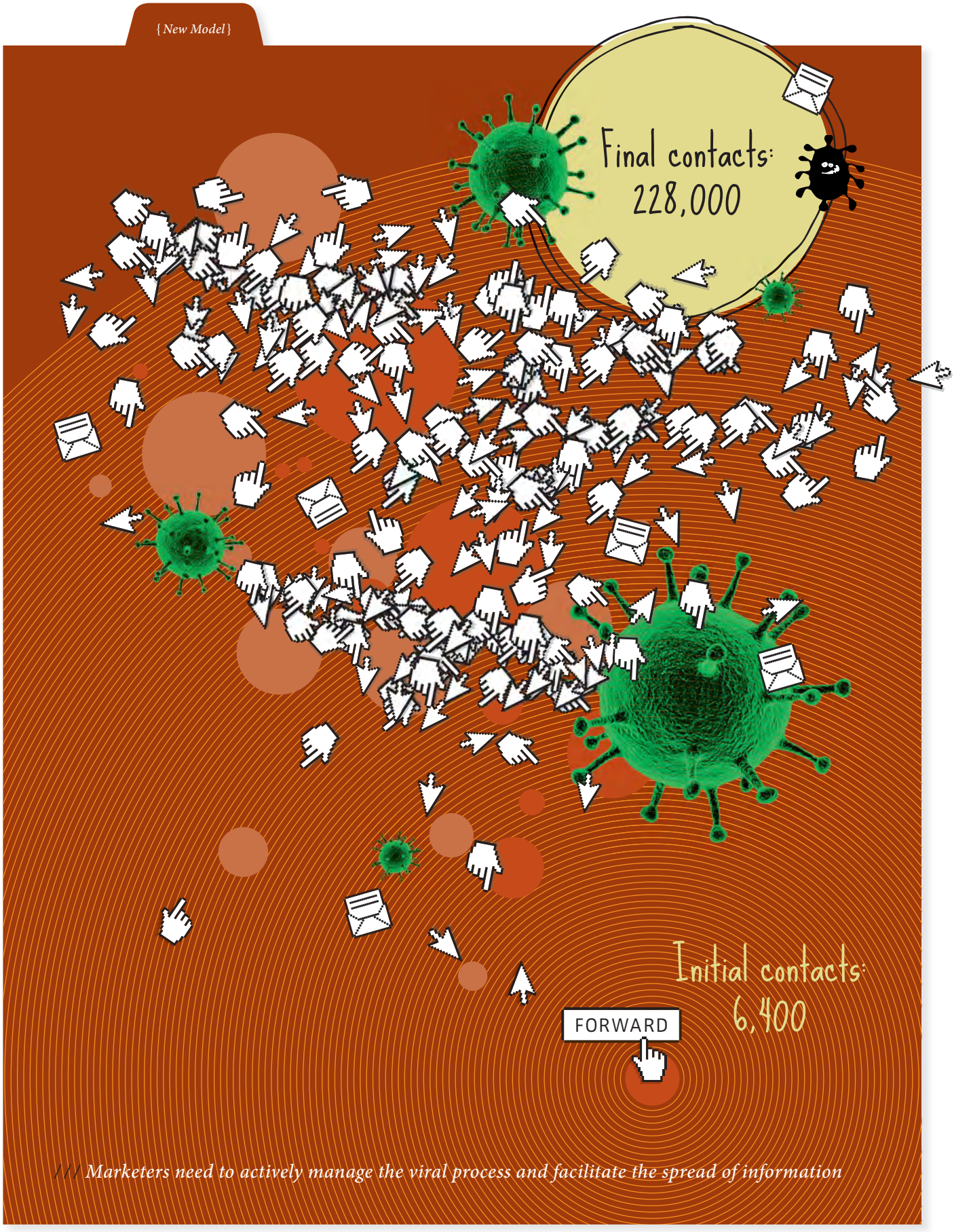
DE GRUYTER — doi 10.2478 / gfkmir-2014-0039
OPEN

{*New Model*}



/// Marketers need to actively manage the viral process and facilitate the spread of information

Seeding a Message to Harvest Reach
PREDICTING AND OPTIMIZING THE SPREAD OF ELECTRONIC WORD-OF-MOUTH

Ralf van der Lans, Gerrit van Bruggen, Jehoshua Eliashberg and Berend Wierenga

In a viral marketing campaign organizations stimulate customers to forward marketing messages to their contacts. To optimize a viral campaign it is necessary to predict how many customers will be reached, how this reach evolves, and how it depends on promotion activities. A new Viral Branching Model can provide these results. It is based on insights from epidemiology and the spread of viruses and was adapted to a marketing context and an electronic environment. The model is applied to an actual viral marketing campaign in which over 200,000 customers participated during a six-week period. The results show that the model quickly predicts the actual reach of the campaign and serves as a valuable tool to support marketing decisions related to online campaigns.

THE AUTHORS

Ralf van der Lans, Associate Professor of Marketing, Hong Kong University of Science and Technology, Hong Kong, rlans@ust.hk.

Gerrit van Bruggen, Professor of Marketing, Rotterdam School of Management, Erasmus University, gbruggen@rsm.nl.

Jehoshua Eliashberg, Sebastian S. Kresge, Professor of Marketing and Professor of Operations and Information Management, The Wharton School, University of Pennsylvania, eliashberg@wharton.upenn.edu.

Berend Wierenga, Professor of Marketing, Rotterdam School of Management, Erasmus University, bwierenga@rsm.nl.

This article is based on the version of *Van der Lans, Ralf; van Bruggen, Gerrit; Eliashberg, Jehoshua; Wierenga, Berend* (2011): „A Viral Branching Model for Predicting the Spread of Electronic Word-of-Mouth", Marketing Science, Vol. 29 (2), 348 – 365.

Tell a Friend ...

In October 2006, Unilever launched a 75-second viral video film ''Dove Evolution''. This campaign generated over 2.3 million views in its first 10 days, and three times more traffic to its website than the 30-second commercial aired during the Super Bowl. More recently, Comic Relief, a British charity organization, achieved 1.16 million participants in the first week after launching their viral game ''Let it Flow'' that promoted Red Nose Day, their main money-raising event. These two examples illustrate a new way of marketing communication in which organizations encourage customers to send emails to friends containing a marketing message or a link to a commercial website. Because information spreads rapidly on the internet, viral marketing campaigns have the potential to reach large numbers of customers in a short period of time. In a viral marketing campaign an organization develops an online marketing message and stimulates customers to forward this message to members of their social network. These contacts are subsequently motivated to forward the message to their contacts, and so on. Such campaigns are powerful because messages from friends have more impact than advertising messages. Not surprisingly many companies such as Microsoft, Philips, Sony, Ford, BMW and Procter and Gamble have gone viral. However, not all viral marketing campaigns are successful and due to competitive clutter, they need to become increasingly sophisticated in order to be effective.

Get Viral Marketing Campaigns Going

The ultimate goal of a viral marketing campaign is that the information spreads automatically ''by itself''. However, marketers need to actively manage the viral process and facilitate the spread of information. It needs to attract customer attention and interest (e.g., by providing videos, online games or other forms of interaction). Tools to easily forward emails to friends, such as ''Tell a Friend'' or ''Share Video'' buttons should facilitate

{ *Box 1* }

MONITORING CONSUMER BEHAVIOR IN A VIRAL MARKETING CAMPAIGN

In order to manage viral marketing campaigns, marketers need to monitor the stages represented in Figure 1 for each individual customer. The following information should be recorded:

1. the source of the invitation
2. if and when a customer arrives at each stage of Figure 1
3. how many friends a customer invites.

This leads to a dynamic database in which each row represents a customer and in which corresponding variables are updated when a customer switches to the next stage. New rows are added when new customers are invited. Such a database can be automatically generated in real time during the process of a viral marketing campaign.

the viral process. Sometimes prizes or other monetary incentives are used to motivate customers to forward messages. Although increasing customers' motivation to forward messages to friends has a strong impact on the reach of the viral campaign, it is necessary to take initiatives to kick-start the campaign by seeding the message in most cases. In general, marketers can choose from three distinct categories to seed their viral marketing campaign:

Seeding emails are usually sent by the company itself or by a specialized marketing agency to customers who have given permission to receive promotional emails. Using this seeding tool, a marketer can target a specific group of customers that are potentially interested in the campaign. The design and content of the emails are crucial since customers easily categorize such emails as spam and quickly delete them. For this reason, seeding emails are expected to be less effective than viral emails sent by friends or acquaintances of the recipient.

Online advertising is another important seeding tool that marketers can use to influence the viral process. The effectiveness of online advertising may differ depending on the customers and the websites on which the ads are placed. Marketers can directly observe when a specific online ad generates a visitor to the viral campaign. Hence, the effectiveness of online advertising can be monitored accurately, and based on its performance marketers can decide to adapt their online advertising strategy. Because online ads may be perceived as less obtrusive than promotional emails, this seeding tool may be very attractive. Further, contracts where organizations pay for each click or guarantee a predetermined number of clicks to the campaign website are comparatively cost-effective.

Finally, besides online seeding tools, marketers may still use ***"traditional" offline advertising*** to seed their campaigns, like magazine and TV ads, package labels or coupons that refer to the website of the viral marketing campaign. However, offline seeding is less popular and expected to be less effective because customers cannot directly visit the campaign website by clicking a link.

Monitoring Campaigns

The appropriate strategic decision on the marketing activities depends on the spread of the process, especially during the initial period and the effectiveness of each marketing communication tool. Therefore, marketers need to closely monitor the spread of information in viral marketing campaigns. It is also important that marketers are able to predict the returns on their expenditures and thus how many customers they will reach (see Box 1).

It gets more and more common to measure the ROI and thus the effectiveness of all marketing campaigns no matter if they concern TV, print or event activities. Being able to measure the ROI of viral campaigns is an important step towards a profound discussion of its optimization and value.

How Consumers Participate in Viral Campaigns

As a starting point of a campaign, a customer receives an invitation from some source to participate, e.g. a viral email from a friend or of one of the seeding tools of a company. At the end of this stage, the customer decides with some probability to go to the second stage and

read the invitation with a certain probability to exit the campaign by deleting or ignoring the invitation. This probability depends on the source of invitation, as customers are less likely to open and read a seeding email from a company than a viral email from a friend. After reading the invitation, a customer decides to accept it with a certain probability by clicking a link to the landing page of the campaign website. After arriving on the landing page a customer decides to participate or not to participate. Participation may consist of watching a video, playing a game, and/or subscribing to a service. Finally, a customer decides to forward the message to a certain number of friends. Figure 1 summarizes this five-stage process that a customer may go through during a viral marketing campaign. The sequence of stages is quite generic for most viral marketing efforts and might have to be adapted to the specific structure of individual campaigns.

Obviously, the number of customers receiving an email is usually not the same as the number of customers who ultimately participate in the viral campaign. The probability of taking the respective next step depend on marketing activities such as the attractiveness of the subject line, the content of the invitation and the design and content of the website (see Figure 1).

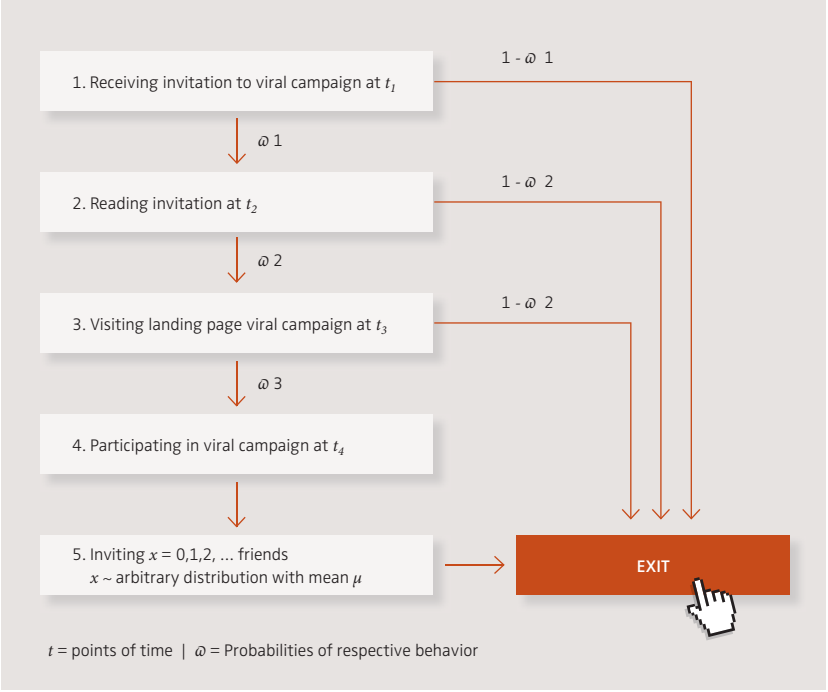


FIGURE 1:
Decision Tree to Participate in Viral Marketing Campaign

The Evolution of Viral Campaigns: Learning from Epidemics and the Spread of Viruses

To effectively manage a viral campaign it is helpful to be able to predict how many customers a viral marketing campaign will reach, how this reach evolves, and how it depends on marketing activities. To understand and model the spread of marketing messages in viral marketing, insights from epidemics about the spread of viruses as well as from biology on the growth of populations over time are useful.

In fact, messages may spread like a virus. An "infected" customer forwards a message to another customer, who may become "infected" as well by joining the campaign. Similarly, the spread of information can also be compared to the growth of families and populations over time. In this case, a female may give birth to children who may give birth to their children, etc. According to these insights, the number of customers should grow exponentially if the average number of "infected" friends exceeds one for each customer that joins the campaign. This average number is called the infection rate and plays a crucial role in the spread of viruses, growth of populations and therefore also in the reach of viral campaigns.

Because detailed data of viral processes has only recently become available, researchers traditionally used deterministic and aggregate level models (diffusion models, e.g., the Bass model) to describe the spread of viruses. They typically assume a specific process, but do not include actual information on this process at the individual level. Such information becomes readily available in viral marketing campaigns and can be used to describe the process accurately. Further, the Bass model, the most prominent "traditional" model in this field, especially in business, assumes that every customer who has adopted the product increases the probability of others adopting in each time period after adoption. However, in viral marketing campaigns customers only influence each other right after participation when they invite their friends.

Therefore, disaggregate level or branching process models using individual-level information are expected to reflect the process more accurately. They assume that customers only influence each other right after partici-

FIGURE 2:
Spread of a Message in a Viral Marketing Campaign as a Branching Process

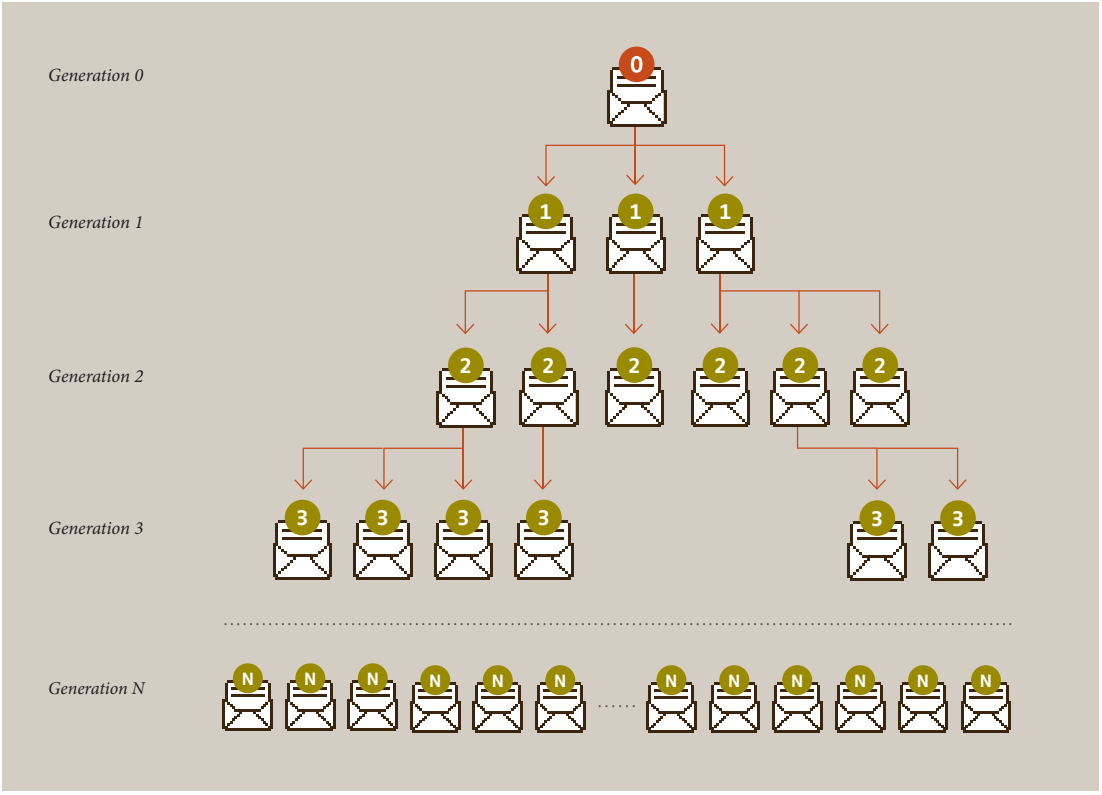
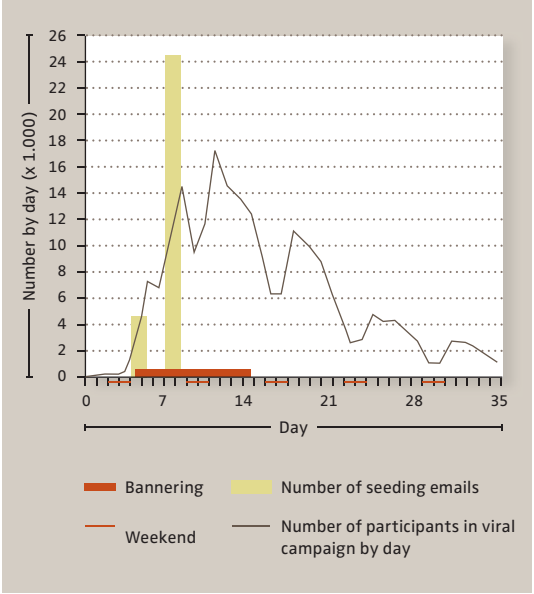


FIGURE 3:
Events and Number of Participants by Day During the Viral Campaign



pation by infecting a fixed number of others. Further, new information can subsequently be introduced into the calculations to infer the reach. Another advantage of using the underlying viral process is that parameters can be estimated based on larger quantities of data.

As a comparison, suppose a small campaign is online for ten days and generates a thousand customers. In this case, deterministic models such as the Bass model would have 10 observations to derive the underlying parameters, i.e. each day serves as one observation. In contrast, the parameters of the branching process are determined by one thousand observations as in this case customers are treated as the unit of observation. More interestingly, however, the parameters of branching processes have a one-to-one relationship with the underlying viral process. Hence, researchers get a much better understanding of why some campaigns are more successful than others.

Viral Marketing as a Branching Process

Figure 2 graphically depicts how information spreads according to a branching process adapted to the specifics of viral marketing campaigns. In this Figure, each email symbol corresponds to one customer in the viral marketing campaign. The viral process is started by one customer in generation 0, and grows rapidly in each subsequent generation. As becomes clear from this Figure, a viral marketing campaign contains two types of customers. First, there are customers who join the campaign as a result of seeding activities of the company. These customers (in Figure 1 there is only one such customer) initiate the viral process by forwarding an invitation to join the campaign by emailing some of their acquaintances. These acquaintances can become customers as well and they may forward an email to invite their friends to join the campaign who again might invite friends.

The key measure in such a branching model is the infection rate, which is the multiplication of the forward rate (the average number of invited customers), and the acceptance rate (mean probability that an invited person joins). Unlike in an epidemic, the infection rate in viral marketing campaigns is generally smaller than 1 which means that the spread of information dies out quickly as each customer generates on average less than one new customer. In such situations, marketers can use marketing tools to influence the viral process by either stimulating the infection rate or by increasing the number of seeded customers.

Developing a Viral Branching Model to Describe and Predict Reach

Although the standard branching model is useful to understand the underlying process in viral marketing campaigns, a more detailed model is needed to accurately describe and predict the actual spread of information. The model parameters are inferred from the individual level click stream data recorded in a dynamically generated database. Based on this daily data, expected results for the key figures of the campaign can be inferred and the number of participants over time can be predicted. The adapted model is characterized by the following characteristics.

{ *Box 2* }

CHARACTERISTICS AND KEY MEASURES OF THE MODEL

The model uses continuous-time Markov processes to predict the status of the campaign at a certain time in the future. Transition times from one state to another follow an exponential distribution. Differential equations determine the values of the interrelated states.

Observed data input

Time and number of seeding emails or banner clicks
Immigration: number of consumers that participate because of other sources

Observed Markov States at each time:
- Unopened seeding emails
- Unopened viral emails (from friends)
- Actual participants (step 4 in figure 1)

Process parameter

- Number of invited friends per participant (mean)
- Probability of accepting an invitation
- Transition time (time between receiving and opening emails)

Forecasting (calculation)

The formulas for the conditional expectation of customers receiving invitations (by seeding instrument or friend) and of actually participating can be used to calculate forecasts for the evolution of the campaign. The individual values can be taken from the dynamically generated database of the campaign. The calculations are easy to implement in a spreadsheet program like Excel.



{ *Box 3* }

PROMOTING FINANCIAL SERVICES IN A VIRAL CAMPAIGN

Campaign and promotion activities:

Customers participated in the campaign while playing a game during which they answered questions that led to a career profile. Then, in return for a guaranteed prize, participants could fill out an online form requesting personal information. After filling out this information, participants were informed that they could win bigger prizes if they invited one or more of their friends to the campaign by sending emails via the "send to a friend" button.

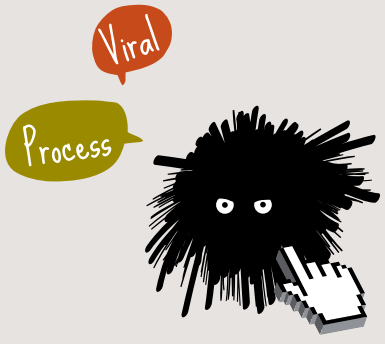
To seed the campaign, the organization bought 6,400 banner clicks to the campaign website. Of the 6,400 visitors, 2,200 people decided to participate in the viral campaign. Furthermore, the marketing agency sent 4,500 (wave 1) and 24,258 (wave 2, 3 days later) seeding mails, to customers who agreed to receive promotional emails.

The viral process:

These marketing activities and the resulting viral process resulted in a total of 228,351 participants by Day 36. Figure 3 summarizes the marketing activities around the viral campaign and the resulting number of participants by day over time. Figure 3 (page 36) shows that the daily number of participants grew rapidly in the first 11 days, after which it slowly decreased over time. The number of participants was lower at weekends, which is due to the fact that during these days customers read their emails less frequently compared to weekdays. Information on each participant was documented in a data set. The 228,351 lines contain the identity of the participant, the date of participation, the source of invitation, the date on which the participant received the invitation, the number of emails that were sent to friends, and how many of these friends already participated or were already invited.

On average, participants sent out over four (4.15) viral emails to friends. The probability that these friends started participating after receiving such an email was 0.26, on average. This led to an average infection rate of 1.08 at the start of the campaign and showed that this particular viral campaign was extremely successful. As the infection rate was larger than one, the number of participants grew exponentially. The proportion of emails sent to customers who had already received an invitation or already participated gradually increased over time. Consequently, at the end of the campaign the average infection rate was smaller than one and equaled 0.87, which means that the number of additional participants decreased over time as shown in Figure 3. However, the infection rate was still substantially higher than other campaigns.

The source of the email strongly influenced its effectiveness. The probability of participation after receiving an email from a friend (0.26) was substantially higher than the probability of participation after receiving a seeding email sent by a company (0.12). Interestingly, the probability of participation after a banner click was even higher than that of customers who received a viral email from a friend (0.34). The bannering approach seemed to benefit from a self-selection mechanism. People who clicked on a banner may have had an interest in the campaign and were then also more likely to participate and send viral emails to their friends. Still, 66 % of these customers decided not to participate and quickly left the campaign's landing page. The source of the email also affected the amount of time that people participate in the viral campaign. This was more than two times shorter when the email was received from a friend rather than from a company (1.64 days vs. 3.88 days during weekdays).



- Customers can participate at any moment in time. Transition times from one stage to another are variable and based on a stochastic distribution. As a result, the model can for instance capture the effect that customers read their emails less frequently during the weekend.

- Second, two different types of marketing seeding activities – banners and seeding emails – are introduced and can be tracked (e.g., the number of unopened seeding emails at each time).

- Third, the model not only counts the number of "infected" customers (i.e., customers who received emails but did not participate further or deleted the emails), but also the cumulative number of customers who actually participate by forwarding the message.

- Parameters are allowed to vary over time to reflect for instance the fact that a participant invites a friend who already received an invitation or already participated. When the campaign progresses, time probability for such instances increases and in turn the overall probability that recipients accept an invitation to participate in the campaign decreases (see Box 2).

Applying the Viral Branching Model in Practice

To test the model it was applied on a real life viral campaign to promote financial services. This campaign illustrates how the model can be used to forecast the reach of campaigns and how marketing decisions can be supported by "what-if analyses" or benchmarking (see Box 3, see Figure 3).

The viral campaign started on a Friday and was online for 36 days. On day 4, the number of participants grew rapidly due to marketing activities. On this day, the company sent 4,500 seeding emails and placed banners on websites that generated 200 participants each day for 11 consecutive days. On day 7, the company sent an additional set of 24,258 seeding emails to further promote the viral campaign.

Forecasting the Spread of the Campaign

All data from the 36 days of the campaign was compared to forecasts that used only the first part of the data. The Viral Branching Model was already able to predict the spread of the campaign quite accurately on day 7, when the campaign was still not fully seeded. After day

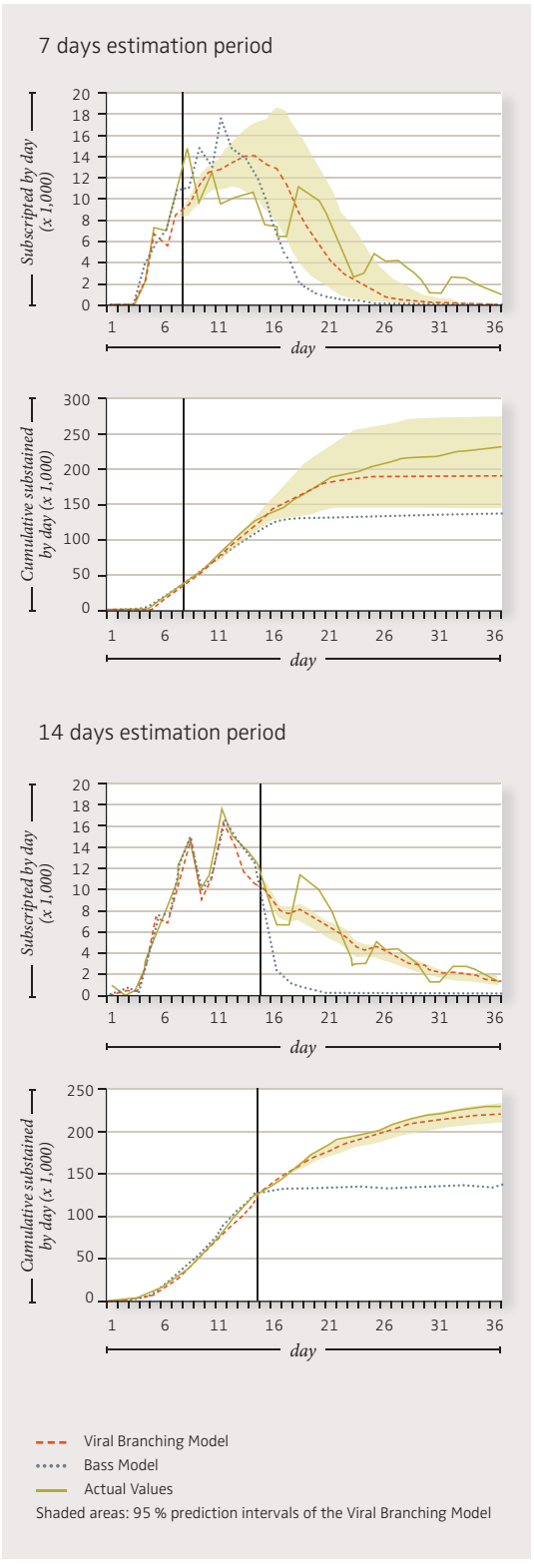
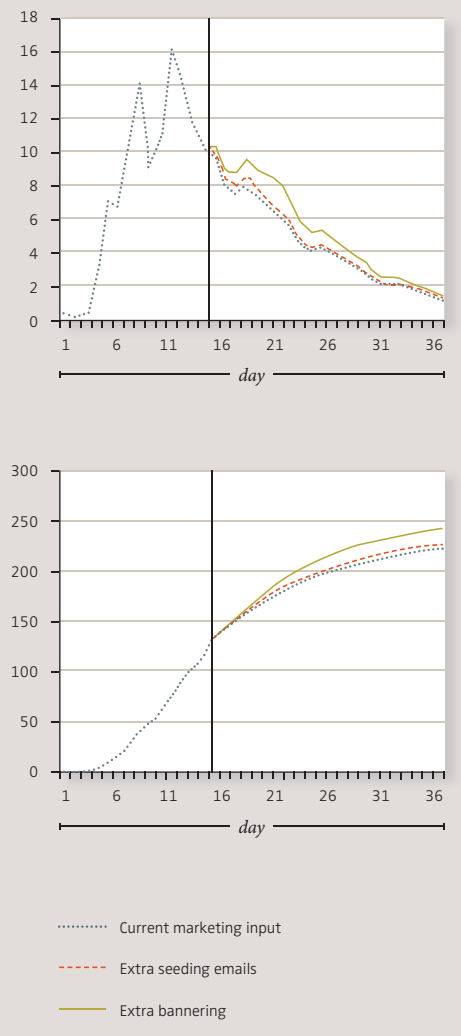


FIGURE 4:
Forecasting the Reach of the Campaign on Day 7 and Day 14

TABLE 1:
Predicted Effects of What-if Scenarios

Marketing activity on day 15	Predicted cumulative number of participants on day 36	Predicted number of additional participants	Predicted number of additional participants per click/seed
Actual marketing strategy	221,429	–	–
Extra bannering for one week: 10,000 clicks	242,595	21,166	2.17 participants/click
Extra seeding: 10,000 emails	227,640	6,211	0.62 participants/seed

FIGURE 5:
Results of What-if Analyses



14, the extended Bass model hugely under predicts at 134,682 whereas the prediction of the Viral Branching Model is 221,429, which is very close to the true ultimate level of 228,351 (see Figure 4). As a matter of fact, the Bass model was not able to accurately predict the number of participants at an early stage of the process, nor was a basic branching model that did not take into account the specifications of the adapted model. Both only start improving at the end of the campaign when the viral process has almost died out and does not attract many new customers. An extended version of the Bass Model starts to predict well only after day 21, the standard model only after day 28.

What-if Analyses: Estimating Marketing Tool Effectiveness

The Viral Branching Model does not only allow for predicting the spread of the viral marketing campaign over time, it also makes it possible to forecast the spread of alternative marketing. Thereby it can support decisions for modifying the campaign in order to reach their objectives. For example, a company might be interested in exploring the effects of two alternative marketing activities to spur on the reach of a campaign. Should they rather use 1) an additional 10,000 seeding emails or 2) an additional 10,000 click-throughs are bought. The seeding emails would be sent out on day 15 of the campaign, the banner would be online for one week from day 15 to day 22. Table 1 summarizes the effects of these two alternative marketing campaigns as of day 36 of the campaign (see Table 1).

The additional 10,000 seeding emails result in an additional reach of 6,211 participants. This means that on average 0.62 additional participants will be reached for

every seeding email. This is the number of people that directly participate by responding to the seeding email and indirectly through receiving a viral email with an invitation from a friend. It is remarkable that the effect of buying 10,000 additional banner clicks is substantially higher. This leads to an additional reach of 21,166 participants at the end of the campaign and means that the additional reach for every click is 2.17. Again, this is the sum of people who start participating directly after they have clicked the banner and subsequently invited contacts through viral emails. These effects originate from the different probabilities of participating after receiving a seeding email and after clicking on a banner. Once the rates of banner clicks and seeding emails are known, a company can determine which seeding method is most cost-effective. The company can also put a dollar value on a customer that participates (e.g. customer lifetime value) to determine if it is profitable to carry out a particular additional seeding. Figure 5 graphically shows the difference in the spread of the campaign if the two alternative scenarios are executed.

Benchmarking

Of course, the reach of a campaign will also depend on the quality of the mailing database and the email itself, the characteristics of the website where the banners are placed and the costs of these seeding tools. Comparing key measures of a campaign like seeding acceptance or forward rates can help to modify a campaign with the right lever by improving individual elements. Further, the switching probabilities from one step to the next in Figure 1 can easily be compared. Depending on where a bottleneck is identified, different activities could improve the situation. If, for example, the forward rate of a campaign falls behind others, it might help to facilitate forwarding or to offer incentives. If the acceptance rate is low compared to other campaigns, it might be more advisable to improve the content or headline of the email.

Summary

Though viral marketing is becoming increasingly popular, it is difficult to predict the success of single campaigns. The viral branching model applied here is a new tool that does very well in predicting the evolution and reach of viral campaigns. Accurately estimating their evolution and reach can help to decide whether additional seeding activities are necessary, which alternatives perform best and how resources can be allocated in an efficient way. •

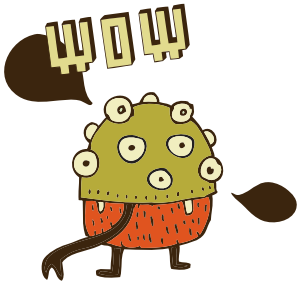
FURTHER READING

Bass, F. M. (1969), "A New Product Growth for Model Consumer Durables", Management Science, 15 (5), 215 – 227.

De Bruyn, A. and G. L. Lilien (2008), "A Multi-Stage Model of Word of Mouth Influence through Viral Marketing", International Journal of Research in Marketing, 25 (3), 151 – 163.

Phelps, J. E., R. Lewis, L. Mobilo, D. Perry, and N. Raman (2004), "Viral Marketing or Electronic Word-of-Mouth Advertising: Examining Consumer Responses and Motivations to Pass Along Email", Journal of Advertising Research, 44 (Dec), 333 – 348.

Watts, D. J. and J. Peretti (2007), "Viral Marketing for the Real World", Harvard Business Review, May, 22 – 23.



KEYWORDS:

Online Marketing, Viral Marketing, Word-of-Mouth, Reach, Branching Processes, Forecasting, Markov Processes