# Agent-based approach to modeling the process of viral marketing

**I H Utakaeva** [1,2]

[1] Financial University under the Government of the Russian Federation, ul. Shcherbakovskaya, 38, Moscow, Russia

[2] utakaev@yandex.ru

**Abstract.** The space of social networks and the Internet challenge classic marketing tools and offers new, modern, effective technologies that improve consumer value. The high rate of technology change poses the problem of effective marketing. In the context of digitalization, new technologies are rapidly forming new market platforms and ways to satisfy their requirements. In this regard, marketers are forced to constantly upgrade their tools and approaches. One effective approach to improving consumer demand is viral marketing. In a highly competitive environment, many more companies are resorting to the viral marketing apparatus. Viral marketing can be based on sound marketing principles, but at the same time become a dangerous tool. Classical epidemic spread models do not take into account the features of modern viral marketing technologies and social networks. The article discusses the concept of viral marketing as a modern tool to increase the consumer value of goods and services. The advantages and disadvantages of this advertising tool are revealed. The main stages of building a viral campaign are described. The conditions under which viral marketing is developing successfully are identified. An agent-graph model of the spread of viral information based on the epidemiological approach that takes into account the basic principles of interaction in a social network is built.

## 1. Introduction

Viral marketing is an impact on the target audience, with the aim of selling goods and services with the help of the target audience itself, which voluntarily participates in the dissemination of viral information about the marketing product. In some cases, viral marketing is more a stochastic process than a planned action. Creating a message to launch viral marketing is not a guarantee of successful marketing; other factors are also necessary, such as the distribution and behavior of a network member. The message becomes viral at the moment when a person passes it further on the network, which means that it should be aimed at a specific audience. Another success factor is that the message should not turn into spam or negative information. This problem is often solved by mailing from a well-known recognized sender, or by a preliminary request for information. Despite vulnerabilities, viral marketing remains attractive for small businesses in the face of companies with a limited budget, which is explained by the ability of this technology to provide good results within a short time and at low advertising costs. The disadvantage of viral marketing technology is the ability to change the content of the message before it is distributed. The viral message should be interesting enough to have a desire to distribute it without creating inconvenience and transmission costs. The real magic of viral marketing begins when consumers spread the message on their own.

Social networks are an effective method of influencing public opinion and have a growing impact on business, politics and society. In social networks, as well as in society, free interaction between people takes place, which is why the viral spread has much in common with epidemic processes.

Identifying the advantages and disadvantages of viral marketing technology, developing a mathematical apparatus that describes the process of viral marketing, identifying the main characteristics that affect the spread of viral messages, as well as the conditions for preventing the spread of spam or unwanted information on the network. The methods of the study are agent-based modeling, analysis, forecasting, imitation modeling tools.

## 2. Models and Results

The advantages and disadvantages of the technology were identified, factors positively and negatively affecting the success of viral marketing were identified, a multi-agent mathematical model was developed that identifies vulnerable vertices and edges of the graph, calculates the epidemic threshold of percolation, predicts the development of the network with certain sets of characteristics, identifies social network clusters infected with viral information.

Communication in social networks and society is often random, but the social network continues to grow dynamically. The main problem is the interest of network participants in viral information, as well as its transmission. As a modeling tool, a pre-fractal graph is selected. A multi-agent simulation model is built. The model formalizes the process of viral marketing with the help of a mathematical apparatus and therefore allows introducing numerical characteristics of the influence on the successful dissemination of information and conduct numerical experiments. A social network, in turn, can be considered as an ordered pair of many agents and many relationships (for example, friendship, contacts on work, acquaintance). Thus, a social network can be described using a graph. Let V be the set of vertices of the graph (the set of participants in the social network - agents), E - the set of edges of the graph (set of connections, interactions, contacts between participants in the social network).

So, in the context of viral marketing, an agent is a basic participant in a social network. In particular, agents are people with a certain social status. Each agent has a set of individual parameters and characteristics that can vary under various conditions and certain factors. For an agent, we define certain numerical and categorical characteristics that together define "marketing status" (level of involvement). Numeric: v is for age, s is for gender, d is for income level. Categorical: b is for speed of reaction, e is for education, q is for marital status, g is for social and professional criterion. Thus, "marketing status" is a characteristic that reflects the resistance of a network member to advertising. However, the same agent may have different "marketing statuses" regarding different products. For example, one person may be interested in products for children and become an active participant in the viral marketing process, and may be a participant in hidden advertising, but will remain indifferent to cosmetics.

An elementary participant in a certain marketing network is an agent in one of two states:

- active agent $p = 1$, they show interest and spread viral information on the network; if we draw a parallel with epidemic processes, then they are "infected" with a marketing virus;

- passive agent $p = 0$, they received a virus message, but do not transmit it further. The state of the agent will be determined by an individual set of characteristics.

In our model, the artificial target audience of agents is in the form of a social network with the same set of properties, but different values. We reproduce the stages of viral advertising. If the new company plans to attract agents and increase consumer demand through viral marketing. At the initial stage, friends, family members, acquaintances who will invite their friends and acquaintances to the network, can act as participants in the marketing network.

The agent model works on the following input data: $V_0$ are agents at the stage of launching viral marketing; $n$ is the number of agents receiving viral information or an invitation from one agent; $V_k$ is the percentage of invitations that later turn into new agents (depends on the individual characteristics of each agent that received the viral invitation).

Each agent of a social network is able to attract a certain number of new customers, we define this indicator as a viral coefficient $K = n * v_k$.

Let the vertex $v_{00}$ be the vertex of the graph (vertex of zero rank) corresponding to the zero participant in the company network, in particular, the owner or director of the company. Then the vertices $v_{0i}, i = \overline{0,n}$ (vertex of the first rank) will correspond to the agents who received an invitation from participant zero. In turn, agents of the first rank can be in two states: active ($\rho\,(v_{0i}) = 1$) or passive ($\rho\,(v_{0i}) = 0$) in relation to a certain marketing product. Sending a virus message from an agent of rank $i = \overline{0,n}$ to an agent of rank $i+1$ will be illustrated on the graph with a directed edge. An ordered pair of the set of vertices W and the set of edges Q, is a graph, $H = (W, Q)$. Let us term it as "seed".
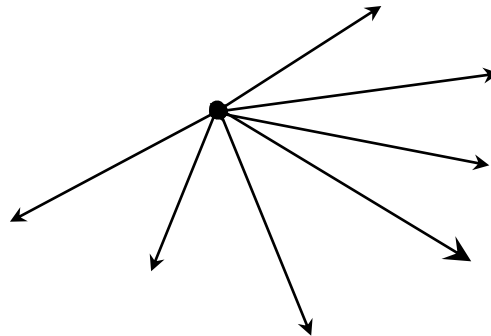


**Figure 1.** Example of a "seed" *H=(W,Q).*

Continuing the process of "infection" with the marketing virus, active agents of the first rank, having become interested in the advertised product, send out new invitations to potential participants in the network of the second rank. Thus, each vertex of the first rank of the graph $H = (W, Q)$,, is connected with *n* vertices of the second rank $v_{1i}, i = \overline{1, n^2}$. Some of the agents who received the invitation will not be interested in the product, will remain inactive and will not send the message further over the network. Continuing the process, each active agent of rank *i* sends out viral advertising information to agents of the next level, etc.

The model developed during the research is implemented in a software product. The program allows the user to:

- simulate the process of spreading viral information on the network;
- take into account the specific marketing status of each agent;
- build a network, taking into account the viral threshold;
- identify the root of the marketing virus or the original source of spam messages;
- identify clusters of viral product infections;
- determine the percolation threshold for a specific marketing virus;
- the ability to determine the necessary quarantine measures in order to stop the spread.

The resulting graph grows dynamically, vertices and edges of new ranks appear, but the graph structure will be stationary, will consist of self-similar parts. Self-similar parts are seeds of various ranks, and the model itself is a pre-fractal graph. At each stage, the active agent that received the virus message forwards it further through the network. This action in the graph describes how to perform the operation of replacing the vertex corresponding to the given active agent by seed. Continuing further the description of the processes occurring at $l \to \infty$, the topology of bonds between agents is described by a fractal graph $G = (V, E)$. It should be emphasized that there may be cases when a person may receive viral product information from several agents. In some cases, having received reviews about a good product from several acquaintances, he will become more trusted in the advertised product and will switch to an active state, and in other cases, on the contrary, obsession can frighten off consumer interest.

Undoubtedly, in practice there are no ideal cases, therefore it is advisable to generalize to the case when the vertex is replaced by a variety of seeds $H = \{H_1, H_2, \ldots, H_s\}$, $s \in Z$.
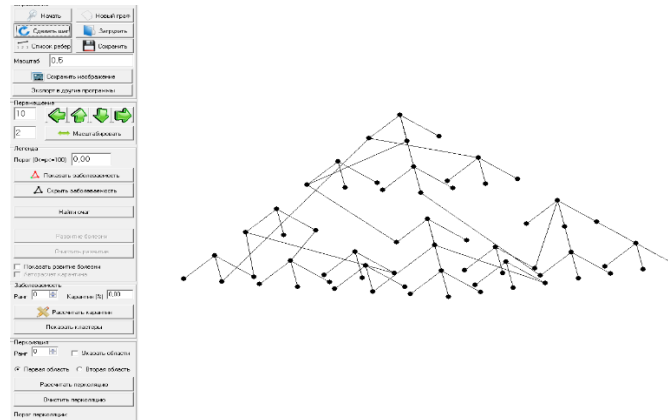


**Figure 2.** Generation of a prefractal graph with a star seed.

For the viral marketing model to work properly, a number of factors need to be additionally considered. In the process of virus infection of the network, the defining behavior of the network is the "virus threshold" (percolation threshold); we will denote it as $\lambda_{\tilde{n}}$. This is the critical indicator of engagement above which the marketing virus will cover the entire network, and a pandemic will occur. A virus with a specific viral score $K < \lambda_n$ propagates exponentially; a virus with $K < \lambda_n$ exponentially "fades". Suppose that a participant in a marketing network has a certain level of involvement $u \geq \lambda_{\tilde{n}}$: having received viral information from an active agent, he will also become active. If the level of involvement $u < \lambda_{\tilde{n}}$, agent remains passive. During the experiment, each vertex corresponds to a set of characteristics that forms the level of involvement. In the graph model, the level of involvement will be set as the weight of the vertex w. Let us assign random vertices to the vertices of the graph using a random number generator.
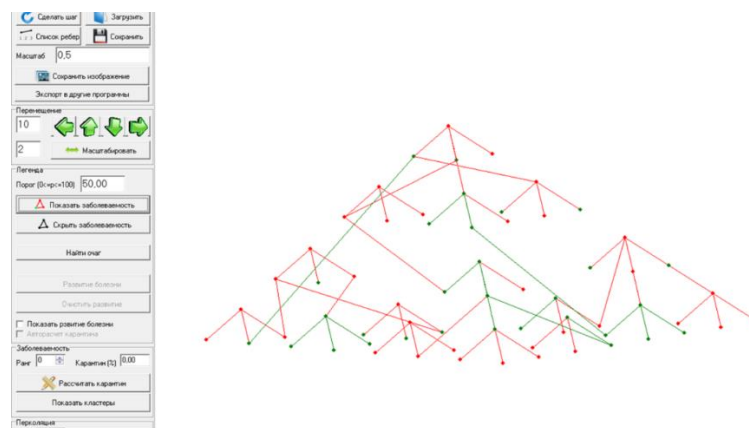


**Figure 3.** Generation of a prefractal graph with a star seed.

Let us we define the concept of network percolation with viral information. Typically, the percolation process is described as "leaking" or "seepage" in a certain direction. For a more specific picture, *drain* and *source* are needed here.

In the theory of percolation, the processes from bottom to top or from left to right are considered, we define the concept of permeability" on a weighted prefractal graph between certain vertices that must

be taken as the source and sink. A "flow" between the two different fixed vertices of the source and the drain on the pre-fractal graph will mean that there is a route connecting them.

### 2.1. Percolation threshold search algorithm

1. $\beta_1 := 1$;
2. searching for routes connecting the source and the drain at $\beta_1$;
3. if there the route exists, return to the previous step at $\beta_1 := \beta_1/2$, otherwise $\beta_2 := \beta_1$, and proceed to the next step;
4. $\beta_1 := \dfrac{\beta_1 + \beta_2}{2}$. If $|\beta_2 - \beta_1| < \xi$, assume $\beta := \beta_1$, the algorithm ends; if $|\beta_2 - \beta_1| \geq \xi$ go to step 2.

Let us denote the resulting numerical value of $\beta$ trough percolation threshold $\lambda_c$.

### 3. Discussion

If the problem of stopping the spread of malicious information on the network is considered, certain measures should be taken. Turning to the model, it is necessary to determine the rank of the seed and indicate the proportion of active participants in the network—nodes at which quarantine measures must be introduced. When the specified rank of the specified share of active agents is reached in the seed, the blocking of all incident edges is applied. These measures are applicable if the viral information is false, panic, or, for example, contributes to the spread of illegal goods and services. In other cases, these edges will be of particular value for the development of the marketing network, and should especially be strengthened.
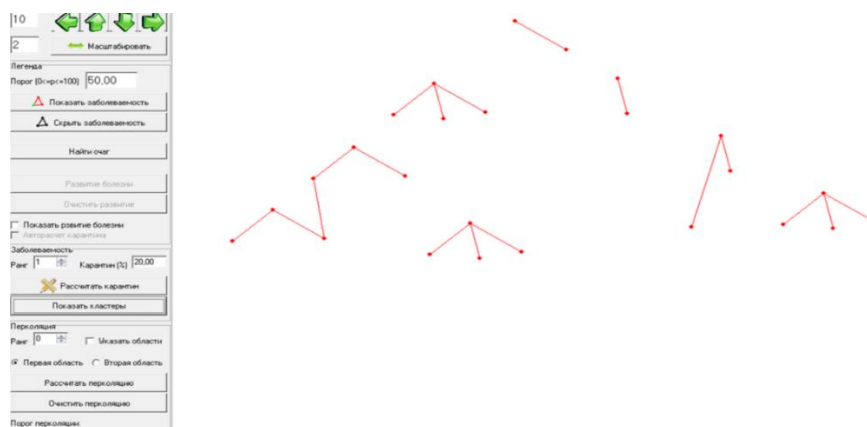


**Figure 4.** Pre-fractal graph under the influence of quarantine measures

Let us consider the distribution structure of advertising information. The source of the virus transmits information. At the stage of $l=1$, the graph structure is a $(n_1+1)$-vertex star $H_1 = (W_1, Q_1)$. Putting the mathematical description of the ongoing process in correspondence with the verbal one, the transition to the next level of distribution will in line with the following rules:

1. graph will be generated by a variety of seeds $H = \{H_1, \dots H_l, \dots H_s\}$, which corresponds to an arbitrary number of infected members of the $r$-rank network;
2. if a certain vertex is not hanging, then the vertex replacement operation is not applied to it;
3. the vertex for performing the vertex replacement operation is selected from the hanging vertices of rank r, taking into account its weight, while the edge becomes incident to the center of the star—the seed;
4. if the hanging vertex is not replaced by a seed, then it is called "frozen", the vertex replacement operation is not applied to it at further stages.

After making $L$ transitions, we get a pre-fractal root tree $G_L = (V_L, E_L)$, which is a tree of distribution of marketing information.
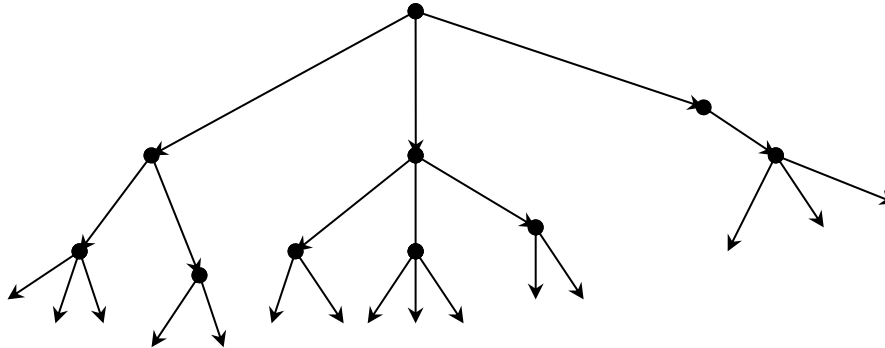


**Figure 5.** Root tree structure

## 4. Conclusion

The importance of viral marketing in today's market economy is difficult to overestimate. Today, it is one of the most effective and tools of marketers, which enables the rapid spread of viral information over the network. The key role of successful product promotion is played by social networks. Thanks to communication on social networks, people can receive not only the virus message itself, but also exchange reviews and discuss the advertised goods and services, and analyze prices and characteristics. For a successful advertising campaign using viral marketing, it is necessary to carefully organize all stages, starting with the development of a quality viral message and choice. If a viral message is created efficiently, then people themselves will be interested in distributing it voluntarily and for free. Viral marketing can be based not only on social networks and the Internet. At the same time, viral marketing, of course, is not an ideal means of advertising, since, like any other advertising tool, it has its drawbacks and problems. A mathematical model of the process of distribution of advertising information in the course of viral marketing of a certain product has been developed. The model is implemented in the form of a software product that solves the tasks posed during the study. The resulting multi-agent simulation model allows simulating the process of distribution of marketing information on the network, calculating the viral threshold and percolation threshold for a specific product and audience, selecting adequate measures to strengthen the most valuable contacts and flows for the distribution of the marketing virus or take measures for quarantine in order to destroy the distribution network.

## References

[1] Chen W, Lu W, Zhang N 2012 Time-critical influence maximization in social networks with time-delayed diffusion process In *AAAI* pp. 592–598.
[2] Gomez-Rodriguez M, Balduzzi D, Scho¨lkopf B 2011 Uncovering the temporal dynamics of diffusion networks In *ICML* pp. 561–568,
[3] Gomez-Rodriguez M, Scho¨lkopf B 2012 Influence maximization in continuous time diffusion networks In *ICML* pp. 1–8.
[4] Grifoni P, D'Andrea A, Ferri F 2013 An integrated framework for on-line viral marketing campaign planning *International Business Research* **6**(1) 22–30.
[5] Krause A, Golovin D 2013 Submodular function maximization *Tractability*.
[6] Van Leeuwen M, Ukkonen A 2015 Same bang, fewer bucks: efficient discovery of the costinfluence skyline. In *SDM*.
[7] Liu B, Cong G, Xu D, Zeng V 2012 Time constrained influence maximization in social networks. In *ICDM*.

[8]   Liu B, Cong V, Zeng Y, Xu D, Chee Y M 2014 Influence spreading path and its application to the time constrained social influence maximization problem and beyond *TKDE* **26**(8)1904–1917.

[9]   Lu W, Bonchi F, Goyal A, Lakshmanan L V S 2013 The bang for the buck: Fair competitive viral marketing from the host perspective. In *KDD* pp. 928–936.

[10] Moore S 2015 Film Talk: An investigation into the use of viral videos in film marketing *Journal of Promotional Communications* **3**(3) 380–404.

[11] Utakaeva I K 2019 Mathematical model of viral marketing *Journal of Physics: Conference Series* **1353**(1) 012122.

[12] Utakaeva I K Directions and features of application of the blockchain technology *Journal of Physics: Conference Series* **1353**(1) 012103.