An IS_mS_nR marketing information spreading model in online social networks considering spammer existence

Zhi-Hong Li, Yun-Bei Zhuang

School of Business Administration, South China University of Technology, Guangzhou, Guangdong, China

Abstract **— Modeling information spreading with spammer consideration is becoming increasingly important, since spammers can interfere individuals' online shopping decisions. To solve this problem, an Ignorants-Spammers-Spreaders-Stiflers (IS_mS_nR) model in this paper, which extends the classical Susceptible-Infected-Removed (SIR) rumor spreading model by spiting the spreaders into spammers and normal spreaders and defining the infectious probabilities as functions of user influence. Then, a mean-field equation that describes the dynamics of the IS_mS_nR model is derived. Besides, a steady-state analysis is conducted to analyse the stability of the equilibrium solutions. Furthermore, numerical study is conducted to investigate the final information spreading size under different initial conditions, groups' user influence, decay rates, and average degree of the network. The results based on a numerical study show that: 1) the information diffusion of marketing under spammer existence shows a normal spreading trend as other circumstances, which is without spammers; 2) the number of the initial seeds, the average degree, the groups' user influence and the decay rates not only can affect the spreading size of the information diffusion, but also can affect the spreading speed and period. Those findings offer a preliminary implication to marketing managers on how they can increase their promotional profit through online networking marketing, and also provides a novel approach to predict the information diffusion process.**

Keyword - Information spreading; Marketing; Normal spreaders; Online social networks; Stiflers; Spammers

I. INTRODUCTION

With the rapid proliferation of economy and technology, online shopping is playing an increasingly important role in daily life. According to a survey conducted by "CNNIC"[1], the number of Chinese online shopping users has reached 347 million, which accounts for 45% of the total marketing, as of June 2015. Besides, the total amount of online shopping has reached ￥2814.51 billion, which accounts for 10.6% of the total sales of consumer goods over the same period. Among this total amount, social commerce attributes to ￥96 billion, 33.8% of the total online shopping users once purchased through online social networking sites (OSNs) (such as, We-chat, Sina weibo etc.), and other 27.8% users once spread sales information through OSNs. Therefore, social commerce is a useful supplement of online shopping and might be of substantial value to electronic marketers.

"Word-of-Mouth(WoM)" is the main factor influencing consumers' purchase decision, CNNIC shows 71% buyers would firstly consider the goods' "WoM" when shopping on OSNs [1]. However, with the ease of access and low awareness of protection, spammers are flooding in OSNs, who are mostly employed by "Public Relation (PR)" firms, sending kinds of spams, such as rumors, ads, pornographic information or phishing links, through repetitively copy and paste, to achieve specific PR goals [2,3]. There has been over 1,614,860 users who once spread spam in “Sina weibo” according to their official service account[2], as of September 2015. Almost every celebrity user has been attacked by spammers. The existence of spammers in OSNs brings both positive and negative effect to social commerce. On one hand, spammer behaviour is threatening users' mutual trust and loyalty to the brand [4,5], affecting users' decisions, especially purchasing decision and recommending decision [6].

On the other hand, spams spread faster than normal information through multiple spammers' collaboration, thus spams on a fixed brand/firm can attribute to the brand awareness [5]. At this juncture, modeling information diffusion [7,8] with spam existence is becoming an urgent challenge.

To date, considerable work on information diffusion of OSNs has been undertaken [9-12]. These works are mainly based on epidemic spreading and complex networks theories [13]. Classic models include Susceptible-Infected-Removed (SIR) model [14], Susceptible-Infected-Susceptible (SIS) model [15], Linear Threshold (LT) model [16] Independent Cascade (IC) model [17]. Inherited from the SIR model, Daley and Kendall proposed the basic DK model, the beginning of rumor spreading modeling, in the 1960s [18] They divided the population into three groups: spreaders, ignorants, and stiflers, which represent those who spread the rumor, those who are unaware of the rumor, and those who are aware of the rumor but choose not to spread, respectively. From 1960s to 2010s, with the development of complex networks, many researchers began to modify the states, the spreading mechanism or the transition probability between different states of DK or SIR model, to apply to different types of network (small-world network [19], scale-free network [20]). For instance, Laijun Zhao et al. modified a flow chart of the rumor spreading process with the SIR, and they believe that ignorants will inevitably change their status to spreaders or stiflers once they are made aware of a rumor. In addition, the probabilities that a spreader becomes a stifler are differentiated in accordance with reality [21]; Jun-Jun Cheng et al. modified the infectious probability of DK model from ignorants to spreaders as a function of the strength of ties, and simulated the model with dataset from a social blog directory website [13]; Liang'an Huo et al divided the states into ignorants-incubations-spreaders-stiflers, where incubation means those who know the information but need a period of time to understand and decide whether to spread or not [22].

However, an important shortcoming of the above group of information diffusion models is that their researchers believe the spreaders would be stiflers when they encounter someone (spreader or stifler) who has known the information with a fixed probability. While in fact, there are a group of spammers who maintain to be spreaders until they are informed against to the community management agencies or lost interest to the PR work. Those spammers have not been studied before in information diffusion of OSNs. Besides, most of the above mentioned studies (except the work of Jun-Jun Cheng et al. [13]) regard the transition probability as a constant, which is impractical because the transition probability can be influenced by many factors, such as participants' personal attributes, mutual trust, attention degree, activity degree and so on.

In order to alleviate above mentioned shortcomings, this paper extends SIR model by dividing the spreaders into spammers and normal spreaders. Spammers indicate persons who only spread spams until be informed or lost interest to the message, normal spreaders indicate persons who only spread compliant information and inform spammer's illegal behaviour to the community management agency whenever he/she encounters and identifies a spammer. Besides, the infectious probabilities of spreaders (spammers and normal spreaders) are modified by functions of their average user influence, which is a combination of various factors. Thus, the Ignorants-Spammers-(normal)Spreaders-Stiflers (IS_mS_nR) model is proposed. Furthermore, we derive the mean-field equations that describe the dynamics of the IS_mS_nR model, and analyze the steady-state of the model in Section 2. In Section 3, numerical simulation is investigated to analyze the impact factors under different parameters. Conclusions and discussions are given in Section 4.

II. MATERIALS AND METHODOLOGY

In this section, we present our "Ignorants-Spammers-(normal)Spreaders-Stiflers (IS_mS_nR)" Model, and provides its equilibrium solutions with steady analysis.

A. *The Model*

Consider an online social site, $G=(V,E)$ consisting of N ($|V|=N$) individuals, which can be subdivided into four groups including ignorants, spammers, normal spreaders, and stiflers, and their population number are described in terms of $I(t)$, $S_m(t)$, $S_n(t)$ and $R(t)$, at time t ($t\in(0,T]$). Besides, at time t, there are only two kinds of marketing information, one kind are spams (such as ads, URLs with virus and messages that libel or slander competitors), the other kind are non-spams (normal messages on that marketing topic, such as user sharing experience):

(A) $I(t)$ ignorants who are ignorant of the marketing information, until they are infected by normal spreaders or spammers (initially, $I(0)=N-K_1-K_2$);

(B) $S_m(t)$ spammers who are employed by PR firms, they maintain to spread spams until being informed against by other persons or lose interest to PR work (initially, $S_m(t)=K_1$);

(C) $S_n(t)$ normal spreaders who are interested on the marketing topic and spread normal information. Besides, normal spreaders can inform spammer's illegal behaviour to the community management agency with a probability, when he/she encounters and identifies a spammer (initially, $S_n(t)=K_2$);

(D) $R(t)$ stiflers who are aware of the marketing information(both the normal information and the spam) but choose not to spread (initially, $R(0)=0$).

(A)and(D), whose members are called ignorants and stiflers [18], respectively, are similar to the classes of susceptibles and "removed cases" in the model commonly used in epidemic theory.

As the population size, N, is fixed, at any time t:

$$I(t)+S_m(t)+S_n(t)+R(t)=N \qquad (1)$$

According to the perception and reaction of a individual to a marketing information on the OSNs, the IS_mS_nR model for the marketing information spreading process is shown in Fig. (1).

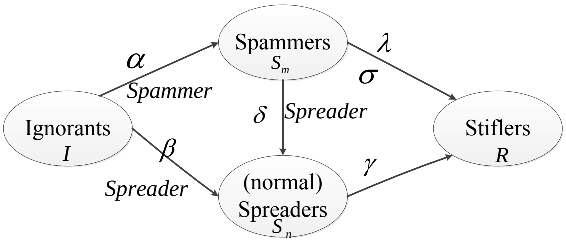

Fig.1 IS_mS_nR Model for The Marketing Information Process with Spammers Existence

As shown in Fig. (1), the contacts between ignorants, spammers, normal spreaders and stiflers are governed by the following parameters:

1) α indicates the infectious probability of an ignorant by spammers, which is the ratio of an ignorant becomes a spreader when the PR firm successively induces him/her to spread spams;

2) β is the infectious probability of an ignorant by normal spreaders, which indicates the probability of an ignorant becomes an normal spreader when the ignorant contacts with a normal spreader;

3) δ indicates the conversion probability of a spammer, which is the probability of a spammer becomes an normal spreader when the normal spreader contacts with a spammer. This implies an assumption that some spammers who transform from ignorants can be corrected to be normal spreaders with interaction with normal spreaders;

4) λ indicates the inform probability of a spammer, which is the probability of a spammer becomes a stifler when an normal spreader contacts with a spammer, detect the spammer and inform against it to the community management agency;

5) σ indicates the decay rate of a spammer, which is the decay rate of an spammer becomes a stifler, based on the assumption that individuals may halt the behavior of information spreading spontaneously just because they could lose their interests on the PR work or forget to spread the message or criticized by forces apart from G such as news report);

6) γ indicates the decay rate of an normal spreader, which is the probability of an normal spreader becomes a stifler because they lose their interests on the message or they forget to spread the message, where $\gamma \le \sigma$.

As we all know, the contacting probability between two individuals is affected by various factors, such as the network structure, such as interactivity, network structure and other attributes. In order to integrate those factors, we apply the initial average user social influence of ignorants, spammers, normal spreaders and stiflers, respectively, to measure the above mentioned parameters as follows:

$$\alpha(t) = \alpha_0 \frac{\bar{k}}{N}$$

Where $\alpha_0 = \dfrac{\sum_{v\in(B)} UI_0(v)}{\sum_{v\in V} UI_0(v)}$, $UI_0(v)$ is the user influence of node v, $\bar{k}$ is the average degree of G and α_0 is the average user influence of spammers) at the initial time. We suppose that a spammer can contact $\bar{k}$ participants per unit time, thus at time t , a spammer can contact $\dfrac{\bar{k}}{N}$ ignorants and can infect $\alpha(t) = \alpha_0 \dfrac{\bar{k}}{N}$ ignorants per unit time.

Similarly,

$$\beta(t) = \beta_0 \frac{\bar{k}}{N} = \frac{\sum_{v\in(C)} UI_0(v)}{\sum_{v\in V} UI_0(v)} \bullet \frac{\bar{k}}{N}$$

$$\lambda(t) = \delta(t) = \delta_0 \frac{\bar{k}}{N} = \frac{\sum_{v\in(C)} UI_0(v)}{\sum_{v\in(B)} UI_0(v)} \bullet \frac{\bar{k}}{N}$$

$$\sigma = \eta + \theta$$

where η is the probability of a spammer becoming a stifler because he/she lost interest on the information or forget to spread the message. θ is the parameter to measure the forces of government and news media on this transition between spammers and stiflers. In this paper, we assume all the parameters are normalized into $(0,1]$.

In the light of the IS_mS_nR marketing information spreading process elaborated as Fig. (1), the mean-field equations can be acquired as follows:

$$\begin{cases} \dfrac{dI}{dt} = -\alpha I(t)S_m(t) - \beta I(t)S_n(t) \\ \dfrac{dS_m}{dt} = \alpha I(t)S_m(t) - (\delta+\lambda)S_m(t)S_n(t) - \sigma S_m(t) \\ \dfrac{dS_n}{dt} = \beta I(t)S_n(t) - \delta S_m(t)S_n(t) - \gamma S_n(t) \\ \dfrac{dR}{dt} = \lambda S_m(t)S_n(t) + \sigma S_m(t) + \gamma S_n(t) \end{cases} \quad (2)$$

From Equation (1), we can get $R(t) = N - I(t) - S_m(t) - S_n(t)$, thus the main equation can be determined by the above first three equations in Equation (2), with the initial condition:

$I = N - K_1 - K_2, S_m = K_1, S_n = K_2$.

B. Steady State Analysis of IS_mS_nR Model

The model IS_mS_nR considered as Equation (2), has the following three possible non-negative equilibrium solutions: $E_0 = (I_0, 0, 0, N - I_0)$; $E_1 = (I_1, S_{m1}, S_{n1}, N - I_1 - S_{m1} - S_{n1})$;where $I_0 \ge 0$ is a real number,

$$I_1 = \frac{N(2\alpha_0\gamma - \beta_0\sigma)}{\alpha_0\beta_0\bar{k}}, \quad S_{m1} = \frac{N(\sigma\beta_0 - \gamma\alpha_0)}{\alpha_0\beta_0\bar{k}},$$

$$S_{n1} = -\frac{\alpha_0}{\beta_0} S_{m1} .$$

Since α_0 , β_0 , σ , γ are in the range of $(0,1]$, and $I(t)$, $S_m(t)$, $S_n(t)$, $R(t)$ are non-negative numbers of population at time t , as mentioned, we derive the existence condition of E_1 (obviously, E_0 always exists and when $I_0 \ge 0$). In fact, E_1 's existence equals to:

$$I_1 = \frac{N(2\alpha_0\gamma - \beta_0\sigma)}{\alpha_0\beta_0\bar{k}} \ge 0 \text{ , } S_{m1} = \frac{N(\sigma\beta_0 - \gamma\alpha_0)}{\alpha_0\beta_0\bar{k}} \ge 0 \text{ ,}$$

$S_{n1} = -\frac{\alpha_0}{\beta_0} S_{m1} \ge 0$ and $N - I_1 - S_{m1} - S_{n1} \ge 0$, that is $\gamma = \frac{\beta_0\sigma}{\alpha_0}$ and $\sigma \le \bar{k}\alpha_0$.

Therefore, $E_1 = (0,0,0,N) \in E_0 = (I_0, 0, 0, N - I_0)$, and Equation (2) has a group of equilibrium solutions as $E_0 = (I_0, 0, 0, N - I_0)$ ($I_0 \ge 0$).

Suppose $E^* = (I^*, S^*_m, S^*_n, N - I^* - S^*_m - S^*_n)$ is an equilibrium solution of Equation (2), and let $X = I - I^*$, $Y = S - S^*_m$, $Z = S - S^*_n$, then we can get the following Equation (3) from Equation (2), where the equilibrium solution $E^* = (I^*, S^*_m, S^*_n, N - I^* - S^*_m - S^*_n)$ of Equation (2) equals to the zero solution $E' = (0,0,0)$ of Equation (3).

$$\begin{cases} \frac{dX}{dt} = -\frac{\bar{k}}{N}\Big[(\alpha_0 S^*_m + \beta_0 S^*_n)X(t) + \alpha_0 I^* Y(t) + \beta_0 I^* Z(t) \\ \quad + \alpha_0 X(t)Y(t) - \beta_0 X(t)Z(t)\Big] \\ \frac{dY}{dt} = \alpha_0 \frac{\bar{k}}{N} S^*_m X(t) + (\alpha_0 \frac{\bar{k}}{N} I^* - 2\beta_0 \frac{\bar{k}}{N} S^*_n - \sigma)Y(t) \\ \quad - 2\beta_0 \frac{\bar{k}}{N} S^*_m Z(t) + \alpha_0 X(t)Y(t) - 2\beta_0 Y(t)Z(t) \\ \frac{dZ}{dt} = \beta_0 \frac{\bar{k}}{N} S^*_n [X(t) + Y(t)] + [\beta_0 \frac{\bar{k}}{N}(I^* + S^*_m) - \gamma]Z(t) \\ \quad + \beta_0 \frac{\bar{k}}{N}[X(t) + Y(t)]Z(t) \end{cases} \tag{3}$$

From Equation (2) and (3), we can derive the following theorem.

Theorem 1: When $I_0 \in \left[\max(\frac{\sigma N}{\alpha_0\bar{k}}, \frac{\gamma N}{\beta_0\bar{k}}), \frac{(\sigma+\gamma)N}{(\alpha_0+\beta_0)\bar{k}}\right]$, $E_0 = (I_0, 0, 0, N - I_0)$ of Equation (2) is global asymptotically stable.

Proof: The Jacobian coefficient matrix of Equation (3) is:

$$V = \begin{pmatrix} v_0 & -\frac{\bar{k}}{N}\alpha_0 I^* & -\frac{\bar{k}}{N}\beta_0 I^* \\ \frac{\bar{k}}{N}\alpha_0 S^*_m & v_1 & -2\frac{\bar{k}}{N}\beta_0 S^*_m \\ \frac{\bar{k}}{N}\beta_0 S^*_n & \frac{\bar{k}}{N}\beta_0 S^*_n & v_2 \end{pmatrix} \tag{4}$$

where,

$$v_0 = \frac{\bar{k}}{N}(\alpha_0 S^*_m + \beta_0 S^*_n) \text{ , } v_1 = \frac{\bar{k}}{N}\alpha_0 I^* - 2\frac{\bar{k}}{N}\beta_0 S^*_n - \sigma$$

$$v_2 = \frac{\bar{k}}{N}\beta_0 I^* + \frac{\bar{k}}{N}\beta_0 S^*_m - \gamma .$$

For $E_0 = (I_0, 0, 0, N - I_0)$, from Equation (4), we can obtain its Jacobian coefficient matrix as,

$$V_0 = \begin{pmatrix} 0 & -\frac{\bar{k}}{N}\alpha_0 I_0 & -\frac{\bar{k}}{N}\beta_0 I_0 \\ 0 & \frac{\bar{k}}{N}\alpha_0 I_0 - \sigma & 0 \\ 0 & 0 & \frac{\bar{k}}{N}\beta_0 I_0 - \gamma \end{pmatrix}$$

The characteristic equation of V_0 is

$\lambda^3 + [\sigma + \gamma - \frac{\bar{k}}{N}(\alpha_0 + \beta_0)I_0]\lambda^2 + (\frac{\bar{k}}{N}\alpha_0 I_0 - \sigma)(\frac{\bar{k}}{N}\beta_0 I_0 - \gamma)\lambda = 0$ thus, according to Hurwitz Theorem, $E_0 = (I_0, 0, 0, N - I_0)$ is a stable solution of Equation (2) equals to:

$$\Delta_1 = \sigma + \gamma - \frac{\bar{k}}{N}(\alpha_0 + \beta_0)I_0 > 0$$

$$\Delta_2 = \begin{vmatrix} \Delta_1 & 1 \\ 0 & (\frac{\bar{k}}{N}\alpha_0 I_0 - \sigma)(\frac{\bar{k}}{N}\beta_0 I_0 - \gamma) \end{vmatrix} > 0$$ that is,

$$I_0 \in \left[\max(\frac{\sigma N}{\alpha_0\bar{k}}, \frac{\gamma N}{\beta_0\bar{k}}), \frac{(\sigma+\gamma)N}{(\alpha_0+\beta_0)\bar{k}}\right]$$

Therefore, when $I_0 \in \left[\max(\frac{\sigma N}{\alpha_0 \bar{k}}, \frac{\gamma N}{\beta_0 \bar{k}}), \frac{(\sigma+\gamma)N}{(\alpha_0+\beta_0)\bar{k}}\right]$, $E_0 = (I_0, 0, 0, N - I_0)$ is a global stable equilibrium solution of Equation (2) .

In summary, $E_0 = (I_0, 0, 0, N - I_0)$ is globally stable means as the time $t \to \infty$, the final state of Equation (2) will be $(I_0, 0, 0, N - I_0)$. That is, all the spreaders (spammers and normal spreaders) in the equation will turn into stiflers in the end, and the ignorants can be any number in $\left[\max(\frac{\sigma N}{\alpha_0 \bar{k}}, \frac{\gamma N}{\beta_0 \bar{k}}), \frac{(\sigma+\gamma)N}{(\alpha_0+\beta_0)\bar{k}}\right]$.

However, when do marketing on the OSNs, the company want to maximize its profit at a fixed time t, how to maximize the promotional profit of the marketing at a given time t is still need to be studied.

III. RESULTS

Because Equation (2) is a set of nonlinear differential equations, which has no displayed solution. In this section, we investigate the influence of the above parameters on the information diffusion at time t by numerical study. Those parameters including initial condition (number of spammer seeds, K_1 , number of normal spreaders seeds, K_2), infectious rates (α_0 and β_0), conversion rate (δ_0), informing rate (λ_0), decay rates (σ and γ) and the average degree of the network ($\bar{k}$). Besides, Equation (2) is not stiff, so the simulation will be carried out using Runge-Kutta method and MATLAB.

Firstly, function F_s and F_c are defined as follows:

$$F_s(t) = S_m(t) + S_n(t) \quad (5)$$

$$F_c(t) = S_m(t) + S_n(t) + R(t) = N - I(t) \quad (6)$$

Where F_s is the size of spreaders (spammers and normal spreaders) in the diffusion process, F_c implies the population of individuals who have known about the marketing information. In some ways, F_s can be seen as the "heat energy" of an information diffusion process. Because more spreaders can cause more interactions, and the marketing product could be discussed more seriously, thus the marketing can gain more profit. Especially in OSNs, more discussions could lead to the explosion of "WoM". At the same time, F_c can somehow be a index to measure the promotional profit of the marketing information diffusion process. In the following part, suppose $G = (V, E)$ is an closed homogeneously OSN network with $N = 10^6$, we will detect how those parameters (as mentioned above) affect F_s and F_c in G .

A. *The Effect of K_1 and K_2*

As mentioned in Section 2, K_1, K_2 are the numbers of initial spammers and normal spreaders, respectively. In order to investigate the influence of K_1 , we set $K_1 = 1, 10, 100, 1000$, successively, and set other parameters as constants, where $K_2 = 10 = \bar{k}$, $\alpha_0 = \beta_0 = \sigma = 0.5$, $\gamma = \sigma - 0.1\varepsilon$ (ε is a random number). For K_2, we do the same setting as K_1 , except $K_1 = 10 = \bar{k}$, $K_2 = 1, 10, 100, 1000$ successively. The Fig.(2) and Fig. (3) below show the simulation results.

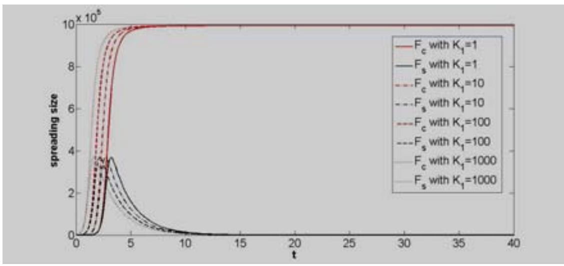

Fig.2 F_S and F_C under different numbers of initial **pammers, K_1 ,**

with $K_2 = 10 = \bar{k}$, $\alpha_0 = \beta_0 = \sigma = 0.5$,

, $\gamma = \sigma - 0.1\varepsilon$

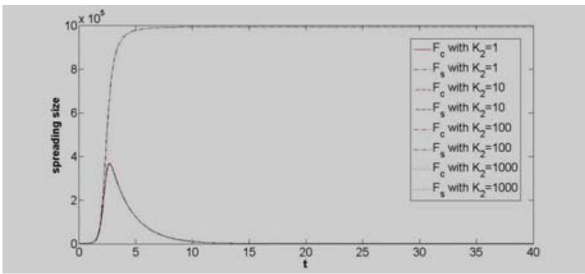

Fig.3 F_S and F_C under different numbers of initial normal spreaders,

K_2 ,with **$K_1 = 10 = \bar{k}$, $\alpha_0 = \beta_0 = \sigma = 0.5$, $\gamma = \sigma - 0.1\varepsilon$**

From Fig.(2) and Fig. (3), we can see that: 1) as time increases, F_C firstly experiences a rapid growth period, then turn into stable, as $t \to \infty$. While, F_S experiences a rapid growth period, followed by a slow decrease period, then turn into stable like F_C . These phenomena follow our analysis discussed in Section 2.2: $E_0 = (0,0,0, N)$ is a stable state of the equation; 2) with K_1 's growth, F_C gets bigger, while F_S first grows at the initial time period then decreases; 3) as K_1 grows bigger, the spreading period of F_C and F_S are both advanced; 4) the diffusion process has no change under different K_2, as which can be seen from Fig. (3), all the curves which describe F_C and F_S under various K_2 coincide into two curves.

Therefore, the number of initial spammers can advance the spreading period, while the number of the initial normal spreaders has no impact on the diffusion process. This analysis implies that through control the number of initial spammers, the marketing managers can somehow advance the spreading period of their marketing information and also increase the final information's spread range, which will cut off their promotional costs and improve their promotional profits.

B. The Effect of α_0 and β_0

As mentioned in Section 2, α_0 and β_0 are the average user influence of spammers and normal spreaders, at the initial time, respectively. In this part, we set $K_1 = 10 = K_2 = \bar{k}$, $\beta_0 = \sigma = 0.5$, $\gamma = \sigma - 0.1\varepsilon$ as constants, and $\alpha_0 = 0.1, 0.3, 0.5, 0.7$, successively, in order to observe α_0 's effect on the information diffusion process.

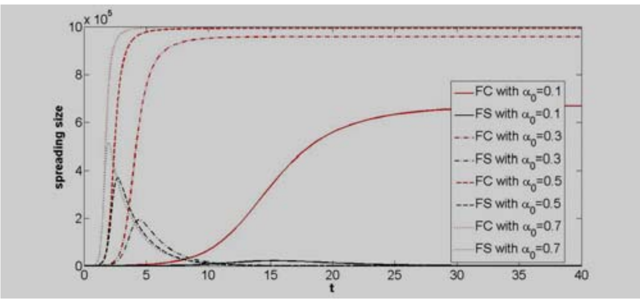

Fig.4 F_S and F_C under different average user influence of spammers, α_0 , with $K_1 = 10 = K_2 = \bar{k}$, $\beta_0 = \sigma = 0.5$, $\gamma = \sigma - 0.1\varepsilon$

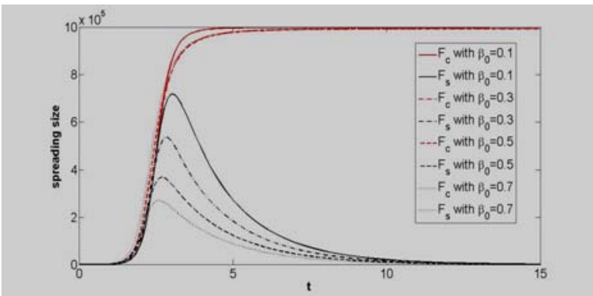

Fig.5 F_S and F_C under different average user influence of normal spreaders, β_0 with $K_1 = 10 = K_2 = \bar{k}$, $\alpha_0 = \sigma = 0.5$, $\gamma = \sigma - 0.1\varepsilon$

From Fig. (4), we can see that: 1) for $\alpha_0 = 0.1, 0.3$, F_C turns into a non-zero constant, as $t \to \infty$, while for all $\alpha_0 = 0.1, 0.3, 0.5, 0.7$, F_S turns into 0 as $t \to \infty$ .This verifies $E_0 = (I_0, 0, 0, N - I_0)$ ($I_0 > 0$) is also a stable state of Equation (2) as we discussed in Section 2.2, and implies that α_0 can affect the system's convergent state; 2) with α_0 's growth, F_C and F_S both get bigger at the same time; 3) the spreading speed (shown as the slopes of the curves) of the diffusion process is increased with α_0 's growth, but the growth rate (the increase difference between every two curves) is decreased with α_0 ; 4) as α_0 grows bigger, both the spreading period of F_C and F_S are shortened.

For β_0 , we did the same setting as α_0 , except $\alpha_0 = 0.5$, $\beta_0 = 0.1, 0.3, 0.5, 0.7$, successively. The Fig. (5) shows the simulation of the diffusion process under different β_0 . And we can find that β_0 does little help on the diffusion process, except that β_0 's growth slightly decrease F_S 's spreading size and speed.

Therefore, the spammers' average user influence (α_0) plays a positive impact on the diffusion speed and can shorten the spreading period, while the normal spreaders' average user influence (β_0) plays a slightly negative impact on the spreading size, which is consistent with our common sense. This implies that through adding the average influence of spammers, managers can increase their promotional benefit and control the spreading period or speed of the information. To be specific, there are many ways to increase one's user influence, such as deliver more interesting and trustworthy [21], be more active, and making more friends, etc, managers can apply those methods to add the average user influence of spammers.

C. The Effect of σ and γ

σ and γ are the decay rates of spammers and normal spreaders, which indicate the probability of spreaders (spammers and normal spreaders) becoming stiflers. To investigate how these two decay rates influence the diffusion process, we successively set $\sigma = 0.1, \; 0.3, \; 0.5, \; 0.7$, with $\gamma = \sigma - 0.1\varepsilon$, $K_1 = 10 = K_2 = \bar{k}$, $\alpha_0 = \beta_0 = 0.5$. The following Fig.(6) illustrates σ 's effect on the diffusion, that σ 's growth slightly decreases the spreading size of F_C and F_S, and shortens the spreading period of the diffusion process. Fig.(7) below shows the effect of γ, which is exactly the same with σ, with the same setting of σ except $\gamma = 0.1, \; 0.3, \; 0.5, \; 0.7$, successively, and $\sigma = \gamma + 0.1\varepsilon$. Besides, when $\sigma = 0.5, \; 0.7$ or $\gamma = 0.5, \; 0.7$, F_C turns into a non-zero constant, as $t \to \infty$. While for $\sigma = 0.1, \; 0.3, \; 0.5, \; 0.7$ or $\gamma = 0.1, 0.3, 0.5, 0.7$, F_S turns into 0, as $t \to \infty$. This implies that σ and γ can also affect the system's convergent state as α_0.

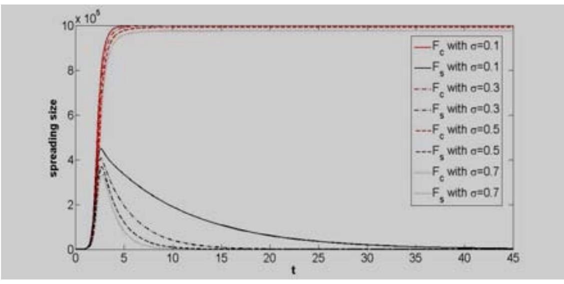

Fig.6 F_S .and F_C under different decay rate of spammers, σ,

with $\gamma = \sigma - 0.1\varepsilon$, $K_1 = 10 = K_2 = \bar{k}$

, $\alpha_0 = \beta_0 = 0.5$

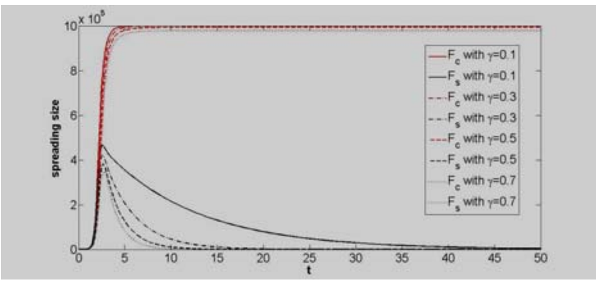

Fig.7 F_S and F_C under different decay rate of normal spreaders, γ,

with $\sigma = \gamma + 0.1\varepsilon$, $K_1 = 10 = K_2 = \bar{k}$, $\alpha_0 = \beta_0 = 0.5$

Therefore, the decay rates (σ and γ) play a negative impact both on the spreading size and period. Through this implication, managers can control the spreading period of the information diffusion process by controlling the decay rates of their information spreaders (spammers and normal spreaders), which can be achieved by making their promotional information more interesting and easy-reading.

D. The Effect of $\bar{k}$

As mentioned in Section 2, $\bar{k}$ is the average degree of the participants in G. The following Fig.(8) shows the effect of $\bar{k}$ on the diffusion, with other parameters as constants ($K_1 = 10 = K_2$, $\alpha_0 = \beta_0 = \sigma = 0.5$, $\gamma = \sigma - 0.1\varepsilon$). From the Fig.(8), we can reach the similar conclusion as α_0: 1) $\bar{k}$'s growth increase the spreading speed of the diffusion with a monotonically decreasing growth rate; 2) with $\bar{k}$'s growth, F_S and F_C grow at the same time; 3) as $\bar{k}$ grows, the spreading period are advanced. In other words, $\bar{k}$ plays a positive impact on the diffusion speed and shortens the spreading period.

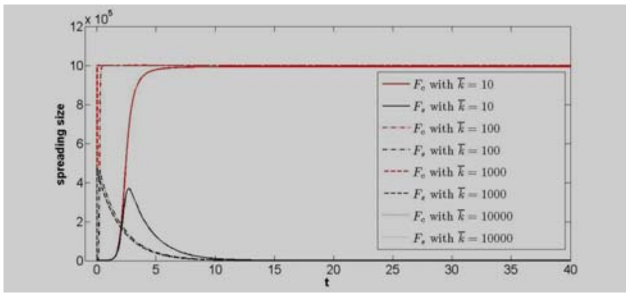

Fig.8 F_S and F_C under different average degree of the network, $\bar{k}$,with $K_1 = 10 = K_2$, $\alpha_0 = \beta_0 = \sigma = 0.5$

, $\gamma = \sigma - 0.1\varepsilon$

In summary, from Fig.(2) to Fig.(8), we can get that:

1) under specific conditions

($I_0 \in \left[\max(\frac{\sigma N}{\alpha_0 \bar{k}}, \frac{\gamma N}{\beta_0 \bar{k}}), \frac{(\sigma + \gamma) N}{(\alpha_0 + \beta_0)\bar{k}} \right]$),

$E_0 = (I_0, 0, 0, N - I_0)$ is a stable equilibrium solution of Equation (2); 2) the diffusion process of marketing information under spammer existence is similar to regular information diffusion, where at the initial time, the information experiences a rapid growth and as the time $t \to \infty$, the diffusion turns into a stable state with all spreaders (both spammers and normal spreaders) turn into stiflers; 3)the growth of K_1 , α_0 and $\bar{k}$ can advance the spreading period of the diffusion process. While σ and γ 's growth can shorten the spreading period; 4) at the same time, β_0 , σ or γ play negative impacts on the spreading size and spreading period of F_c and F_s ; 5) α_0 , σ , γ , $\bar{k}$ play positive impacts on the spreading speed of the diffusion .

The growth of F_c , F_s and the spreading speed, can increase the number of potential customers, thus they can promote the promotional benefit. Besides, the shortening and advancing of the spreading period can cut off the promotion costs. Therefore, managers can increase their promotional profit through adding the number of initial spammer seeds, increasing spammers' average user influence or making the information more interesting and easy-reading. The simulation we did in Section 3 offers a preliminary analysis on how these parameters affect online marketing promotional benefit maximization as F_c and F_s .

IV. DISCUSSION AND CONLUSIONS

In this paper, a IS_mS_nR model is provided to model the marketing information diffusion process under spammers' existence in OSNs. Compared to SIR rumor spreading model, IS_mS_nR model is novel in the following aspects: 1) IS_mS_nR model divides the group of spreaders into spammers and normal spreaders; 2) ignorants can turn into spammers or normal spreaders with two infectious rates which are functions of their average user influence, respectively; 3) spammers only deliver spams until be informed by normal spreaders or lose interest to the message and become stiflers; 4) normal spreaders deliver norm messages and can inform spammers until they lost interest to the message. To the best of our knowledge, this paper attempts for the first time to study the information diffusion under spammer existence. The numerical simulation results show that: 1) the information diffusion of marketing under spammer existence shows a normal spreading trend as other circumstances, which is without spammers; 2) the number of the initial spammer seeds, the average degree, the groups' average user influence and the decay rates not only can affect the spreading size of the information diffusion, but also can affect the spreading speed and period.

With those meaningful findings, this paper can guide some applications in practice. For example, marketing managers can increase the number of initial spammers, add their user influence and make their information more interesting or easy-reading to get more users known their promotion, thus to improve their promotional benefit. Besides, managers can predict the diffusion process of their marketing information with accurate calculation of users' influence and appropriate setting of the decay rates.

CONFLICT OF INTEREST

The authors confirm that this article content has no conflicts of interest.

ACKNOWLEDGMENT

This research was supported by the National Natural Science Fundation of China (71571073), the Natural Science Foundation of Guangdong Province (2014A030313243) and the Twelfth Five-year Planning Project for Guangdong Philosophy and Social Science (GD14CGL09).

REFERENCES

[1] Zhang K Z K, Cheung C M K, Lee M K O. “Examining the moderating effect of inconsistent reviews and its gender differences on consumers’ online shopping decision” [J]. International Journal of Information Management, 2014, 34(2): 89-98.

[2] Stringhini G, Kruegel C, Vigna G. “Detecting spammers on social networks” [C] //Proceedings of the 26th Annual Computer Security Applications Conference. ACM, 2010: 1-9.

[3] Benevenuto F, Rodrigues T, Almeida V, etal. “Detecting spammers and content promoters in online video social networks” [C] //Proceedings of the 32nd international ACM SIGIR conference on Research and development in information retrieval. ACM, 2009: 620-627.

[4] Feijóo C, Gómez-Barroso J L, Voigt P. “Exploring the economic value of personal information from firms’ financial statements” [J]. International Journal of Information Management, 2014, 34(2): 248-256.

[5] Chou E Y, Lin C Y, Huang H C. “Fairness and devotion go far: Integrating online justice and value co-creation in virtual communities” [J]. International Journal of Information Management, 2016, 36(1): 60-72..

[6] Benevenuto F, Rodrigues T, Almeida V, et al. “Detecting spammers and content promoters in online video social networks”

[C]// Proceedings of the 32nd international ACM SIGIR conference on Research and development in information retrieval. ACM, 2009: 620-627..

[7] Martin A, Lakshmi T M, Venkatesan V P. “An information delivery model for banking business”[J]. International Journal of Information Management, 2014, 34(2): 139-150.

[8] Durugbo C, Tiwari A, Alcock J R. “Modelling information flow for organisations: A review of approaches and future challenges” [J]. International Journal of Information Management, 2013, 33(3): 597-610.

[9] Trpevski D, Tang W K S, Kocarev L. “Model for rumor spreading over networks” [J]. Physical Review E, 2010, 81(5): 056102.

[10] Miritello G, Moro E, Lara R. “Dynamical strength of social ties in information spreading” [J]. Physical Review E, 2011, 83(4): 045102.

[11] Lü L, Chen D B, Zhou T. “The small world yields the most effective information spreading” [J]. New Journal of Physics, 2011, 13(12): 123005.

[12] Guille A, Hacid H, Favre C, et al. “Information diffusion in online social networks: A survey” [J]. ACM SIGMOD Record, 2013, 42(2): 17-28.

[13] Cheng J J, Liu Y, Shen B, et al. “An epidemic model of rumor diffusion in online social networks” [J]. The European Physical Journal B, 2013, 86(1): 1-7.

[14] Kermack W O, McKendrick A G. “A contribution to the mathematical theory of epidemics” [C]// Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences. The Royal Society, 1927, 115(772): 700-721.

[15] Kermack W O, McKendrick A G. “Contributions to the mathematical theory of epidemics—II. The problem of endemicity” [J]. Bulletin of mathematical biology, 1991, 53(1): 57-87.

[16] Granovetter M. “Threshold models of collective behavior” [J]. American journal of sociology, 1978: 1420-1443.

[17] Kempe D, Kleinberg J, Tardos É. “Maximizing the spread of influence through a social network” [C]// Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, 2003: 137-146.

[18] Daley D J, Kendall D G. “Stochastic rumours” [J]. IMA Journal of Applied Mathematics, 1965, 1(1): 42-55.

[19] Zanette D H. “Dynamics of rumor propagation on small-world networks” [J]. Physical review E, 2002, 65(4): 041908.

[20] Moreno Y, Nekovee M, Pacheco A F. “Dynamics of rumor spreading in complex networks” [J]. Physical Review E, 2004, 69(6): 066130.

[21] Zhao L, Cui H, Qiu X, et al. “SIR rumor spreading model in the new media age” [J]. Physica A: Statistical Mechanics and its Applications, 2013, 392(4): 995-1003.

[22] Huo L, Huang P, Guo C. “Analyzing the dynamics of a rumor transmission model with incubation” [J]. Discrete Dynamics in Nature and Society, 2012, 2012.