

# Predict Individual Choice With Internal and External Factors In OSN

Wanning Jiang

School of Electrical and  
Information Engineering  
Shanghai Jiao Tong University  
Email: jwn930@sjtu.edu.cn

Songjun Ma

School of Electrical and  
Information Engineering  
Shanghai Jiao Tong University  
Email: masongjun@sjtu.edu.cn

Xinbing Wang

School of Electrical and  
Information Engineering  
Shanghai Jiao Tong University  
Email: xwang8@sjtu.edu.cn

**Abstract**—Social networks, like Twitter and Chinese Weibo, play a fundamental role in the diffusion of information. However, there are different ways of how information diffuses from the several source users to the large scale knowledge. Internal factors includes the interactions amongs online friends, while external influence comes from news, other OSNs, offline friends, etc. In addition, it is more difficult because influence does not necessarily lead to a behavior to spread the piece of information. To predict individual's attitude and action towards a piece of information, we derive a time variant epidemic model based on the survival analysis and machine learning. In this model, information can reach a node via the links of the social network or through the influence of external sources.

**Keywords**—social, network, machine learning, behavior, information.

## I. INTRODUCTION

The Online Social Network (OSN) is the latest evolution of the Web platform where information gets spread at a more rapid speed so that a larger scale of people are covered than ever before. The unique features like mobility and instantaneity make OSN highly open and users tightly interacted. Thanks to the basic functions of main OSNs, information is generated very quickly, consumed by millions of users, and updated quickly by others via commenting and reposting [1]. Therefore, new approaches are needed to take advantage of OSN to predict the intension of people to be attracted by a piece of breaking events, new products, or other informations, for the sake of better grasp the pulse of the times and innovative points to split the information.

We often think of information, a rumor, or a piece of content as being passed over the edges of the underlying social network [3] [4]. Following this mechanism, information spreads over the edges of the network like an epidemic, just in the same way that diseases diffuse through populations by various infections, or computer viruses and worms attack operating systems via networks and portable disks. However, the OSN is influenced also by external factors due to the emergence of mass media, like newspapers, TV stations and online news sites [2]. So the information not only reaches us through the links of our social networks but also through the influence of exogenous out-of-network sources [6]. From the early stages of research on news media and, more generally, information diffusion, there has been the tension between global effects from the mass media and local effects carried by socail structure. Today, mass media as well as the social

networks both exist in the same Web ecosystem, which means that it is possible to collect massive online social media data and at the same time capture the effects of mass media as well as the influence arising from the social networks. This allows us to study processes of information diffusion and emergence in much finer detail than ever before. Therefore, we make three contributions based on the question “How do ‘intrinsic spread’ and ‘external spread’ dynamically make the information accepted or rejected by the OSN users?”

- 1) We study the network graph, and derive a variant exposure probability of each user in this system, which indicates the chance that the user will see the information, respectively from intrinsic spreaders and external sources.
- 2) We proposed a machenism of attitude transformation in the process that user accept the information, and model it within a realistic scenery.
- 3) We figure out the most weighted parameters that make a user popular and believable. We also know the most influential external mass media through computation using the training set, thus we can create an optimally efficient profile of external influence.

**Intrinsic spreading.** It is not uncommon that we know a piece of breaking news because a glasp of a tweet sent by your friend on weibo.com, which is the Chinese version of twitter. This is the basic way that we consider this intrinsic spreading spreading. But each user only have a limmited budget of attention [7], namely we only have a limited frequency to log in our weibo account to see the fresh tweets from friends. Also, we will not check every tweets from every friends because perhaps the omitted tweets do not contain interesting keywords and the omitted friends are not close with us in offline social activities. Those factors makes a uniform intrinsic spreading pridiction become impossible, so we take considerations of specific features of every pair of friends and the empirical distribution of the intrinsic spreading probabilities. As we all know, the cumulative probability of being exposed by the information will increase as the time shifting. Apart from that, the steady state is depended, because for any piece of information, there are always some people immune towards it, and whether one will be immune should be embodied in our system model. Survival analysis, after some change to the logistic finity, are proper to discribe this kind of distribution, as in IV we will further discribe how we choose the model.

**External spreading.** On Twitter, users often post links to various webpages most often these are links to news articles, blog posts, funny videos or pictures. Generally there are two fundamental ways how users learn about these URLs and tweet them. One would be due to the exogenous out-of-the-network effects. For example, one can imagine a scenario, where one checks news on CNN.com, finds an interesting article and then posts a tweet with a URL to the article. In this case CNN is the external influence that caused that URL to emerge onto a particular Twitter user. In order to accurately model the emergence of content in Twitter we need to consider the activity of the invisible out-of-network sources that also transmit information to the nodes of the Twitter network (via channels, like TV, newspapers, etc.). We present a probabilistic generative model of information emergence in networks, in which information can reach a node via the links of the social network or through the influence of the external source. Developing such a model is important. As for the mathematical model, although the trend and variance looks similar with the intrinsic process, there are many other difficulties. The most significant one will be the gap between exposure and infected in epidemic scenario, namely seeing and reposting or commenting in the social network scenario.

**User behavior.** We try to answer the question “How many times have one seen the information before deciding to act towards it like spread or comment on?” in this part. It should be a conditional probability because at this time we only care about the deep reason that the user act towards this piece of information. The only unknown thing is the gap between two visible events: seeing and acting, which we call them exposure and infect in epidemic spreading research. The model is highly close to the human cognitive and learning science, because when we spread something in a community, first we are bound to be aware of it, by means of several times of repeat pay attention to the very information. Then at a certain time point, we decided to take some voluntary action to diffuse it. The concept is analogous to the strength of resistance to a disease. Some people are able to endure more times of exposure to viruses than others without contagion. The relationship between internal spreading, external spreading, and user behavior is showed in Figure 1. What calls for special attention is that the internal infection should not have a stage of being exposed but still not infected, since the infection probability almost relies on the impact of the very information and the very friend, but the repeating times of information. So the gap function is only necessary in external infection process.

Furthermore, we develop an efficient parameter estimation technique. We are given a network and a set of node infection status. We then infer the event profile and intrinsic influence coefficients. We also infer the gap function that models the probability of infection as a function of the number of external exposures of a node. Our model accurately distinguishes external influence from network diffusion.

## II. RELATED WORKS

Work on the diffusion of innovations [9] provides a conceptual framework to study the emergence of information in networks. Conceptually, we think of an (often implicit)

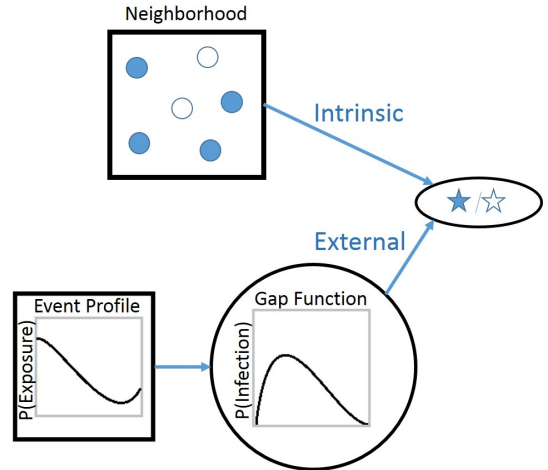


Fig. 1. Our model of infection. A node (denoted by a star) is exposed to information through external sources, and is influenced by status of neighbors.

network where each node is either active (infected, influenced) or inactive, and active nodes can then spread the contagion (information, disease) along the edges of the underlying network. A rich set of models has been developed that all try to describe different mechanisms by which the contagion spreads from the infected to an uninfected node [10] [11] [4] [12]. However, most of those models only focus on the diffusive part of the contagion adoption process, while neglecting the external influence. In this regard our work introduces an important dimension to the diffusion of innovations framework, where we explicitly model the activity and influence of the external source. Although there are some works related with the external influence [15] [14] [13], they focus on the borrowing the physical theory onto the OSN research but ignore that the uniform strategy is unproper to a high degree.

## III. DATASET

Our dataset is from Chinese Sina Weibo. Since the experiment is not launched, the actual size of the data is unsure. But we will use enough large dataset to validate the algorithm and modify the model according to the results.

## IV. SYSTEM MODEL

Here, we develop in detail our novel information diffusion model that incorporates both the spread of information from node to node along edges in the network as well as the external influences acting on the network. Additionally, our model reconciles the gap between a stream of external exposures arriving in continuous time and a series of discrete decisions leading to infection. The rest of the system model part is developed in this sequence. First we explain 1 in a more systematic way, after which we discuss the model for each part in detail, followed by the joint expressions that combine the three rectangular in 1 together, and can be applied into our dataset to train the parameters.

The whole system is based on a god eye view, which indicates that when we want to decide one user’s status, we have imply his neighbors’ status by means of observations in a very short time which could be omitted. In another words,

only the central node's status which is waiting for prediction is unassured. The scenario is unrealistic in the real OSN because no mechanism can assure an existence of god eye view, however, the hypothesis is useful and enough to deal with the predictive problem. For the stochastic network graph, we will focus on it as next step, and the highlight will be the big picture instead of the precise individual status. In 1, at each time slot, a node  $i$  consults the  $n$  of his neighbors, and get contagion at a different rate from each. Meanwhile the node receives a stream of varying intensity of external exposures, governed by the event profile of outside world, including mass media and offline friends as long-range travelers in dynamic graph theory [8].

#### A. Survival Analysis

Consider a single contagion. In our model, we call the intrinsic infection and the external exposure a uniform mechanism because of their similarities, that is the information contagion. It may occur at a variant rate when a neighbor of a node becomes infected, or a piece of news was released outside the online social friendship networks. Then the information, like the disease, is transmitted after a random interval of time. The process highly resembles the patient survive or die in an epidemic spreading process. So we borrow the survival analysis and add to some amendment to be accord with our problem.

The traditional survival analysis is the basement of our research of internal and external infection mechanism, so this subsection is meant to introduce the utilized part of survival analysis, the physical meaning of this mathematical model in our empirical problem and the intention of our change to the traditional algorithm.

Let  $T$  be the random variable denoting death time of a patient, and at time  $t$  which is continuous, the probability of infection is denoted as

$$F(t) = Pr(T \leq t).$$

Correspondingly, the probability of healthy is denoted as

$$S(t) = Pr(t < T).$$

$Pr$  here is the abbreviation of probability. That is, the survival function is the probability that the time of death is later than some specified time  $t$ . Usually one assumes  $S(0) = 1$ , although it could be less than 1 if there is the possibility of immediate death or failure, which is not possible in our model. An important property of survival function is non-increasing:  $S(u) \leq S(t)$  if  $u \geq t$ , so the model we prepare for survival function must satisfy the property. Given the survival probability, we can imply the density function and the lifetime distribution function easily.

As we pointed in I, one user may never be enough interested in a piece of information so there should be a possibility that at the end of time he is still out of infection. Driven by this intention, we modify the survival analysis. The survival function is usually assumed to approach zero as age increases without bound, i.e.,  $S(t) \rightarrow 0$  and  $F(t) \rightarrow 1$  as  $t \rightarrow \infty$ . In our model, we let  $F(\infty) = h(\vec{X})$ , where  $\vec{X}$  depicts the characters of the user. Thus related quantities are defined in terms of the survival function, and will be showed in the following subsections.

#### B. Intrinsic Influence

We denote the whole online social network as a graph  $G = (V, E)$ , and  $V$  means the users,  $E$  means the directional friend relationship. Assuming we are doing a research towards the same information, then time and individual are the essential variants in modeling the intrinsic influence. A person will show different sensitivity towards different friends, which means the user may have 80% probability to be impacted by friend A, while only 20% probability by friend B. The percentage should certainly be between zero and one, so in our model we use a logistic function to derive the impact probability from  $j \in N(i)$  to  $i$ :

$$h(\vec{x}_{ij}) = g(\theta^T x_{ij}) = \frac{1}{1 + e^{-\theta^T x_{ij}}}$$

where  $N(i)$  means the neighbors of node  $i$ . The logistic function satisfies the constraints as  $h(\vec{x}_{ij}) \rightarrow 0$  when  $\theta^T x_{ij} \rightarrow -\infty$  and on the contrary  $h(\vec{x}_{ij}) \rightarrow 1$  when  $\theta^T x_{ij} \rightarrow \infty$ . To represent the user's property we assume:

$$\vec{x}_{ij} \in \mathbb{R}^4 = \begin{cases} Act(i) = \frac{w(i)}{\varepsilon_w} \\ Act(j) = \frac{w(j)}{\varepsilon_w} \\ Wil(i) = \frac{m(j)}{m_\Sigma} \\ Wil(j) = \frac{m(i)}{m_\Sigma} \end{cases}$$

and here  $Act(i)$  is the active level if user  $i$ .  $x$  should be composed of both the sender's activity and the receiver's activity, defined by the frequency of sending weibo in the past 14 days.  $w(i)$  is the weibo numbers and  $\varepsilon_w$  is the a reference level to normalize the parameters, noted that if  $w(i) > \varepsilon_w$ , then let  $Act(i) = 1$ .  $Wil(i)$  and  $Wil(j)$  is the emotional tendency that  $i$  and  $j$  are infected by each other.  $m_\Sigma$  is also the normalizing item.

After the steady status of the survival analysis, we should infer the transient status of the dynamic system. Traditionally, exponential distribution is used to depict the non-increasing survival probability. In our model, we let lifetime distribution be in the multi-infinity exponential form:

$$F(t) = -h(\vec{x}_{ij})(e^{-\rho t} - 1)$$

, in this form  $F(\infty) = h(\vec{x}_{ij})$ ,  $F(0) = 1$  and  $F(t)$  is non-decreasing.

Since we focus on the differential distribution, we define the derivative as:

$$\lambda_{int}^i(t) = \frac{\rho \cdot \exp(\theta^T x_i - \rho t)}{1 + \exp(\theta^T x_i - \rho t)}.$$

The physical sense is the hazard of node  $i$  getting infected by node  $j$  at time  $t$ . To simplify the training part, we let  $\lambda_{int}^{(i)}(t)$  be 'average' infection hazard, where  $\vec{x}_i$  is composed by the highest 4 parameters among  $\{Act(i), Act(j), Wil(i), Wil(j)\}$  for all neighbors  $j$  respectively. To correspond with the next subsections, we let  $\Phi_{int}^i(t)$  denote  $F(t)$ , the cumulative infection distribution. The relationship between  $h(\vec{x}_i)$  and  $\Phi_{int}^i(t)$  is showed in Figure 2.

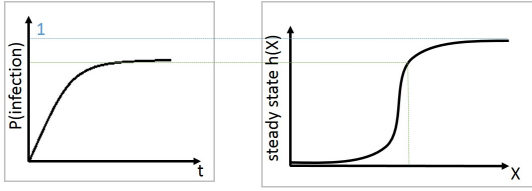


Fig. 2. The relationship between the asymptotic function and the dynamic function

### C. External Exposure

Similar with the intrinsic infection part, in analysing external exposure, we should consider not only the event profile which is governed by the external influential coefficients, but also the user's sensitivity towards the event profile. However, the internal continuous function is no longer proper because of there existing a obvious gap between virtual world and the online social network, which is like a transition of two different social platforms. In this subsection we only talk about the event profile, and in IV-D we will focus on the individual resistance in changing from outside world to online social network.

If we choose a period, like when a video is popular, than in this period we can assume that in every slot, the influence of event profile is uniform:

$$\Lambda_{ext}^i = \frac{1}{1 + \exp(-\alpha y)}$$

, where  $\Lambda_{ext}^i$  means the probability of node (i) get exposed by external information in a time slot and  $\vec{y}$  is the event profile defined as follow:

$$\vec{y}_i \in \mathbb{R}^{3+d} = \begin{cases} news = \frac{com(item)}{com_{\Sigma}} \in \mathbb{R}^d \\ 2ndplatform = \frac{traffic}{\varepsilon_t traffic} \\ advertisement = (0, 1) \\ receivergender(i) = (0, 1) \end{cases}$$

Here *news* is a d-dimension vector composed of the percentage of news in every column, and *plat* is a normalized parameter of the traffic in another platform. It is realistic because a second OSN may have convolutional impact on the main OSN. For example, when Facebook is in the heat of a video, sooner or later a lot of Twitters will share the url of the video. The item of  $q(i)$  seems so easy, but in applications it can be modified to more complex forms. We only aim to provide the algorithm that can solve a individual dependent event profile and can be easily generalized.

### D. Individual Habit

Define a gap function which means the probability that user  $i$  is infected by external influence immediately after the  $r$ st external exposure. Due to the discrete definition of external exposure in a time slot, the numbers of exposes after  $z$  time slots can be calculated as it obeys Bernoulli Distribution. To be correspond with the intrinsic infection rate, we rewrite the

discrete time probability:

$$\begin{aligned} \Phi_{ext}^i(t) &= \sum_{z=1}^{\infty} \left\{ P_{exp}^i(z; t) \left[ 1 - \prod_{r=1}^z (1 - Q(r)) \right] \right\} \\ &\approx \sum_{z=1}^{\infty} \left[ \binom{k}{z} (\Lambda_{ext}^i)^z (1 - \Lambda_{ext}^i)^{k-z} \right] E \\ &= \sum_{z=1}^{\infty} \left[ \binom{k}{z} \frac{\exp[-(k-z)(\alpha^T y)]}{[1 + \exp(-\alpha^T y)]^k} \right] E \end{aligned}$$

where

$$E = 1 - \prod_{r=1}^z (1 - Q(r))$$

So the infection probability combined intrinsic and external influence should be

$$\Phi^i(t) = 1 - [1 - \Phi_{int}^i(t)][1 - \Phi_{ext}^i(t)].$$

To divided the continuous time into discrete time slots, we should choose the proper length of the time slot, and thus  $k = \left\lfloor \frac{t}{\Delta T} \right\rfloor$  in the above equations.

## V. TRAIN THE MODEL

Next we develop a method of inferring the model parameters for a given network and tract of a single contagion. We are given the network and the infection times for each node that got infected with the contagion under consideration. We then need to infer the coefficients  $\vec{\theta}$  of event profile  $\vec{x}$ , the coefficients  $\vec{\alpha}$  of event profile  $\vec{y}$ , and the individual habits  $Q(r)$ . Maximize the training data likelihood, and we can infer the parameters. The training data joint likelihood is:

$$L(\theta, \alpha) = \prod_{i \in \mathbb{D}} \left\{ \Phi^i(t)^{s^i} [1 - \Phi^i(t)]^{1-s^i} \right\},$$

and the log-likelihood is correspondingly:

$$\sum_{i \in \mathbb{D}} \{ s^i \log \Phi^i(t) + (1 - s^i) \log [1 - \Phi^i(t)] \}.$$

Here  $s^i$  is the status of node  $i$ . If  $i$  is infected, let  $s^i = 1$ , else  $s^i = 0$ .

However, there is a latent parameter  $z$ , the times of external exposures before infection, with a distribution  $Q(z)$ , in the likelihood expression, so it is difficult to directly infer the parameters from the single expression. Our method is EM algorithm. First assume

$$\sum_{z=1}^4 = 1,$$

that is

$$Q(z) \begin{cases} = 0 & z \neq 1, 2, 3, 4; \\ \neq 0 & z = 1, 2, 3, 4. \end{cases}$$

It means for every user, he or she might only be infected during the first to fourth time he or she sees the information. If the user reposts the information, on and after the fifth time he or she is bound to have repost it. The two step of EM algorithm

is shown as below, and the algorithm is listed in algorithm 1.

**E-step:**

$$Q_i(z^i = j) := \frac{P(s^i | z^i = j) \cdot Q_i(z^i = j)}{\sum_{l=1}^4 P(s^i | z^i = l) \cdot Q_i(z^i = l)}$$

$$= \frac{P(s^i, z^i = j)}{\sum_{l=1}^4 P(s^i, z^i = l)};$$

where

$$P(s^i, z^i = j) = \begin{cases} Q_i(z^i = j) \Phi_{int}^i(t) + A & s^i = 1 \\ Q_i(z^i = j) [1 - \Phi_{int}^i(t)] - A & s^i = 0 \end{cases},$$

and

$$A = [1 - \Phi_{int}^i(t)] P_{exp}^i(z^i; t) \left\{ 1 - \prod_{r=1}^{z^i} [1 - Q(r)] \right\}$$

and

$$P_{exp}^i(z^i; t) = \binom{k}{z^i} (\Lambda_{exp}^i)^{z^i} (1 - \Lambda_{exp}^i)^{k - z^i}$$

In E-step the iteration is done easily, however in M-step it is more difficult because we borrow the gradient descent algorithm to derive the coefficients in logistic functions.

**M-step:**

$$\theta, \alpha := \arg \max_{\theta, \alpha} \sum_i \sum_{z^i} Q_i(z_i) \log B$$

where

$$B = \begin{cases} \Phi_{int}^i(t) + \frac{1}{Q_i(z^i)} A & s^i = 1 \\ 1 - \Phi_{int}^i(t) - \frac{1}{Q_i(z^i)} A & s^i = 0 \end{cases}$$

Using the gradient descend algorithm (take  $\vec{\alpha}$  for example to illustrate the iteration idea):

$$\alpha_j = \alpha_j + \Delta \frac{\partial}{\partial \alpha} \sum_i \sum_{z^i} Q_i(z_i) \log B.$$

and the result is:

$$\alpha_j = \alpha_j + \Delta C \frac{\partial}{\partial \alpha_j} P_{ext}^i(z_i; t)$$

where

$$C = \frac{1}{Q_i(z_i)} [1 - \Phi_{int}^i(t)] \left[ 1 - \prod_{r=1}^{z^i} (1 - Q_i(r)) \right]$$

and

$$\frac{\partial}{\partial \alpha_j} P_{ext}^i(z_i; t) = \binom{k}{z^i} [z_i (\Lambda_{ext}^i)^{z^i} (1 - \Lambda_{ext}^i) - (k - z^i - 1) (1 - \Lambda_{ext}^i)^{k - z^i} \Lambda_{ext}^i] y_j$$

Here the iteration step length  $\Delta$  is adjustable.

---

**Algorithm 1** EM algorithm to solve the problem

---

```

1: for each do
2:   Initialize  $\vec{\alpha}, \vec{\theta}, Q(z), l_{max}$ 
3: end for
4: while not converged do
5:   for each  $i \in V$  do
6:     // Do E-step and update  $Q$ 
7:   end for
8:   // Do M-step
9:   while  $l > l_{max}$  do
10:    for each  $\alpha_j$  in  $\vec{\alpha}$  do
11:      // Update  $\alpha_j$ 
12:    end for
13:    for each  $\theta_j$  in  $\vec{\theta}$  do
14:      // Update  $\theta_j$ 
15:    end for
16:     $l_{max} \leftarrow l$ 
17:  end while
18: end while

```

---

## VI. PROPOSED EXPERIMENTS

We plan to do the future jobs and the experiments in this summer. The outline of experiments are listed here.

- 1) Decide the length of the time slot, namely  $K$  in the above discussions.
- 2) Decide the proper iteration step length in E-step, and make sure that the algorithm will converge in a fast and accurate way.
- 3) Compare the probability of intrinsic influence and external influence, and learn the mechanism of information diffusion in social networks. Depend on the results to dicide the further steps.

## VII. CONCLUSION

Emergence of information has traditionally been solely modeled as a diffusion process in networks. however, it is identified that only around 71% of URL mentions on Twitter can be attributed to network effects, and the remaining 29% of mentions seem to be due to the influence of external out-of-network sources [2]. Therefore we present a model in which information can reach a node via the links of the social network or through the influence of external sources. We are going to apply this model to the emergence of URLs in the Weibo network in the approaching summer and see if there are any difference in the infection pattern from previous conclusions. Applying the algorithm to the real data and we can also get the coefficients of the external profile. Make use of the conclusion and the spread of the information can be modified to be more effective and to the point. We should emphasize that our model does not only reliably capture the external influence, but, as a consequence, also leads to a more accurate description of the real network diffusion process.

For future work it would be necessary to do the proposed experiment for the first step. Then we will change our point of view to be macroscope. We will release the assumption of the god eye view by observe the online social network as a stochastic process graph. Applying the mean field theory to

the graph we will get more dramatic conclusions. Another inspiration is to focus on the second social network's influence. In addition, it will be interesting to move further to research into the convolutional mutual influence of internal and external influence.

#### REFERENCES

- [1] A. Guille and H. Hacid. *A Predictive Model for the Temporal Dynamics of Information Diffusion in Online Social Network*. In *WWW 2012*.
- [2] S. Myers, C. Zhu and J. Leskovec. *Information Diffusion and External Influence in Networks*. In *KDD 2012*.
- [3] J. Leskovec, M. McGlohon, C. Faloutsos, N. Glance, and M. Hurst. *Cascading Behavior in Large Blog Graphs*. In *SDM 2007*.
- [4] D.J. Watts. *A Simple Model of Global Cascades on Random Networks*. In *PNAS*.
- [5] H.W. Hethcote. *The Mathematics Of Communication: An Up-To-Date Report on an Hypotheses*. In *SIAM Review*,42(4):599-653,2000.
- [6] A.B. Longman *Classics in Political Science*, seventh edition, 2006.
- [7] B. Jiang, N. Hegde, L. Massoulie and D. Towsley. *How to Optimally Allocate Your Budget of Attention in Social Networks*. In *INFOCOM 2013*.
- [8] A. Gopalan, S. Banerjee, A.K. Das, S. Shakkottai. *Random Mobility and the Spread of Infection*. In *INFOCOM 2011*.
- [9] E. Rogers. *Diffusion of Innovations*. Fress Press, 4th Ed, 1995.
- [10] D. Centola and M. Macy. *Complex contagions and the weakness of long ties*. *American Journal of Sociology*, 2007.
- [11] D. Cosley, D. P. Huttenlocher, J. M. Kleinberg, X. Lan, and S. Suri. *Sequential influence models in social networks*. In *ICWSM*, 2010.
- [12] D. M. Romero, B. Meeder, and J. M. Kleinberg. *Uncovering the Temporal Dynamics of Diffusion Networks*. In *WWW*, 2011.
- [13] J. Yang and J. Leskovec. *Modeling Information Diffusion in Implicit Networks*. In *ICDM*, 2010.
- [14] A. Goyal, F. Bonchi and L. V. S. Lakshmanan. *Learning Influence Probabilities In Social Networks*. In *WSDM*, 2010.
- [15] S. Banerjee, A. Gopalan, A. K. Das and S. Shakkottai. *Epidemic Spreading with External Agents*. In *INFOCOM*, 2011.
- [16] D. Wang, H. Park, G. Xie, S. Moon, M. Kaafar and K. Salamation. *A Genealogy of Information Spreading on Microblogs: a Galton-Watson-based Explicative Model*. In *INFOCOM*, 2013.
- [17] X. Cheng, H. Li, J. Liu. *Video Sharing Propagation in Social Networks: Measurement, Modeling, and Analysis*. In *INFOCOM*, 2013.