Modeling the Views of WeChat Articles by Branching Processes

Lin Zhang* and Yuhong Liu†
* School of Science, Beijing University of Posts and Telecommunications, Beijing, China
† Department of Computer Science and Engineering, Santa Clara University
E-mail: yhliu@scu.edu Tel/Fax: +1-408-551-3513

Abstract—**In cyber security, the temporal patterns of information propagation in online social networks is of crucial importance, especially for false information detection and defense. The mechanism of information propagation is the fundamental of online security, such as predicting and suppressing false information dissemination. The dynamics of information popularity in online social networks open the possibilities for understanding the mechanisms of information spreading.**

In this paper, we study the temporal patterns associated with online contents generated by WeChat subscription accounts. In order to reveal the popularity dynamics of WeChat articles, we formulate a probabilistic model using branching process to reveal the mechanism of views of WeChat articles. Specifically, a non-Markovian age-dependent branching process is introduced for the purpose of considering the heterogeneousness of human behavior and the tree-like dissemination structures. Different from the Markovian case with exponentially distributed lifetime of particles in the branching system, the heavy-tailed power law distributed inter event time is one of the key ingredients for our model. Moreover, the limitation of users' attention span, and the attractiveness of articles are also considered in our model.

We demonstrate our approach on the real data of WeChat articles' popularity time series. The branching model is successful in presenting the temporal patterns of the real evolution. Our findings offer insights into the temporal patterns of information popularity in online social networks, which provides references for further prediction and control of information propagation concerning cyber security.

I. Introduction

The prosperity of online social media has dramatically changed the way people work, interact and entertain. Online social media and user-generated contents is becoming increasingly dynamic. There are many social media platforms supporting the publishing and sharing of information, behavior, and opinions, for example, Facebook, Twitter, Flicker, YouTube, Sina Weibo, WeChat, etc. [4], [12], [15], [29].

Increasing attention from both academic and industry has paid on how to use such platforms for greater societal benefits, such as in emergence management, politics and economics [7], [27]. On the other hand, a recent study found that false information spreads significantly farther, faster, deeper, and more broadly than the truth in all categories of information [24]. The propagation of false information online has become a severe threat to the security of online social network and attracted extensive attention. Investigating the mechanism of temporal pattern of information propagation is an crucial topic in cyber security.

Modeling the information diffusion process is of outstanding interest since it provides the fundamentals to understand and manage the dynamics of content on social media [8]. Scholars have studied online information diffusion from various perspectives, which can be roughly classified into three categories.

Studies in the first category have been mainly developed based on the epidemiological processes. They classify nodes into several states and focus on the evolution of the proportions of nodes in each state. SIR and SIS are the two seminal models [11], where S stands for "susceptible", I for "infected" (i.e. adopted a piece of information), and R for "recovered". Nodes in S state switch to the I state with a fixed probability β, and backward with probability γ in SIS case. In the SIR case, nodes in state I switch into R state permanently. The proportion of each state is expressed by simple differential equations under the assumption of homogeneous randomness.

The second category is graph theory based studies. For example, the Linear Threshold Model [19], [22], [25] assumes that a node is influenced by each of its neighbors, and if the number of its infected neighbors exceeds a certain threshold, this node will also be infected. Another example is the Independent Cascade Model [5], [20], [26], which assumes that an infected node at each discrete step randomly chooses one of its neighbors for propagation with a certain probability of success, which is independent of the propagation history. The target neighbor will be infected if the propagation succeeds. The process runs until no more infection is possible.

In addition, machine learning based approaches are also proposed recently to learn information propagation from large-scale data [17], [18], [21], [30], [28]. However, these approaches are often data-driven and cannot theoretically answer the how and why questions for information propagation.

These existing studies, however, often underestimate the behavior patterns of different users, such as the heterogeneousness of response time, the limitation of users' attention span, and the attractiveness of different articles. As a result, some major questions remain open. (1) How will human behavior patterns influence the propagation patterns? (2) Can we mathematically model the real time propagation patterns?

In this paper, we propose to model the real time information spreading in social media as a branching process. In particular,

This paper is partially supported by the NSFC (No. 61671005) and CSC (No. 201806475024).

considering the heterogeneous human behavior pattern, the limitation of attention span, as well as the attractiveness of articles, a non-Markovian age-dependent branching model is applied to reveal the temporal pattern of the popularity of WeChat article at collective level. In the non-Markovian case, the heavy-tailed power law distributed response time is applied instead of the exponential distribution in the Markovian one. Finally, we match our model to the real WeChat articles' time series of popularity data and achieve high accuracy, which reveals the validation of modeling information propagation by branching process.

II. Data

WeChat is a Chinese multi-purpose messaging, social media and mobile payment app developed by Tencent [9]. It was first released in 2011, with more than 1.1 billion monthly active accounts in the first quarter of 2019. WeChat subscription account is a basic type of account, allowing content and notification updates for subscribed followers displayed in the subscription area of a user's WeChat messages. The contents updated by accounts can be viewed by the subscribers and other users once shared in "Moments", which is WeChat's brand name for its social function of friends updates. Moments allows users to post images, post text, post comments, share music, share articles, etc. Only the friends from the user's contact are able to view their Moments contents and comments.

The dataset used in this paper is the views of articles released by subscription accounts. The views increase with time after its release. Therefore, the data appears as a time series of views for each article. We have collected 620 articles from 140 WeChat subscription accounts [28]. The views data of each article are collected every 4 minutes shortly after their release time. The recorded data are saved as time series of views, which are illustrated as in Table I.

The scatter plot of the time series of views are illustrated in Fig. 1. The x-axis is time and the y-axis is the amount of views. As it is shown in Fig. 1, the views increase as time goes by. The total amount of views for different articles are quite different. In the paper, we are aiming at revealing the mechanism of the temporal pattern of the views of WeChat articles by a branching model. In our model, the real data time series and the model simulation match quite well. Moreover, the impact of the parameters are investigated, considering the

TABLE I
Example of time series of views

Time	Views	Time	Views	⋯
10:18	454	10:15	1383	⋮
10:22	663	10:19	1832	⋮
10:26	820	10:23	2227	⋮
⋮	⋮	⋮	⋮	⋮
23:36	11699	23:07	15071	⋮

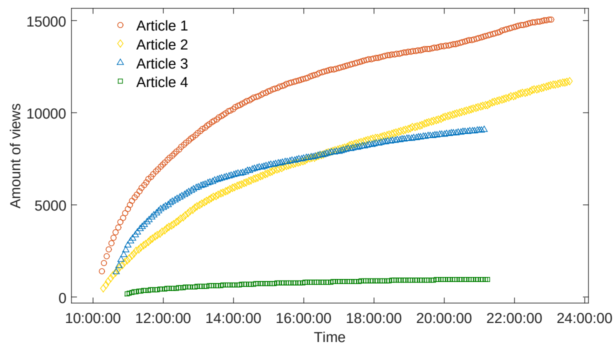

Fig. 1. Scatter plots of views time series of four typical articles.

heterogeneousness of human behaviors, the limitation of users' attention, and the attractiveness of articles.

III. Model

Many studies reveal that the trace of information propagation possessing a cascading tree-like structure, which inspires us to model the temporal pattern by branching processes. Reference [15] shows the retweet structure is tree liked. In online blog communities, there is also a diffusion tree [23]. Rumors and meres on Facebook also possesses cascades [6], [2]. The evolution tree model also appears in fake news [13]. Based on the existing work, we explore a branching model to reveal the spreading process of WeChat articles.

The dynamics of popularity of an article are modeled based on a branching structure, considering three influential factors, (1) the heterogeneousness of human behavior, (2) the limitation of uses' attention span, and (3) the attractiveness of articles. In the rest of this section, the theory of branching processes is introduced firstly. Then, the corresponding relationship between branching structure and WeChat popularity are linked by our key parameters from the above three influential factors.

The theory of branching processes is an area of mathematics that describes situations in which an entity exists for a time and then may be replaced by one, two or more entities during its lifetime. It is a well developed and active area of research with theoretical interests and practical applications [1], [10], [14]. In branching processes, a particle lives for a random time, and at some point during its lifetime, produces a random number of progenies. An illustration of the process is shown in Fig. 2.

The corresponding relationship between branching structure and the popularity is as follows. In the WeChat subscription system, after its release, the article can be read by the subscribers of the account. Then the subscribers will decide whether or not to share it by posting it on the Moments, where their friends can decide to read and/or share it further based on its attractiveness. From this point of view, the number of views of an article also possesses a tree structure.

There are two key random features for a branching process. The first one is the offspring distribution, which gives the

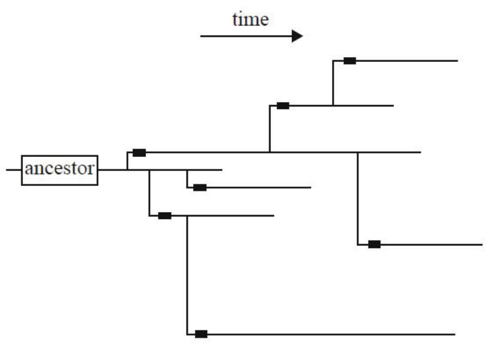

Fig. 2. An illustration of an age-dependent branching process. Black rectangles depict individuals, horizontal lines depict lifetimes. Vertical lines are added to link individuals to their parents. The length of vertical lines is arbitrary [14].

number of progenies that one particle can produce during its lifetime. In the WeChat system, it equates to the number of views of the article after one user's sharing in his/her Moments. The other key random feature is the distribution of the lifetime. The lifetime of particles in a branching process is used to model the waiting time for friends' views after one shares the article in the Moments. The temporal pattern of WeChat article popularity relies on the two key random features above. Moreover, the three influential factors are considered simultaneously.

For the offspring distribution, the limitation of users' attention span and the attractiveness of articles are modeled. The offspring distribution is modeled by binomial distribution Bin $(n(t),\ p)$ with parameters $n(t)$ and p, where $n(t)$ is the average number of friends in one's Moments at time t that can see the title and general information of the shared article, p is the probability of viewing the article. Specifically, the offspring distribution ξ_t at time t is given as follows.

$$P(\xi_t = k) = \binom{n(t)}{k} p^k (1-p)^{n(t)-k}, \quad k = 0, 1, \cdots, n(t).$$

Furthermore, considering the limitation of attention, we assume that $n(t) = A * t^{-\alpha}$, where A is the average total number of friends, and α is a positive real number, serving as the parameter of attention decay. The attractiveness of the articles is modeled as the view probability p, which is fixed, while the number of friends that can see the shared article decays with time due to the merging of new contents and limitation of attention of each user.

For the distribution of lifetime, the temporal pattern originates from human behavior is considered. The distribution of waiting time can not be modeled by exponential distribution, although it is the guarantee of homogeneity, which can make the branching model much easier. However, many real data evidence reveals that human behavior, no matter online or offline, is quite heterogeneous [3], [16]. As a consequence, the distribution of inter event time, which is the time between two views of one's Moment, is a heavy-tailed distributions. One of the commonly used distributions for human dynamics is the power law distribution. The density function of the inter event time between two successive events possesses the polynomial decay tail:

$$f(\tau) \sim \tau^{-\beta},$$

where $\beta > 0$ is the parameter of power law distribution. However, the inter event time distribution is not the waiting time of a shared article to be read by friends. It is the residual time T, which can be derived from the inter event time distribution as

$$f_w^{\beta}(T) = \frac{1}{\langle \tau \rangle} \int_T^{\infty} f(\tau) d\tau \sim T^{-\beta+1}. \tag{1}$$

Now we are ready to build our age-dependent branching process model. At the beginning, there is only one individual. It can be used to model the release of an article. Then, at a random time, the article can be viewed and shared. After sharing, the article can be viewed and shared further at other random times, as the structure shown in Fig. 2. Specifically, the random time is power-law distributed with density distribution f_w^{β}. The number of further views of a share happened at time t follows a binomial distribution $\mathrm{Bin}(A * t^{-\alpha}, p)$. The amount of views can be modeled as the total population of branching process evolving with time t. Since $\lim_{t\to\infty} A * t^{-\alpha} = 0$, there will be no more views after quite a long time, and the population goes to a constant eventually.

The three influence factors are modeled by β, α, and p for the heterogeneousness of human behavior, the limitation of attention span, and the attractiveness of articles, respectively. The descriptions for the notations are listed in Table II and the algorithm of the age dependent model used in this paper is in Algorithm 1.

IV. EXPERIMENT RESULTS

The experiment of our paper is done by MATLAB. In order to match the model to the real data, the parameters are selected elaborately for different articles. During the experiment, the average number of friends A is fixed with value 50. The other three parameters p, α, and β are changed case by case. The matching results of the 4 selected articles shown in Fig. 1 are shown in Fig. 3. Since the data are collected during one day shortly after the release of articles, and the views are collected every 4 minutes, the x-axis is set from 0 to 150 or 200, with 1 unit as 4 minutes for simplicity, which is 600 or 800 minutes in total, covering the whole time interval of the view time series we crawled. The same x-axis are also appeared in the following figures, Fig. 4-6. The color of the scatter dots are

TABLE II
TABLE OF NOTATIONS

Notation	Description
A	average number of friends
t	time
α	decay rate of the number of friends that can see the article
p	probability of view an article
β	power of the human behavior
$\mathrm{Bin}(A * t^{-\alpha}, p)$	distribution of views shared at time t
f_w^{β}	density distribution of waiting time for a view

Algorithm 1 Age-dependent Branching Process

Input parameters A, N, p, α, β
$Population(1) \leftarrow 1$
$View \leftarrow$ One random number from $\text{Bin}(A, p)$ for views of the released article at the beginning
$WaitTime \leftarrow View$ random numbers from f_w^β for waiting times at which the article is viewed
for $t = 2$ to N **do**
  $WaitTime \leftarrow WaitTime - 1$
  if $WaitTime = 0$ **then**
    $L \leftarrow$ Logical array of $WaitTime = 0$
    $TempView \leftarrow$ sum of L
    $NewView \leftarrow$ sum of $TempView$ random numbers from $\text{Bin}(A * t^{-\alpha}, p)$ for future views of the article shared at time t
  end if
  $NewWaitTime \leftarrow NewView$ random numbers from f_w^β for waiting times at which the article is further viewed
  $WaitTime \leftarrow$ Delete $WaitTime = 0$
  $WaitTime \leftarrow$ Combine of $WaitTime$ and $NewWaitTime$
  $Population(t) \leftarrow Population(t-1) + TempView$
end for
Output $Population$

the same as in Fig. 1, referring different articles. The black curve in each sub-figure is the simulation result of our age-dependent branching process model. Each curve is the average over 1,000 trails of the sample paths. The parameters are listed in Table III, and also explained in the caption of Fig. 3. The well match between the scatter plot of the real data and the simulation curve validates that the age-dependent branching process model reveals the intrinsic mechanism of temporal pattern hidden in the propagation of online content popularity.

In the following, the impact of the three factors are illustrated by numerical experiments. The impact of each parameter is analyzed by fixing other parameters and turning only one parameter once. For comparison concern, the three parameters are analyzed on the background of the first three articles. Parameters have already listed in Table III.

First, the effect of the attractiveness modeled by view probability p is analyzed based on Article 1. More attractive article gets more views. Therefore, it is obvious to get the influence of p, that is, larger probability makes the views more.

TABLE III
PARAMETERS FOR EACH ARTICLE

	p	α	β
Article 1	0.10 (0.09, 0.11) [a]	0.715	3.65
Article 2	0.09	0.709 (0.65, 0.75) [a]	3.48
Article 3	0.12	0.900	3.70 (3.50, 3.90) [a]
Article 4	0.05	2.200	3.58

[a] Numbers in the parentheses are used for further experiments to analyze the impact of each parameter

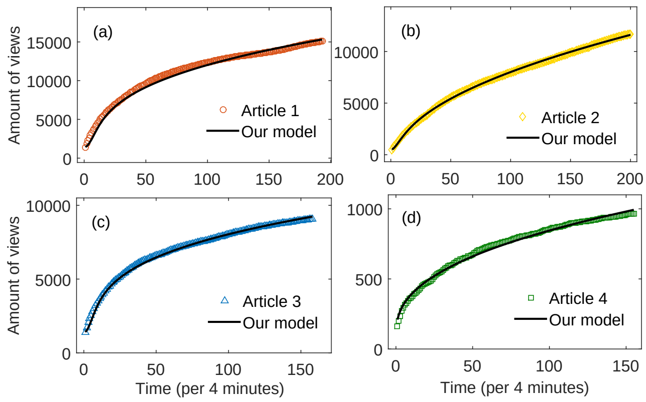

Fig. 3. Model fitting of 4 articles. The x-axis is set from 0 to 150 or 200 with 4 minutes as 1 unit for simplicity, covering the whole time interval of the real data we crawled. The parameter $A = 50$ is fixed. Other parameters are changed case by case. (a) Article 1. The parameters are as follows. $p = 0.10$, $\alpha = 0.715$, and $\beta = 3.65$. (b) Article 2. The parameters are as follows. $p = 0.09$, $\alpha = 0.709$, and $\beta = 3.48$. (c) Article 3. The parameters are as follows. $p = 0.12$, $\alpha = 0.900$, and $\beta = 3.70$. (d) Article 4. The parameters are as follows. $\alpha = 2.200$, $p = 0.05$, and $\beta = 3.58$.

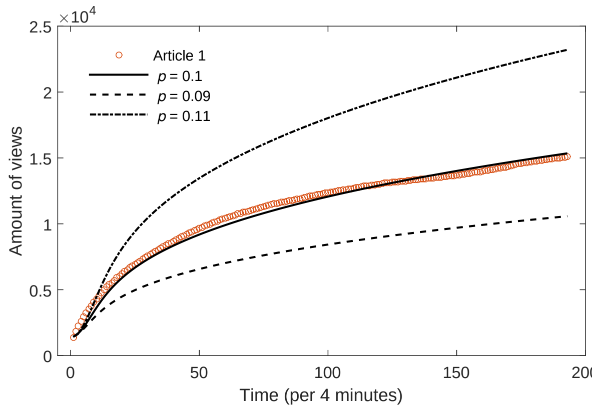

Fig. 4. Effect of parameter p based on Article 1. The fitted parameters are $A = 50$, $p = 0.10$, $\alpha = 0.715$, and $\beta = 3.65$. The turning parameter p is set as $p = 0.09$ and $p = 0.11$.

The experiment result is shown in Fig. 4, which is natural and intuitive.

Second, the impact of the limitation of users' attention span, which is modeled by α, is analyzed based on Article 2. α is the decay rate of the number of average friends $n(t)$ that can see the title and general information of the shared article. Specifically,

$$n(t) = A * t^{-\alpha}.$$

Therefore, larger α indicates faster decrease of $n(t)$, which results in fewer views of the article. The experiment result shown in Fig. 5 supports our analysis.

Finally, the heterogeneousness of human behavior is analyzed, which is modeled by β. Different from the first two factors, the impact of β is not monotone to the amount of views, as the experiment results shown in Fig. 6. There are intersections among the three values of β, indicating the effect

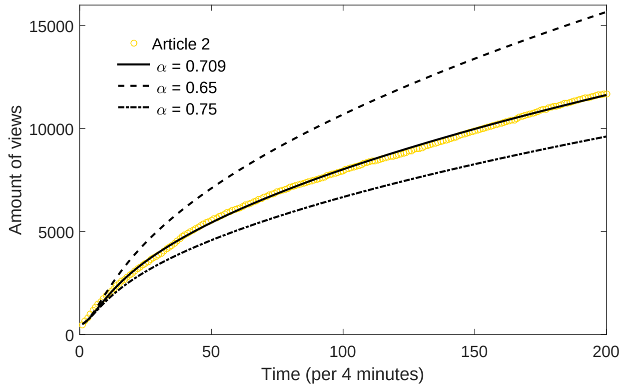

Fig. 5. Effect of parameter α based on Article 2. The fitted parameters are $A = 50$, $\alpha = 0.709$, $\beta = 3.48$, and $p = 0.09$. The turning parameter α is set as $\alpha = 0.65$ and $\alpha = 0.75$.

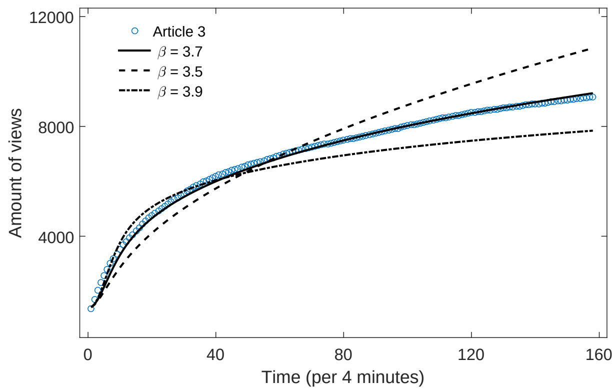

Fig. 6. Effect of parameter β based on Article 3. The fitted parameters are $A = 50$, $\alpha = 0.9$, $\beta = 3.7$, and $p = 0.12$. The turning parameter β is set as $\beta = 3.5$ and $\beta = 3.9$.

of β is not the same at different stage of the propagation.

In order to understand the impact of β, it is important to figure out the impact of β on the heterogeneity. As in (1), smaller β means $P(T > t)$ is larger. Moreover, larger probability for larger waiting time T makes larger heterogeneity. That is, larger possibility to get larger waiting times. To simplify, smaller β tends to get longer waiting times, and larger β for shorter waiting times.

Furthermore, the decay of $n(t)$ is an important ingredient. An illustration of the evolution of share and view is shown in Fig. 7. For Share 1 at time t_1, it will be viewed at time t_2 for shorter waiting time (i.e., larger β) and be viewed at time t_3 for longer waiting time (i.e., smaller β). That is, at the early stage of propagation, more heterogeneity (i.e., smaller β) makes the view process evolves slower. However, at the late stage, the situation is not the same. For the same View 3 at time t_6 in Fig. 7, the share for this view may be Share 2 at time t_4 for longer waiting time (i.e., smaller β), and be Share 3 at time t_5 for shorter waiting time (i.e., larger β). As the decrease of $n(t)$, the views of Share 2 at time t_4 is more

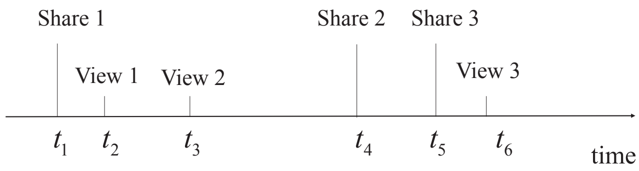

Fig. 7. An illustration of the share and view process.

than Share 3 at time t_5. The heterogeneity slows the decay of $n(t)$ and keeps more views at later stage of the dynamics.

In summary, the age-dependent branching model, considering the heterogeneousness of human behavior, the limitation of attention span, and the attractiveness of article, succeeds in explaining the real time patterns of the views of WeChat articles released by subscription accounts. Moreover, the influence of each parameter is analyzed by numerical experiments.

V. Conclusion and Further Work

In this paper, we present a tree-like-structured branching model to reveal the real time pattern of popularity of contents in online social platform. The proposed scheme relies on integration of several components: the heterogeneity of human behavior, the limitation of users' attention span, and the attractiveness of articles. Our model matches the real data quite well, which in turn validates the rationality of our scheme.

The proposed scheme represents a new philosophy of information propagation modeling, and provides insights into the propagation dynamics for further security concern.

References

[1] K. B. Athreya, P. E. Ney, *Branching processes*, Dover Publications, Mieola, 2004.

[2] L. A. Adamic, T. M. Lento, E. Adar, and P. C. Ng, "Information evolution in social networks," *Proceedings of the Ninth ACM International Conference on Web Search and Data Mining*, pp. 473-482, Feburay, 2016.

[3] A. L. Barabási, "The origin of bursts and heavy tails in human dynamics," *Nature*, vol. 435, pp.207-211, May, 2005.

[4] R. Crane and D. Sornette, "Robust dynamic classes revealed by measuring the response function of a social system," *Proc. Nat. Acad. Sci. USA*, vol. 105, no. 41, pp.15649-15653, Oct. 2008.

[5] L. Dickens, I. Molloy, J. Lobo, C. P. Cheng, and A. Russo, "Learning stochastic models of information flow," *IEEE 28th international conference on data engineering*, pp. 570-581, April, 2012.

[6] A. Friggeri, L. Adamic, D. Eckles, and J. Cheng, "Rumor cascades," *Eighth International AAAI Conference on Weblogs and Social Media*, May, 2014.

[7] N. Grinbery, K. Joseph, L. Friendland, B. Swire-Thompson, and D. Lazer, "Fake news in twitter during the 2016 US Presidential election,"*Science*, vol. 363, pp. 374-378, Jan. 2019.

[8] A. Guille, H. Hacid, C. Favre, and D. A. Zighed, "Information Diffusion in Online Social Networks: A Survey,"*ACM Sigmond Record*, vol. 42, no. 2, pp. 17-28, 2013.

[9] https://en.wikipedia.org/wiki/WeChat.

[10] T. E. Harris, *The theory of branching processes*, Springer-Verlag, Berlin, 1963.

[11] H. W. Hethcote, "The mathematics of infectious diseases," *SIAM review*, vol. 42, no. 4, pp. 599-653, 2000.

[12] A. Java, X. Song, T. Finin, and B. Tseng. "Why we twitter: understanding microblogging usage and communities." *WebKDD workshop*, pp. 56-65, 2007.

[13] S. M. Jang, et al, "A computational approach for examining the roots and spreading patterns of fake news: Evolution tree analysis,"*Computers in Human Behavior*, vol. 84, pp. 103-113, 2018.

[14] M. Kimmel, D. E. Axelrod, *Branching processes in biology*, Springer, New York, 2015.
[15] Q. Kong, W. Mao, G. Chen, and D. Zeng, "Exploring trends and patterns of popularrity stage evolution in social media,"*IEEE Transactions on System, Man, and Cybernetics: Systems*, vol. 99, pp. 1-11, 2018.
[16] M. Karsai, H. H. Jo, and K. Kaski, *Bursty human dynamics*, Springer International Publishing, 2018.
[17] J. Leskovec, M. McGlohon, C. Faloutsos, N. Glance, and M. Hurst, "Patterns of cascading behavior in large bolg graphs," *Proceedings of the 2007 SIAM international conference on data mining*, pp. 551-556, 2007.
[18] J. Leskovec, L. Backstrom, and J. Kleinberg, "Meme-tracking and the dynamics of the news cycle," *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 497-506, 2009.
[19] Y. Lim, A. Ozdaglar, A. Teytelboym, "A simple model of cascades in networks," *Tech. Rep., mimeo*, vol. 1, 2015.
[20] F. Lu, W. Zhang, L. Shao, X. Jiang, P. Xu, and H. Jin, "Scalable influence maximization under independent cascade model," *Journal of Network and Computer Applications*, vol. 86, pp. 15-23, 2017.
[21] L. Liu, B. Qu, B. Chen, A. Hanjalic, and H. Wang, "Modelling of information diffusion on social networks with applications to WeChat," *Physica A: Statistical Mechanics and its Applications*, vol. 496, pp. 318-329, 2018.
[22] A. Nematzadeh, E. Ferrara, A. Flammini, and Y. Y. Ahn, "Optimal network modularity for information diffusion,"*Physical Review Letters*, vol. 113, no. 8, 088701, August, 2014.
[23] S. Pei, L. Muchnik, S. Tang, Z. Zhang, and H. A. Makse, "Exploring the complex pattern of information spreading in online blog communities," *Plos one*, vol. 10, no. 5, e0126894, 2015.
[24] S. Vosoughi, D. Roy, S. Aral, "The spread of true and false news online," *Science*, vol. 359, pp. 1145-1151, March, 2018.
[25] D. E. Whitney, "Dynamic theory of cascades on finite clustered random networks with a threshold rule," *Physical Review E*, vol. 82, no. 6, 066110, 2010.
[26] C. Wang, W. Chen, and Y. Wang, "Scalable influence maximization for independent cascade model in large-scale social networks," *Data Mining and Knowledge Discovery*, vol. 25, no. 3, pp. 545-576, 2010.
[27] L. Wu et.al "Procuct adoption rate prediction in a competitive market," *IEEE Trans. Knowl. Data Eng.*, vol. 30, no. 2, pp. 325-338, Feb. 2018.
[28] J. Xu, Z. Du, J. Guo, X. Fu, Y. Zhang, and Y. Wu, "Empirical and modeling studies of WeChat information dissemination," *Physica A: Statistical Mechanics and its Applications*, vol. 512, pp. 1113-1120, 2018.
[29] Z. Yang, J. Guo, K. Cai, J. Tang, J. Li, L. Zhang, et.al"Understanding retweeting behaviors in social networks,"*Proceedings of the 19th ACM international conference on Information and Knowledge management*, pp. 1633-1636, 2010.
[30] J. Yang, and J. Leskovec, "Patterns of temporal variation in online media," *Proceedings of the fourth ACM international conference on Web search and data mining*, pp. 177-186, Feburary, 2011.