Information Dissemination in scale-free networks: Profusion versus Scarcity

Laurent Brisson, Philippe Collard, Martine Collard, Erick Stattner

Abstract The study of information dissemination in social networks is of particular importance in many areas as marketing, politics and security for example. Various strategies are being developed to disseminate information, those aimed at disseminating information widely and those aimed at disseminating information in a more confidential manner to make it scarce. In this paper, we adapt a model dedicated to spreading rumours by word of mouth in a physical space to the context of social networks. We compare two modes of dissemination based on profusion or scarcity and study the impact of the choice of the initial node. The results obtained show to what extent each mode exploits the social network topology and especially the influence of hubs.

1 Introduction

Dissemination of information within social networks covers many phenomena that are part of everyone's daily life (spread of fake news, buzz effects, propagation of emotions, adherence to projects, etc.). Dissemination of this information is crucial

Laurent Brisson
IMT Atlantique, Lab-STICC, UBL, F-29238 Brest, France,
e-mail: laurent.brisson@imt-atlantique.fr

Philippe Collard
Université Côte d'Azur, CNRS, I3S, UMR 7271, UNS, Parc Valrose, 06108 Nice, France,
e-mail: philippe.collard@unice.fr

Martine Collard
LAMIA Laboratory, University of the French West Indies, Guadeloupe, France,
e-mail: martine.collard@univ-ag.fr

Erick Stattner
LAMIA Laboratory, University of the French West Indies, Guadeloupe, France,
e-mail: erick.stattner@univ-ag.fr

both from a civilian point of view (to make a brand known, to find funding, to convince people of a political ideology, etc.) and from a security point of view (to prevent fake news from spreading, to understand how harmful ideas spread, etc.). Two strategies are observable: trying to saturate networks and spread information massively, or trying to target a small population so that the information is considered an advantage if it remains confidential on the principle that "everything rare is expensive". In this paper we study these two modes of information dissemination. The first, profusion, when the broadcaster relays the information if most of its neighbours are already informed, and the second, scarcity, when the broadcaster relays the information if most of its neighbours are not already informed of it.

Dissemination of information have long been studied in various disciplines as physics, economics, psychology, medical sciences for examples. Numerous studies have looked at the phenomenon of dissemination from the point of view of the spread of epidemics, which has led to the development of models combining epidemiological mathematics and stochastic solutions [11]. Other studies are interested in dissemination from the point of view of rumours spreading. The first approaches, both in epidemiology with the famous SIR compartment model originally defined by Kermack and McKendrick [11] and in rumor modeling with the DK model proposed by [6], were based on differential equations. Most extensions of the original DK models define spot refinements on individual behavior of Spreaders and Stiflers and on pair wise contacts. Zhao et al [14] introduced the concepts of *forgetting and remembering* that is close to *recursiveness*. An individual that was a Spreader but forgot the story, may remember it and becomes again a Spreader. Cheng et al [3] rather considered the quality of the link as the *trustiness* between two individuals as a main factor for spreading or not the story. It defined also a probability for a spreader to have less interest in the rumor independently of the number of Spreaders or Stiflers he meets. Xia and Huang [13] focused on the evolution of *belief* of an agent about rumor and *anti-rumor*. The approach of Borge et al [2] is rather similar since they consider that Spreaders may become *inactive* at times. The ODS model proposed by Collard et al [5] is based on the spread of a rumor from one individual to another individual through word of mouth. It is a compartment model that takes into account the spatial location of mobile agents to model the propagation of the rumor. This work was a first attempt to consider the behavior of an Ignorant in reaction to his/her neighborhood.

More recently, research has focused on the dissemination of information in social networks. Doer et al. [7] have shown that the structure of social networks allows a diffusion of rumors much faster than in other network topologies, including complete graphs. According to them the source of the speedy spread of information is fruitful interaction between the few nodes with many neighbors and the large number of nodes with few neighbors. He et al. [10] investigated the countermeasures to be put in place to restrict the spread of a rumor on a social network. They used a propagation model based on the SIR model and mathematically defined the equilibrium solution to stop a rumor by means of a counter-rumor. Garcia et al. [8] design agent-based models to reproduce and to analyse emotions in online communities. Indeed, although the face-to-face and online interactions are different [12], emotions can emerge collectively, as can be seen through the spread of the Internet memes [9] or conflicts around political debates [4].

In this paper we revisit the ODS model proposed by [5], which models the propagation of rumors by word of mouth in a spatial context, in order to adapt it to handle the dissemination of information in social networks. We are mainly interested in the mode of dissemination and in the choice of the initial node to disseminate the information. For a first step we have limited ourselves to the study of a scale-free network generated by a Barabási-Albert based algorithm [1]. We compare two modes of diffusion respectively based on profusion and on scarcity using 5 indicators: Transmissibility potential, Incidence curve, Number of Stiflers at convergence, Diffusion variability and Influence of hubs (high-degree nodes) during the dissemination. We observe the impact of the choice of the initial infected node with respect to its topological characteristics (betweenness, closeness, pagerank and degree) on the dynamics and the result of the diffusion (i.e. the equilibrium of populations to convergence).

The remainder of this paper is organized as follows. Section 2 presents the ODS model we adapted for diffusion in social network. Section 3 presents computer simulations to study the impacts of the mode of dissemination. Finally, in Section 4 we discuss results and offer our conclusions and hints for future research.

2 A model of information dissemination

In this section we revisit the ODS rumor propagation model proposed in [5] and we adapt it to information dissemination through social networks. Let's consider a graph G of N nodes and M links where a node represents an agent and a link represents the possibility for one agent to contact another. As our aim is to model the information dissemination in a social network, we do not consider mobile agents and a contact between individuals is defined as a node connected to another one by a social link.

2.1 Model description

We adapt the ODS model by redefining the three different compartments (i.e. states of its agents) as follows:

1. Open-minded (O) agents are those individuals who do not have access to the information and are therefore likely to be informed;
2. Disseminator (D) agents are active individuals currently disseminating the information;
3. Stifler (S) agents are the individuals who have had access to the information but no longer disseminate it.

The *ODS* model is defined in two versions ODS_p and ODS_s. In ODS_p, transmission takes place under the condition of profusion, that is, when an Open-Minded individual is surrounded by many neighboring Disseminators. On the contrary, in ODS_s, transmission takes place under scarcity condition, that is, when an Open-Minded individual is surrounded by few neighboring Disseminators. Transitions from compartment O to compartment D and from compartment D to compartment S during one step time characterize the dynamics of the model. The probability that a D-individual k transmits the information to an O-individual upon contact at time t depends on each individual and varies over time; thus it is referred as $\beta_k^{OD}(t)$.

Algorithm 1 Simulation of the generic *ODS* model

$t \leftarrow 0$
Initialize the parameter *DPeriod* $\{\gamma \leftarrow \frac{1}{DPeriod}\}$
Initialize the population size to N
Create N agents
$\{$each agent have a $state$ variable in $\{O, D, S\}\}$
Set all the agents O except one which is D
while $\exists$ one D agent **do**
    // OtoD transitions
    for each $a_i \in D$ **do**
        for each $a_k \in O$ in the neighborhood of a_i **do**
            Compute r_k^D the proportion of D in the neighborhood of a_k $\{r_k^D > 0\}$
            a_k will become D with probability $\beta_k^{OD} = F(r_k^D)$ $\{F$ is a monotonic function from $[0;1]$ to $[0;1]\}$
        end for
    end for
    // DtoS transitions
    for each $a_k \in D$ **do**
        a_k becomes S according a *Poisson* law with mean $\gamma = \frac{1}{Dperiod}$
    end for
    $t \leftarrow t + 1$
end while
Ensure: $\nexists$ D agent

The pseudocode for simulating the *ODS* models is defined in algorithm 1; at the end of the run, there are no O-individuals which could become an D-individual.

2.2 "O to D" transition

A first condition under which the transition "OtoD" may occur is that the two protagonists are connected. It is the potential receiver, and not the transmitter, who decides whether or not he will become itself a transmitter. Now, the question is how an O-individual comes to make the decision to become a disseminator? In the *ODS* model, the likelihood that an O-individual a_k becomes himself a D-individual is function of the rate $r_k^D(t)$ of D-individuals in his neighborhood. On this basis we define two alternative solutions to *ODS*: the first one, called ODS_p, is driven by the *profusion* of information, while the second, called ODS_s, depends on *scarcity*.

2.2.1 Transition based on profusion

In the first instance ODS_p of the model, it is assumed that the higher the rate r_k^D, the higher the probability that the O-individual a_k becomes himself disseminator will be. Let's define the function F (algo. 1 line 4) as $F_p(x) = \frac{1}{1+e^{-c(2x-1)}}$ where c is a constant[1]. Then the value $p_k^D = F_p(r_k^D)$ can be interpreted as the *profusion* of disseminators around the O-individual a_k. Profusion follows an increasing sigmoid curve: the more the profusion, the more the number of D-individuals in the vicinity will be: if $p_k^D = 0$, there are no informed individuals in the vicinity, while if $p_k^D = 1$, all the neighbors are informed. So, the probability that a_k becomes a disseminator is $\beta_k^{OD} = p_k^D$ (algo. 1 line 11).

2.2.2 Transition based on scarcity

The alternative instance ODS_s, is based on the assumption that the higher the rate r_k^D, the lower the probability that the O-individual a_k becomes himself disseminator will be. Let's define the F function as $F_s(x) = 1 - \frac{1}{1+e^{-c(2x-1)}}$, where c is a constant[1]. Then the value $s_k^D = F_s(r_k^D)$ can be interpreted as the *scarcity* of disseminators around the given open-minded individual a_k. Scarcity follows a decreasing s-curve: the higher the scarcity, the lower the number of disseminators will be: if $s_k^D = 0$, all the neighbors are disseminators, while if $s_k^D \approx 1$ there are very few disseminators in the neighborhood. So, the probability that a_k becomes itself disseminator is $\beta_k^{OD} = s_k^D$ (algo. 1 line 11).

2.3 "D to S" transition

The "D to S" transition is common for both instances ODS_p and ODS_s. It is explained in algorithm 1 (lin 15-17) and it is shared with the SIR model. It is based on the fact that the mean period of time that a D-individual remains in his state is fixed to $Dperiod$. Let's note that $\gamma = \frac{1}{Dperiod}$ is the removal or recovery rate.

[1] In the experiments the constant c will be fixed to 5.

3 Experimental results

The experiments were carried out on a Barabasi-Albert graph with 1000 nodes. For a given set of parameters, each node was initially infected 10 times. By varying the DPeriod from 10 to 100 in steps of 10 and comparing the profusion mode with the scarcity mode, each node was infected initially 200 times. Thus, all the results presented in this section were obtained by calculating the median and standard error for the 200,000 experiments performed. In order to compare the two dissemination modes we defined 5 indicators for analyzing the results: **Transmissibility potential**, **Incidence curve**, **Number of Stiflers at convergence**, **Diffusion variability** and **Influence of hubs** (high-degree nodes) during the dissemination. Moreover, throughout this section we will comment on the impact of the initial infected node setup with respect to its topological characteristics (betweenness, closeness, pagerank and degree) on the dynamics and the result of the diffusion (see Table 1).

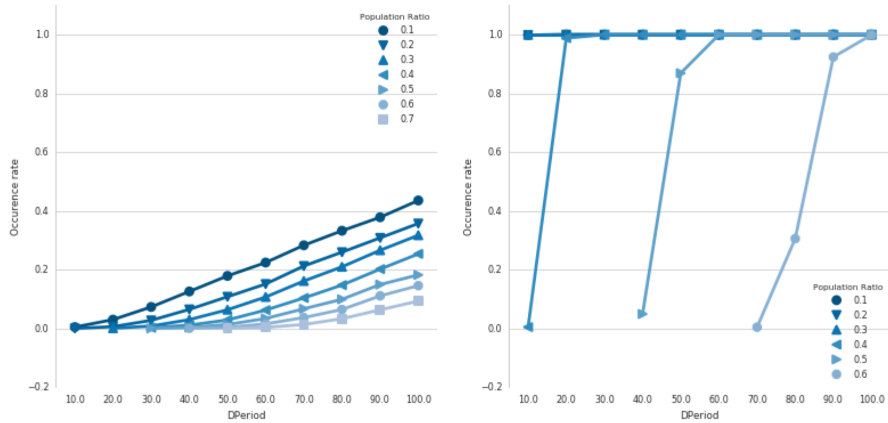
Table 1: Median values of **Pearson correlation coefficients** of one topological characteristics of the initial node (betweenness, closeness, degree, pagerank) and a measure describing the dynamics and the result of the diffusion by setting DPeriod and the dissemination mode. Empty cells mean that no test is statistically significant ($p_{value} > 0.005$).

		Max Disseminators	Time at max Disseminators	Number of Stiflers at convergence	Time at convergence
Profusion	Betweenness	0.17	0.02	0.11	0.06
	Closeness	-0.04	-0.16	0.11	-0.08
	Degree	0.15	-0.01	0.10	0.06
	Pagerank	0.15	-0.01	0.10	0.06
Scarcity	Betweenness		-0.14		-0.08
	Closeness	0.04	-0.64		-0.43
	Degree		-0.18	-0.03	-0.11
	Pagerank		-0.16	-0.03	-0.10

3.1 Transmissibility potential

This indicator shows the extent to which information has been disseminated or not disseminated. This success is characterized by *PopulationRatio* which is the percentage of Stiflers in the population at the end of the diffusion. We define the *OccurrenceRate* as the ratio between the number of times the diffusion has reached at least the *PopulationRation* threshold and the number of experiments performed. We conducted our experiments in order to understand the influence of the mode of dissemination (profusion or scarcity) on the information dissemination in a scale-free network. Figure 1 shows the extent to which information dissemination (*OccurenceRate*) has reached a given percentage of the population (*PopulationRatio*) according to the mean duration (*DPeriod*) an individual remains in the disseminator state.

- **Experimental results with Profusion:** Figure 1a shows that the information has a probability between 0 and 0.4 to pervade up to 70% of the population. This probability increases slowly with *DPeriod*.
- **Experimental results with Scarcity:** Figure 1b shows that the information could only pervade up to 60% of the population. But, for each *PopulationRatio* value, there is a *DPeriod* threshold for which the information systematically pervade.



(a) Profusion (b) Scarcity

Fig. 1: Occurrence rate where information dissemination reached a given percentage of the population *PopulationRatio* by *DPeriod* (each point is calculated on $10,000$ experiments).

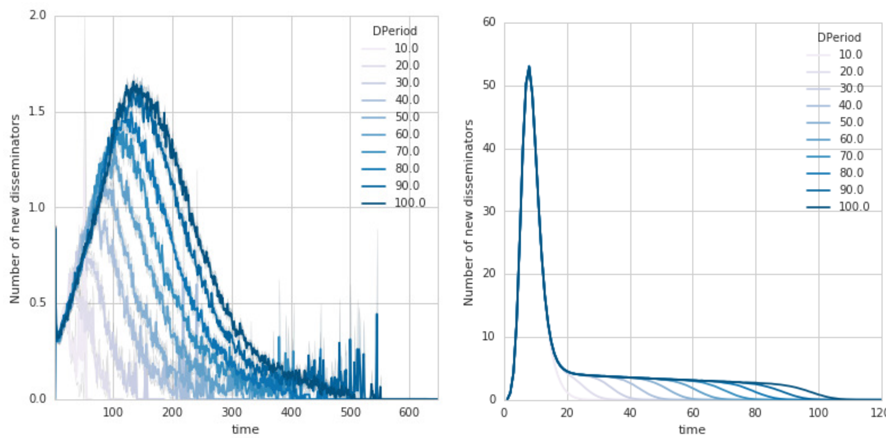
3.2 Incidence curve

The incidence curve shows the evolution of the number of newly informed individuals over time. Figure 2 shows the mean number of new disseminators and the standard error (y-axis) per unit time (x-axis) calculated on $10,000$ experiments.

- **Experimental results with Profusion:** Figure 2a shows a small number of new disseminators over time with a non-zero standard error observable on the curve. The *DPeriod* moderately influence the intensity of the diffusion peak and the time it occurs.

- **Experimental results with Scarcity:** Figure 2b shows a peak with a large number of new disseminators whose average intensity is the same regardless the value of the *Dperiod*. Unlike experiments with Profusion, the low standard error shows that the results are very stable from one run to another.

By analyzing the impact of the choice of the initial node, we can see on Table 1 (1st column) that, in the case of diffusion by profusion, the intensity of the diffusion peak (max Disseminators) is weakly correlated with the degree (0.17), the betweenness (0.15) and the pagerank (0.15) whereas it has no influence on the timing of the diffusion peak. Conversely, in the case of diffusion by scarcity, closeness strongly influences (-0.64) the timing of the diffusion peak (2nd column) but not its intensity (1st column).



(a) Profusion (b) Scarcity

Fig. 2: Incidence curves according *DPeriod* (for each point, median results obtained on 10,000 experiments).

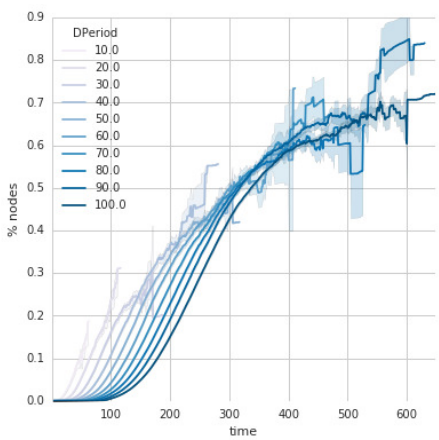
3.3 *Number of Stiflers at convergence*

Another way to observe the outcome of information dissemination is to estimate the final proportion of people who have been informed, i.e. the number of Stiflers at convergence. Figure 3 shows the proportion of Stiflers in the population (y-axis) per unit time (x-axis) according *DPeriod*. This figure shows, for each point, median results with their standard error obtained on $10,000$ experiments.

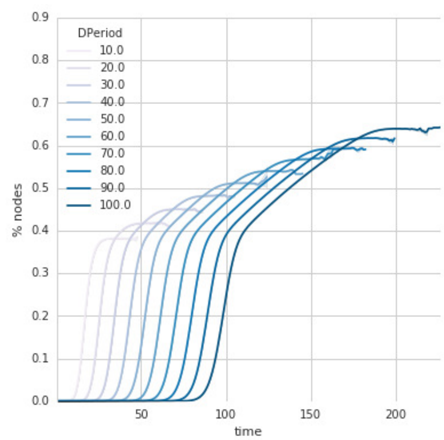
- **Experimental results with Profusion:** Figure 3a shows that if the diffusion lasts long enough (time $\geq$ 400) a large part of the population (from 65% to 80%) can be informed. However, it is also observed that over the time interval [200, 600] the standard error is high, which shows a large variability in the impact of the diffusion.

- **Experimental results with Scarcity:** Figure 3b shows that the duration of a diffusion is shorter than a diffusion by Profusion. The *DPeriod* influences the final number of Stiflers which ranges from 40% to 65% of the population. Again, we can see a very low standard error which highlights the stability of the result.

By analyzing the impact of the choice of the initially infected node, we can see from Table 1 (3th and 4th columns) that in the case of Profusion, once again, this choice does not influence the equilibrium of populations at convergence or the time of convergence. However, in the case of Scarcity, closeness strongly influence the time at convergence (-0.43).



(a) Profusion



(b) Scarcity

Fig. 3: Population of Stiflers according *DPeriod* (for each point, median results obtained on $10,000$ experiments).

3.4 Diffusion variability

This indicator allows us to observe whether the diffusion always occurs in the same way with the same setting. During the study of the incidence curve and the number of Stiflers at the convergence, we have observed a high standard error with the profusion mode. In order to analyze this characteristic, we compare in Figure 4 the mean standard deviation of the number of Stiflers at the convergence (x-axis) by reproducing 10 times for each node a diffusion with the same initial setting. So,

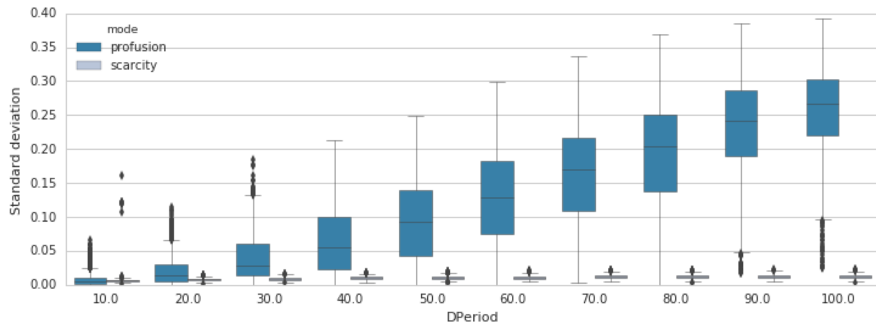


Fig. 4: **Profusion** vs **Scarcity** - Mean Standard deviation of Number of Stiflers at the convergence computed over 10 experiments for each initial infected node with the same diffusion parameters.

each column is the mean result over $10,000$ experiments given the *DPeriod* (y-axis) and the diffusion mode (Profusion or Scarcity). It is explicit that the scarcity mode produces robust and reproducible results from one diffusion to another while the profusion mode produces results whose variability increases with the $DPeriod$.

3.5 Influence of hubs during the dissemination

Figure 5 shows the temporal evolution of the mean degree of the disseminators according to the two dissemination modes ($Dperiod$ is set to 10). The observation of the two curves enables us to say that: (i) with profusion, the agents with a rather low degree (approximately 1.5 on average) disseminate information throughout the duration of the dissemination; (ii) with scarcity, early in the process, information spreads rapidly mainly thanks to hubs (i.e. the agents with a high degree); then, progressively, agents with decreasing degree disseminate to reach the leaves of the network with 1-$degree$.

4 Conclusion

In this paper we have compared two modes of information dissemination in a relational network with a degree distribution following a power law (scale-free network). The ODS model has been adapted to model information dissemination in the context of social networks where individuals are the nodes of the network and their social relationships the links between these nodes. We have assumed that the "acceptance criteria" for one information by somebody depends on the ratio of his direct neighbors who already possess the information. Such hypothesis enables us

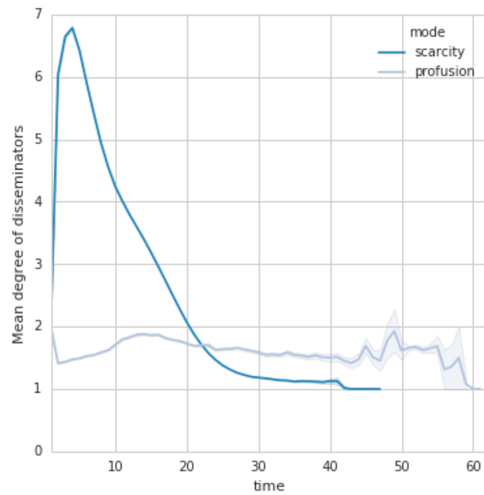


Fig. 5 Evolution of the mean degree of Disseminators over time according to the dissemination mode (with $DPeriod = 10$).

to concentrate on the endogenous properties of the propagation process and so to set aside exogenous properties concerning for instance the nature or the quality of the information. In the first mode the acceptance criteria is based on the profusion of information in the neighborhood while in the second it is the scarcity. The intent was not to decide between the two but rather to determine the respective properties of each of these modes either in term of dynamics or from the fixed-point property. We have conducted agent-based simulations; the following are highlights from the experimental results:

Transmissibility (section 3.1) An information spread by profusion can affect slightly more individuals but with much less certainty than with scarcity.
Incidence curve (section 3.2) When information spreads by scarcity, the intensity of the diffusion peak and when it occurs are not influenced by the DPeriod. However, the higher the closeness of the initial node, the earlier the diffusion peak occurs. On the contrary, when information spreads by profusion, the higher the DPeriod, the later and more intense the peak of diffusion will be. Here, choosing a hub (a node with a high degree or pagerank) as the initial node, influences the intensity of the dissemination peak.
Number of Stiflers at convergence (section 3.3) In both dissemination modes, increasing the DPeriod increases the number of people affected and the duration of dissemination. However, when information spreads by scarcity, it reaches more people faster than by profusion when the DPeriod is low. However, when the DPeriod increases the profusion can reach far more people (sometimes the whole population) than scarcity.
Variability (section 3.4) When information spreads by profusion, we see strong variability in the results as the DPeriod increases. Thus, it is possible to have a significant but more uncertain impact than scarcity, where the number of people affected remains stable in all simulations.

Role of the hubs (section 3.5) The profusion mode tends to flatten the network topology while scarcity takes advantage of the structure by giving a crucial role to the hubs and this from the very beginning of the dissemination process.

All these results show that the choice between profusion and scarcity is the ones that determine to a certain extent how information disseminates in a social network. As such a choice can result from the nature of information (news, rumor, buzz, gossip, fake news, etc), in the future we plan to move the current work forward by trying to establish a link between the method of dissemination and the nature of the information.

References

1. Albert, R., Barabási, A.L.: Statistical mechanics of complex networks. Rev. Mod. Phys. **74**, 47–97 (2002). DOI 10.1103/RevModPhys.74.47. URL `https://link.aps.org/doi/10.1103/RevModPhys.74.47`
2. Borge-Holthoefer, J., Meloni, S., Gonçalves, B., Moreno, Y.: Emergence of Influential Spreaders in Modified Rumor Models. Journal of Statistical Physics **151**, 383–393 (2013). DOI 10.1007/s10955-012-0595-6
3. Cheng, J.J., Liu, Y., Shen, B., Yuan, W.G.: An epidemic model of rumor diffusion in online social networks. The European Physical Journal B **86**(1), 29 (2013). URL `http://link.springer.com/10.1140/epjb/e2012-30483-5`
4. Chmiel, A., Sienkiewicz, J., Thelwall, M., Paltoglou, G., Buckley, K., Kappas, A., Hołyst, J.A.: Collective emotions online and their influence on community life. PLoS ONE **6**(7) (2011). DOI 10.1371/journal.pone.0022207
5. Collard, M., Collard, P., Brisson, L., Stattner, E.: Rumor spreading modeling: Profusion versus scarcity. In: ASONAM 2015 : IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, pp. 1547–1554 (2015)
6. Daley, D.J., Kendall, D.G.: Epidemics and rumours. Nature **204**(4963), 1118 (1964). DOI 10.1038/2041118a0. URL `http://dx.doi.org/10.1038/2041118a0`
7. Doer, B., Fouz, M., Friedrich, T.: Why rumors spread so quickly in social networks. Communications of the ACM **55**(6), 70 (2012). DOI 10.1145/2184319.2184338
8. Garcia, D., Kappas, A., Küster, D., Schweitzer, F.: The Dynamics of Emotions in Online Interaction. To appear in Royal Society Open Science pp. 1–26 (2016). DOI 10.1098/rsos.160059
9. Guadagno, R.E., Rempala, D.M., Murphy, S., Okdie, B.M.: What makes a video go viral? An analysis of emotional contagion and Internet memes. Computers in Human Behavior **29**(6), 2312–2319 (2013). DOI 10.1016/j.chb.2013.04.016. URL `http://dx.doi.org/10.1016/j.chb.2013.04.016`
10. He, Z., Cai, Z., Wang, X.: Modeling Propagation Dynamics and Developing Optimized Countermeasures for Rumor Spreading in Online Social Networks. Proceedings - International Conference on Distributed Computing Systems **2015-July**, 205–214 (2015). DOI 10.1109/ICDCS.2015.29
11. Kermack, W.O., McKendrick, A.G.: A contribution to the mathematical theory of epidemics. Proceedings of the Royal Society of London. Series A **115**, 700–721 (1927)
12. Schweitzer, F., Garcia, D.: An Agent-Based Model of Collective Emotions in Online Communities ties. European Physical Journal B (September) (2008)
13. Xia, Z., Huang, L.: Emergence of social rumor: Modeling, analysis, and simulations. Computational Science pp. 90–97 (2007)
14. Zhao, X., Wang, J.: Dynamical Model about Rumor Spreading with Medium. Discrete Dynamics in Nature and Society **2013**, 1–9 (2013). DOI 10.1155/2013/586867. URL `http://www.hindawi.com/journals/ddns/2013/586867/`