

# The Impact of Social Curiosity on Information Spreading on Networks

Didier A. Vega-Oliveros<sup>\*‡</sup>, Lilian Berton<sup>†</sup>, Federico Vazquez<sup>‡</sup>, and Francisco A. Rodrigues<sup>\*</sup>

<sup>\*</sup>Instituto de Ciências Matemáticas e de Computação,  
Universidade de São Paulo - Campus de São Carlos,  
PB 668, Zip code 13560-970 São Carlos, SP, Brazil.  
Email: davo, francisco@icmc.usp.br

<sup>†</sup> Universidade Federal de São Paulo, Campus de São José dos Campos  
Zip code 12247-014 São José dos Campos, SP, Brazil.

<sup>‡</sup>Instituto de Física de Líquidos y Sistemas Biológicos, IFLYSIB,  
Universidad Nacional de la Plata - 1900 La Plata, Argentina.

**Abstract**—Most information spreading models consider that all individuals are identical psychologically. They ignore, for instance, the curiosity level of people, which may indicate that they can be influenced to seek for information given their interest. For example, the game Pokémon GO spread rapidly because of the aroused curiosity among users. This paper proposes an information propagation model considering the curiosity level of each individual, which is a dynamical parameter that evolves over time. We evaluate the efficiency of our model in contrast to traditional information propagation models, like *SIR* or *IC*, and perform analysis on different types of artificial and real-world networks, like Google+, Facebook, and the United States roads map. We present a mean-field approach that reproduces with a good accuracy the evolution of macroscopic quantities, such as the density of stiffers, for the system's behavior with the curiosity. We also obtain an analytical solution of the mean-field equations that allows to predict a transition from a phase where the information remains confined to a small number of users to a phase where it spreads over a large fraction of the population. The results indicate that the curiosity increases the information spreading in all networks as compared with the spreading without curiosity, and that this increase is larger in spatial networks than in social networks. When the curiosity is taken into account, the maximum number of informed individuals is reached close to the transition point. Since curious people are more open to a new product, concepts, and ideas, this is an important factor to be considered in propagation modeling. Our results contribute to the understanding of the interplay between diffusion process and dynamical heterogeneous transmission in social networks.

## I. INTRODUCTION

Information spreading research relates to understanding how rumors, news and behaviors spread on a large scale in a short time. The diffusion of information is a ubiquitous process in the society, where people are interacting with each other all the time, making contact and exchanging information in the public

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

ASONAM '17, July 31 - August 03, 2017, Sydney, Australia

© 2017 Association for Computing Machinery.

ACM ISBN 978-1-4503-4993-2/17/07?/\$15.00

<http://dx.doi.org/10.1145/3110025.3110039>

transport, at work or school, the Internet and at home [1], [2]. In this way, the maximization of the transmission of some information on social networks and the society is important for scenarios like viral marketing, political and ideological campaigns, among others.

After the establishment of Network Sciences [3], [4], more accurate models have been introduced for analyzing the spread of information as social contagion processes [5]–[9]. Within these models, the *SIR* epidemic model has been applied extensively in research areas like opinion formation, decision making, and information or rumor propagation [5], [6]. The spreading of information can be approached as a psychological contagion where an idea passes from person to person and “contaminates” the mind of many individuals [9]. In the classical approach people are divided into three groups or states: ignorant (those not aware of the information), spreaders (those who spread the information), and stiffers (those who know the information but stop trying to convince others).

However, most models on spreading processes consider homogeneous or static transmission probabilities [4]–[6], i.e., each individual has a fixed likelihood of transmission over time. Moreover, the curiosity has been disregarded in these models until now, although it is expected to have an impact on diffusion processes: the more surprising a rumor, the more rapidly and extensively it is likely to spread. Also, the more people know the information and the more sources that spread it, the higher the likelihood of accepting that information. The curiosity is psychologically defined as the desire of knowledge and experiences that lead to exploratory behavior and the acquisition of new information [10]–[12]. It is also associated with the reward of two kind of triggers, the cognitive curiosity and sensory curiosity [10]. A person often gets curious about some behaviors of his/her friends; a phenomenon known as the social curiosity in psychology [12].

In this paper, we aim to model the information spreading on networks taking into account the human curiosity. We evaluated how information (for instance a rumor) propagates considering the social curiosity, which is high if more friends are talking about the spreading rumor (Figure 1). One moti-

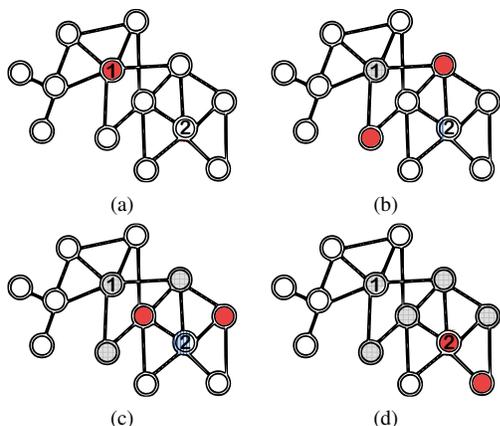


Fig. 1. (Color online) Illustration of four consecutive time steps in the Social Curiosity model for information spreading on networks. (a) Vertex 1 is the only initial spreader (red circle). Vertex 2 is ignorant (white circle) and has no curiosity because none of its neighbors knows about the information. (b) Vertex 1 randomly convinces two neighbors and it turns into stifter (gray circle). Vertex 2 starts to become curious about the information (blue bars inside the circle) since one of its five neighbors is informed. (c) Given that already three neighbors of vertex 2 know the information, its curiosity level and thus the chances to get the information are increased. (d) Eventually, vertex 2 becomes informed due to its curiosity and the contact with spreaders.

vational example is the rapid outbreak of the game Pokémon GO [2], [13] launched by Niantic in July 2016 and that is based on the 1990's Pokémon video games. It became a social phenomenon that achieved a high popularity in a short time, being the app in the iOS App Store with more downloads in one week than any other in the same period, and one of the most commented topics in on-line and traditional magazines, Social Networks, among others [2]. This success is probably due to the fact that the game promotes an interesting mix of virtual reality and socialization activities in public spaces where players can interact, which arouses the curiosity of people that are not part of the game [12]. The probability that a person becomes player increases if more of its neighbors or friends are playing it. In this article, we show by analytical and numerical simulations that including the psychological characteristics of individuals, such as the curiosity, improves the diffusion of information in traditional spreading models.

The major contributions of this work are: *i*) we provide a more detailed and realistic description of information spreading processes with the combination of the curiosity mechanism and the *SIR* model of epidemics. *ii*) We define a dynamical measure of the social curiosity for each subject in its neighborhood. *iii*) A mathematical model for information spreading is presented, including different probabilities of contagion and the dynamical curiosity that evolves over time. *iv*) We provide an analytical solution for the final density of informed individuals, which is  $\mathcal{O}(1)$  and agrees quite well with simulations results. *v*) We analyze the impact of curiosity on information spreading, such as the critical threshold, the peak of spreaders and the final density of informed individuals, in artificial and real-world networks. The results indicate that depending on the transmission parameters, the social curiosity improves the

propagation process in nearly twice the expected value of the non-curiosity case. Besides, the curiosity increases the peak in the number of spreaders reached during the propagation process and speeds up the propagation. These results suggest that the social curiosity is an important feature to be considered in spreading processes, by marketing campaigns and in the analysis of information spreading models.

## II. RELATED WORKS

Several models have been proposed for modeling propagation dynamics on networks [5]–[9], [14]–[17]. These models consider some assumptions about the propagation process and network structure, like the degree correlation or distribution, classes of vertices, among others [5]–[8]. The majority of the spreading models consider only homogeneous or fixed transmission probabilities, i.e., each vertex has the same likelihood of transmitting the information [6]–[8]. For instance, in [9] the authors propose a discrete-time model for rumor propagation with heterogeneous transmission probabilities, but still, each vertex has a constant probability over time. Other transmission models consider characteristics like short-term of immunity or steady active state [4], and spreaders' procrastination in the transmission of information by heterogeneous delays [14]. Also, there are some proposals that introduced disbelieving or forgetting mechanism [15], lost of interest [16], apathy and many others [6]. Regarding the contagion dynamic, the diffusion process can be spontaneous or by different kind of contact interaction or adaptation among the subjects [17].

Previous studies analyzed the effect of the new characteristic in the propagation process and examined numerical and dynamical properties on top of the different type of networks. They looked for critical thresholds, computational cost and accuracy in the results [6]. However, these studies of information spreading assumed fixed parameters and the curiosity mechanism was also disregarded [5]–[8]. Here, we consider the curiosity as an important factor to determinate the speed and coverage of information spreading on networks.

## III. CONCEPTS AND METHODS

An *information system* can be represented as a network (or a graph) consisting of two elements: actors (individuals or objects that are the elements of the system) and the connections (interactions, relationships, or social ties) [4]. Formally, let's consider a network  $G = (V, E)$ , where  $V$  is the set of  $|V| = N$  vertices,  $E$  is the set of  $|E| = M$  edges or connections. The edges can be directed, indicating the flow of the relation, or undirected. Two vertices are called *neighbors* if they are connected by an edge.

Classical propagation models consider that nodes interact in the same way during the dynamic. The model's parameter  $\beta$  (for propagating or transmitting the signal) and  $\mu$  (for recovering or stopping the propagation) are constants and invariant during the process. In the *SIR* model [4], [6], the susceptibles ( $S$ ) are those who remain unaware of the information, the spreaders ( $I$ ) are the influencers who disseminate the information, and the stiflers ( $R$ ) are those who know the

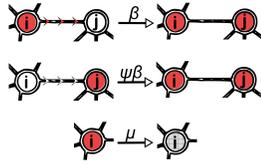


Fig. 2. (Color online) Dynamical rules of the proposed model. Vertex  $i$  is making contact with the neighbor  $j$ , in which: the first line,  $i$  is a spreader (red circles) that convinces its ignorant neighbor (white circles) with probability  $\beta$ ; the second line,  $i$  is an ignorant that asks for information to its spreader neighbor with probability  $\Psi$  and get convinced with probability  $\beta$ . The third line is the spontaneous transition of  $i$  to stop the propagation of the information (stiffer state in gray circles).

information but lose the interest in the spreading process. The influencers and the stifiers are part of the group of informed individuals. Also, the transition to the stiffer state happens spontaneously [6]. Whenever a spreader  $i$  meets a susceptible neighbor  $j$ , the latter become spreader with probability  $\beta$ . Spreaders turn into stiffer with probability  $\mu$ .

The effective spreading rate [6], [18], [19]  $\lambda = \beta/\mu$  is the contagion strength or infectivity level of the propagation. The final density of informed individuals is equivalent to the combination of  $\beta$  and  $\mu$  that hold the same  $\lambda$  value [4], [9], [18]. In epidemics, this is known as the basic reproductive ratio [4] and a global endemic phase is reached by infectivity levels above a critical threshold ( $\lambda_c$ ), while no endemic phase appear whenever the contagion strength is below  $\lambda_c$ . For random and uncorrelated networks,  $\lambda_c = \langle k \rangle / \langle k^2 \rangle$  [6], where  $\langle k \rangle$  and  $\langle k^2 \rangle$  are the mean and the second moment of the degree distribution, respectively.

Another propagation model is the Independent Cascade (*IC*) model [7], [8]. The difference with the *SIR* model yields in the transition to the stiffer (inactive spreaders) state [5], which happens instantly in the next iteration step. In fact, the *IC* can be approached as a variation of the *SIR* model [20]. Both models are spreader-centric, i.e., the contagion occurs given the contact interaction of the spreaders. In the *IC* model, the spreaders try to pass the information only once, which is equivalent to the *SIR* with parameter  $\mu = 1$ . The models are also incremental monotonic functions, and they can consider heterogeneous probabilities in the diffusion [8], [9], [19].

#### IV. CURIOSITY IN SPREADING PROCESS

We present a discrete-time approach including the social curiosity property, centered in the propagation dynamic of each vertex on arbitrary network structure. The approach follows the *SIR* states and the social curiosity is as a parameter that evolves over time. This parameter introduces a susceptible-centric approach, affecting the behavior of each susceptible vertex. We consider vertices with different curiosity levels and this approach is rooted on the observation that activity patterns are essentially heterogeneous [9] and people do not behave in the same way when asking for information.

##### A. Curiosity measurement

The personality, as an inherent characteristic of an individual, has a role in the organization of people on the net-

work, and in turn, they position reinforce specific personality traits [1]. A range of network-based features is correlated with specific characteristics of individuals [1], [2], [21]. For instances, the Extroversion, one of the Five-Factor Model of personality [1], [2], is positively correlated with the number of friends or connections of a person [1]. In particular, the Openness to experience refers to an individual's curiosity and willingness to engage in new experiences. The Extroversion and Openness are positively related with the capacity of higher social engagement [1]. The before is in accordance with a social study about the users of Pokémon GO [2], where was reported that individuals with high level of Openness also were the early adopters.

In the case of epidemic or information spreading process, the "early adopters" are correlated to the network centrality, where the most central nodes are reached at an early stage [20]. This is explained by the higher interaction of individuals with their pairs and the update of their states as a consequence of these interactions. This property is known as the homophily phenomena or assortative mixing [20], the tendency of similar nodes to be connected, which is the main characteristic of social relationship.

**Definition.** (*Curiosity Strength*) The curiosity strength  $\Psi$  is the probability that a susceptible vertex asks a spreader neighbor about some information.

Thus, given that the Openness to experience individuals are part of the group of early adopters and they are somehow positively correlated with the connection structure of the network, we can consider that higher the scores in Openness to experience, the more strongly connected the individual. For a more realistic scenario, we define the social curiosity for a given vertex ( $\Psi_i(t)$ ) as the fraction of informed neighbors over time (Figure 1). Thus, the higher the level of social curiosity, the higher the probability of a susceptible vertex makes contact with a spreader. Therefore, the chances of the susceptible vertices take an action for searching for information depends on the propagation process and the size of their networks.

When  $\Psi = 0$ , we recover the *SIR* and the *IC* models. The Utopian case happens when all the individuals are always 100% curious ( $\Psi = 1$ ). The critical threshold for the maximum curiosity case,  $\lambda_c^1$ , and the critical threshold of non-curiosity case,  $\lambda_c^0$ , are the lower and upper bound of the critical threshold for the proposed social curiosity measure ( $\Psi_i(t)$ ), i.e.,  $\lambda_c^1 < \lambda_c^{\Psi_i(t)} \leq \lambda_c^0$ .

##### B. Monte Carlo simulation

We consider a constant population of  $N$  vertices in all time steps. The spreading dynamic is a stochastic process where each vertex has a probability of being in a possible state over time. The density of susceptible  $S(t)$ , spreader  $I(t)$  or stiffer  $R(t)$  vertices satisfies  $S(t) + I(t) + R(t) = 1$  for all time steps. At each time step, all spreaders uniformly try to infect their neighbors with probability  $\beta$ , or stop the diffusion with probability  $\mu$ . The dynamical rules that define the general diffusion process are depicted in Figure 2.

**Algorithm 1** MC for information spreading with curiosity.

---

**Input:**  $G, S(0), \mathcal{I}(0), \mathcal{R}(0), \beta, \mu, \Psi$   
*Initialization* :  $t \leftarrow 0$

- 1: **repeat**
- 2:    $t \leftarrow t + 1$
- 3:    $S(t) \leftarrow S(t-1), \mathcal{I}(t) \leftarrow \mathcal{I}(t-1), \mathcal{R}(t) \leftarrow \mathcal{R}(t-1)$
- 4:   **for all** vertex  $i$  such that  $i \in \mathcal{I}(t-1)$  **do**
- 5:     **for all** neighbor  $j$  of  $i$  such that  $j \in S(t-1)$  **do**
- 6:       **if**  $i$  infected  $j$  with probability  $\beta$  **then**
- 7:          $S(t) \leftarrow S(t) \cup \{j\}$
- 8:          $\mathcal{I}(t) \leftarrow \mathcal{I}(t) \cup \{j\}$  // update the states
- 9:       **end if**
- 10:    **end for**
- 11:    **if**  $i$  gets recovered with probability  $\mu$  **then**
- 12:        $\mathcal{I}(t) \leftarrow \mathcal{I}(t) - \{i\}$
- 13:        $\mathcal{R}(t) \leftarrow \mathcal{R}(t) \cup \{i\}$
- 14:    **end if**
- 15:    **end for**
- 16:    **for all** vertex  $i$  such that  $i \in S(t-1)$  **do**
- 17:     **for all** neighbor  $j$  of  $i$  such that  $j \in \mathcal{I}(t-1)$  **do**
- 18:       **if**  $i$  gets curious with probability  $\Psi(i)$  **and**  
        $i$  gets infected by  $j$  with probability  $\beta$  **then**
- 19:          $S(t) \leftarrow S(t) - \{i\}$
- 20:          $\mathcal{I}(t) \leftarrow \mathcal{I}(t) \cup \{i\}$
- 21:       **end if**
- 22:     **end for**
- 23:    **end for**
- 24: **until**  $I(t) = 0$
- 25: **return**  $R(t)$

---

The Monte Carlo (MC) algorithm for the simulations consider the following variables at time  $t = 0$ : (i)  $S(0)$  that represents the set of ignorants, (ii)  $\mathcal{I}(0)$  the set of spreaders or seeds and (iii)  $\mathcal{R}(0)$  the set of stiflers. Vertices belong to only one of the states over time, i.e., if  $i \in \mathcal{I}(t)$ , then  $i \notin [S(t) \cup \mathcal{R}(t)]$ , and always  $i \in [S(t) \cup \mathcal{I}(t) \cup \mathcal{R}(t)]$ . In terms of MC implementation, we assume that infection and recovering do not occur during the same time step. The contact interaction happens with all the neighbors of an informed node in each time step, known as reactive process (RP) [22]. The macroscopic densities of susceptible ( $S(t)$ ) over time is calculated as  $S(t) = |S(t)|/N$ , where  $|S(t)|$  is the size of the set. For the other states, we adopt a similar approach.

The end of the simulation is reached when  $\mathcal{I}(\infty) = \{\emptyset\}$ . Thus, the final fraction of informed individuals is  $R(\infty) = |\mathcal{R}(\infty)|/N$  or  $R(\infty) = 1 - |S(\infty)|/N$ . The algorithm for the information spreading with curiosity (Algorithm 1) needs as input the network  $G$ , the set of initial states of the vertices  $S(0), \mathcal{I}(0), \mathcal{R}(0)$ , the transmission probabilities  $\beta, \mu$  and the vector  $\Psi$  of node curiosity. The sets of vertices states can be approached as a structured list of nodes in a specific time step. For illustrative purpose, the steps (7 - 8), (12 - 13) and (19 - 20) are the same procedure for updating the state of the vertex, but it can be generalized in a separated function. Also, the part of propagation by curiosity, steps (16 - 23), can be refactored in a main loop (step 4) that considers all the vertices.

The complexity of the MC algorithm is determined by: (i) the maximum number of steps  $T$  in which  $I(T) = 0$ ; (ii) the number of iterations in the loops of the classical (step 4) and curiosity (step 16) propagation, which are  $|S(t) \cup \mathcal{I}(t)|$  and can be approximated to  $N/2$ ; (iii) the computational cost of

TABLE I  
 TOPOLOGICAL PROPERTIES OF THE NETWORKS CONSIDERED HERE.

	Network	$N$	$\langle k \rangle$	$\langle k^2 \rangle$	$\langle \ell \rangle$	$\rho$
Artificial	ER	1000	11.9	155	3.03	0.001
	BA	1000	11.9	284	2.84	-0.032
	SSF	1000	11.9	243	4.075	0.19
Real	<i>USAroad</i>	6443	3.09	10.41	50.84	0.19
	<i>Google+</i>	23613	3.32	1251.67	4.03	-0.389
	<i>Facebook</i>	63392	25.77	2256.80	4.32	0.177

copying and updating the sets, which can be considered as a constant value  $C_1$ ; and (iv) the cost of contacting the neighbors of the vertices, which in average for the whole networks is equivalent to  $\langle k \rangle$ . Thus, the complexity of the simulation algorithm is  $\mathcal{O}(T * N/2 * C_1 * (2\langle k \rangle + 1)) \equiv \mathcal{O}(T * N * \langle k \rangle)$ .

## V. EXPERIMENTS

The impact of the curiosity in the propagation process was analyzed on top of three network models, presenting different degree distributions, and three real-world networks. We consider the Barabási-Albert [3] (BA), Erdős-Rényi [23] (ER) and Barthélemy spatial scale free [24] (SSF) models, and for real-world networks we adopt the *Google+* [25], *Facebook* [21] and *USAroad* [26], the road network of USA. Their topological features are summarized in Table I, with the corresponding average degree  $\langle k \rangle$ , second moment of degree distribution  $\langle k^2 \rangle$ , average shortest path length  $\langle \ell \rangle$  and assortativity coefficient that measures degree-degree correlation  $\rho$ . For the real-world networks we take the main component and assumed them as undirected networks.

### A. Simulation Setup

We analyze the proposed model of information spreading considering the macroscopic scale of the propagation. For this purpose, the initial setup for each simulation is  $S(0) = 1 - 1/N$ ,  $I(0) = 1/N$  e  $R(0) = 0$ . We calculate the corresponding density of stiflers  $R(t)$  and spreader  $I(t)$  over time by taking each vertex as the only initial seed. For this value was considered an average of 60 realizations. The macroscopic density of stiflers  $R$  is calculated by averaging the number of stifter among all nodes. The macroscopic measures represent the propagation values for the whole network. Without lack of generality, we adopt the recovery probability  $\mu = 1$  [6], [20].

### B. Simulation results

The impact of the social curiosity is evaluated on the information spreading process. Initially, we analyze whether the phase transition for the endemic state is affected by the inclusion of the curiosity parameter. For different values of  $0 < \lambda \leq 1$ , the average density of stifter and the critical threshold are evaluated, as shown in Figure 3.

We observe that independently of the network structure, i.e., scale-free, spatial or Poisson networks, the curiosity phenomenon improves the overall propagation. For the artificial networks (Fig. 3a, 3b, 3c), the critical threshold of the Utopian case ( $\lambda_c^1$ ) is very close to half of the corresponding threshold

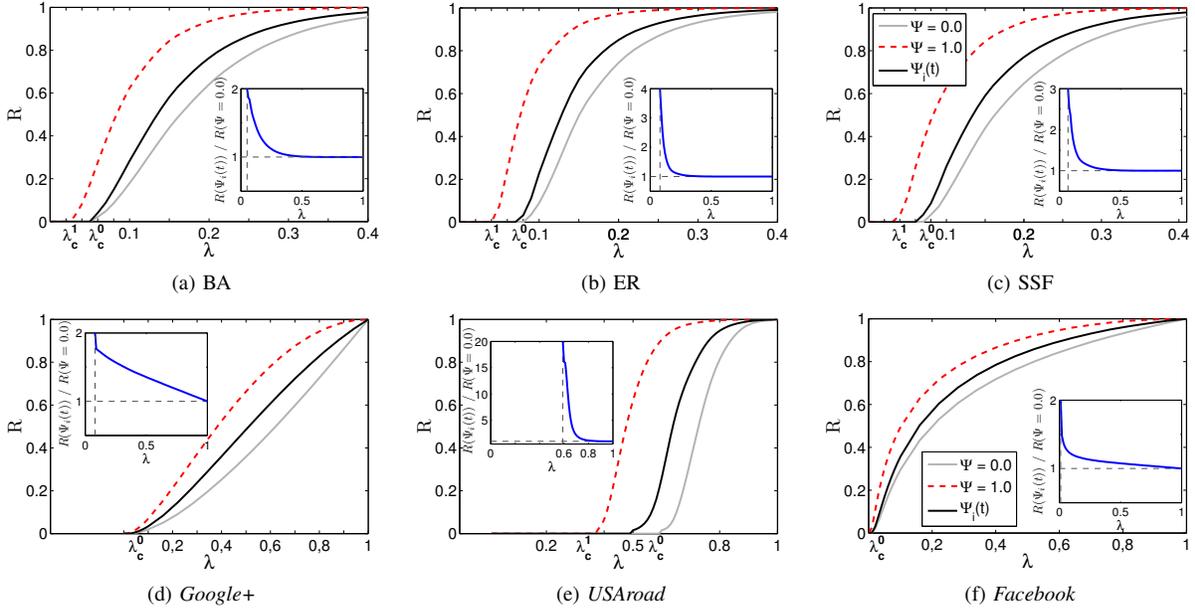


Fig. 3. (Color online) Phase diagram of the macroscopic density of stifiers as a function of parameter  $\lambda$  for (a) a Barabási-Albert (BA), (b) Erdős-Rényi (ER), (c) spatial scale-free (SSF), (d) *Google+*, (e) *USAroad* and (f) *Facebook* networks. The artificial networks have the same size  $N = 1000$  and average degree  $\langle k \rangle = 11.9$ , and the real-world networks are described in Table I. The classical propagation is recovered when  $\Psi = 0$ , whereas the complete curiosity case occurs for  $\Psi = 1$ .  $\Psi_i(t)$  means the dynamical curiosity over time (Figure 1). The  $R$  values are the average of the final density of stifiers among all vertices. The inset figures are the quotient between the results of  $R$  considering the dynamical curiosity and the respective  $R$  of the classical propagation.

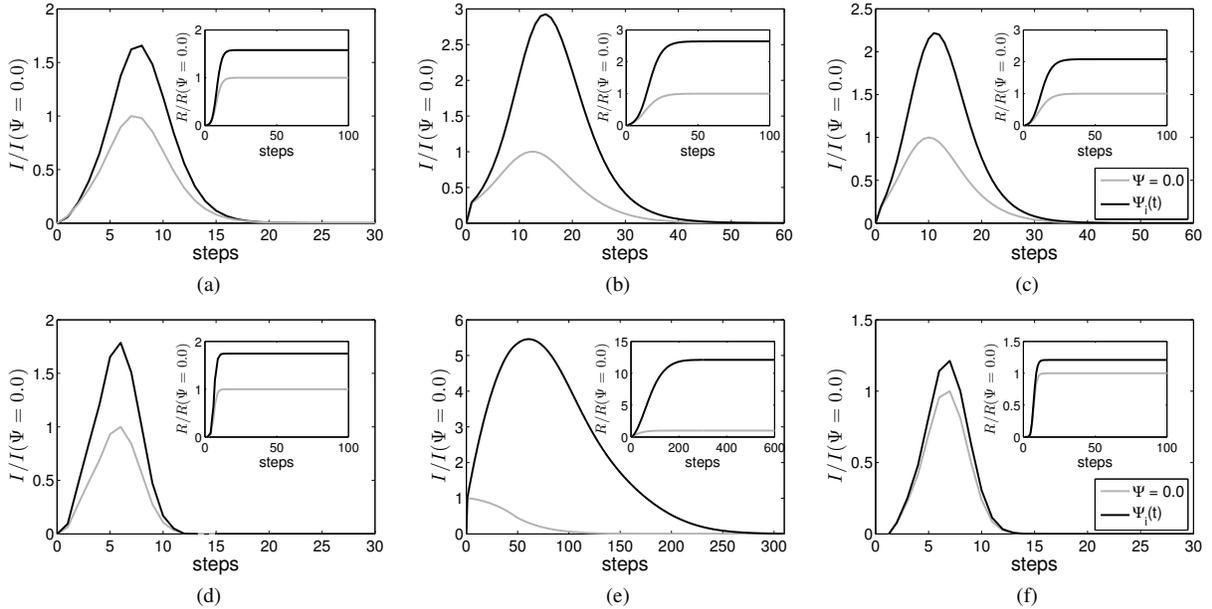


Fig. 4. (Color online) Macroscopic density of spreaders and stifiers over time for (a) a Barabási-Albert (BA), (b) Erdős-Rényi (ER), (c) spatial scale-free (SSF), (d) *Google+*, (e) *USAroad* and (f) *Facebook* networks. Parameter  $\lambda$  was fixed to a value 0.61 for *USAroad* and 0.1 for the remaining networks. In the figures we have the curves of the average peak of spreaders and the insets are the corresponding average density of stifiers over time. The averages are calculated considering each vertex as a initial seed.

of the non-curiosity configuration ( $\lambda_c^0$ ), when  $\Psi = 0$ . In the real-world networks, the critical thresholds  $\lambda_c^0$  and  $\lambda_c^1$  are very similar and tend to zero, (Fig. 3d, 3f). This occurs due to real-world networks have a scale-free organization, characterized by a power-law degree distribution [3], [4] ( $P(k) \sim k^{-\gamma}$  with  $\gamma < 3$ ). In these networks, the critical threshold  $\lambda_c \rightarrow 0$  when

$N \rightarrow \infty$ . The *USAroad* is the most homogeneously distributed network, having low variance in the degree distribution and larger average shortest path length (Table I). The dynamical curiosity ( $\Psi_i(t)$ ) little affects the critical threshold in comparison with the non-curiosity case ( $\Psi = 0$ ). The final density of stifiers is improved in more than twice than the results reached

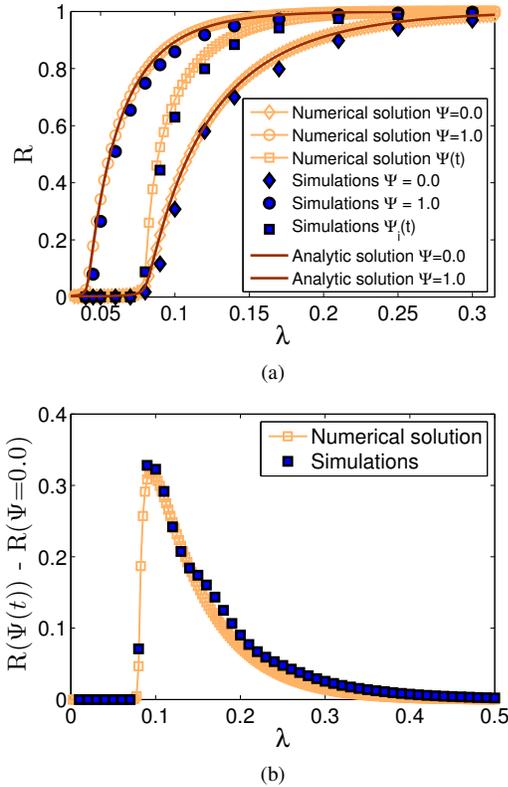


Fig. 5. (Color online) (a) Final density of stiflers  $R$  vs  $\lambda = \beta/\mu$  for curiosity  $\Psi = 0.0$  (diamonds),  $\Psi = 1.0$  (circles) and  $\Psi(t) = I(t) + R(t)$  (squares). Filled symbols correspond to Monte Carlo simulations on Erdős-Renyi networks of size  $N = 10^4$  with initial density of spreaders  $I(0) = 10^{-3}$ , whereas empty symbols represent the numerical integration of Eqs. (1). Solid lines are the analytical solutions from Eqs. (9). (b) Difference between the final density of stiflers in the dynamic case  $\Psi(t)$  and in the  $\Psi = 0$  case, obtained from the data of panel (a).

by the non-curiosity case in lower values of  $\lambda$ , and tend to be ineffective for larger values (inset Figure 3).

To study the evolution of the diffusion process, we evaluate the curves of spreaders achieved in each network, and the insets are the corresponding density of stiflers (Figure 4). Parameter  $\lambda$  is fixed to a value close to  $\lambda_c^0$ , which is 0.61 for *USAroad* and 0.1 for the remaining networks. The peak of spreaders and the final density of stiflers for the social curiosity are superior in all the case to the classical spreading approach ( $\Psi = 0$ ). The  $\Psi_i(t)$  curves of the peak of spreaders and density of stiflers last the same time than the classical approach. This indicates that the curiosity can improve the extent and velocity of the propagation. Specifically, the peak of the spreaders is twice or higher the respective peaks of the classical approach, except for the *Facebook* network, which is 23% higher. The  $\Psi_i(t)$  curves also achieve more than twice the final density of stiflers of the  $\Psi = 0.0$  curves, for the *ER*, *SSF* and *USAroad* networks (inset Fig. 4b, 4c, 4e); it achieves 50% or higher density of stiflers in the *BA* and *Google+* networks (inset Fig. 4a, 4d); and it is 20% higher in the *Facebook* network (inset Fig 4f).

The results suggest that adopting the social curiosity in the diffusion process can improve the expected density of

informed individuals, as well, it increases the peak of spreaders and the velocity of the dynamic. The impact in the results depend on the network topology and the effective spreading rate  $\lambda$  employed. For instances, for spatially distributed networks like *USAroad*, we obtain very significant results than in social networks like *Google+* or *Facebook*.

## VI. MEAN-FIELD ANALYSIS

The system behavior can be described, at the mean-field level, by the following set of coupled recurrence equations for the time evolution of the densities of individuals  $S$ ,  $I$  and  $R$ :

$$S(t+1) = S(t) - \gamma S(t), \quad (1a)$$

$$I(t+1) = I(t) + \gamma S(t) - \mu I(t), \quad (1b)$$

$$R(t+1) = R(t) + \mu I(t), \quad (1c)$$

which allows calculating their values at time  $t+1$  from the previous time  $t$ . These equations describe an infinitely large system, where fluctuations due finite-size effects are neglected. Here

$$\gamma(t) \equiv 1 - [(1 - \beta)(1 - \beta\Psi(t))]^{(k)I(t)} \quad (2)$$

is the probability that a susceptible vertex gets informed from one of its spreader neighbors, and  $\Psi(t)$  is the curiosity level, i. e., the probability that a susceptible vertex requests the information to each spreader. Within a mean-field approach, the curiosity is approximated as the mean fraction of informed neighbors of a given vertex  $i$ ,  $\Psi(t) \simeq I(t) + R(t)$ , while the exponent  $(k)I(t)$  in  $\gamma$  is the estimated mean number of spreader neighbors of  $i$ . In Eq. (1b), the gain term  $\gamma S(t)$  represents the fraction of susceptible nodes that become spreaders, while the loss term  $-\mu I(t)$  corresponds to the fraction of spreaders that become stiflers in a time step.

In Fig. 5(a) we plot the stationary value of  $R$  vs  $\lambda = \beta/\mu$  obtained from Monte Carlo simulations (filled symbols) and compare with results from the numerical integration of Eqs. (1) (empty symbols). Simulations were performed on Erdős-Renyi networks of  $N = 10^4$  nodes, starting with a small fraction  $I(0) = 10^{-3}$  of spreaders uniformly distributed over the network. Diamonds, squares and circles correspond to curiosity  $\Psi = 0$ ,  $\Psi(t) = I(t) + R(t)$  and  $\Psi = 1$ , respectively. We observe that the agreement between theory and simulations is quite good for the entire range of  $\lambda$ . The mean-field solution is also able to capture the transition point  $\lambda_c$  at which the final density of stiflers becomes larger than zero, and gives the same qualitative behavior as the corresponding curves for all networks shown in Fig. 3. Figure 5(b) shows the difference in the values of  $R$  corresponding to the dynamic case  $\Psi(t)$  and  $\Psi = 0$ . We observe that  $R(\Psi(t))$  is larger than  $R(\Psi = 0)$  for the entire range of  $\lambda$ , and that the difference becomes very large close to the transition point. This shows that the spreading of the information is largely increased when the curiosity mechanism is taken into account.

We now perform the analysis of Eqs. (1) starting from the simplest case in which the curiosity  $\Psi$  is constant over time. To solve Eqs. (1) we make two approximations. First, we approximate the difference  $S(t+1) - S(t)$  by the time

derivative  $\frac{dS(t)}{dt}$  in the continuous time limit, and analogously for  $I$  and  $R$ . Next, we assume that  $I(t)$  is small and expand  $\gamma(t)$  from Eq. (2) to first order in  $I \ll 1$  and obtain

$$\gamma(t) \simeq -\langle k \rangle \ln[(1 - \beta)(1 - \Psi\beta)]I(t). \quad (3)$$

Then, Eqs. (1) become

$$\frac{dS(t)}{dt} = \langle k \rangle \ln[(1 - \beta)(1 - \Psi\beta)]I(t)S(t) \quad (4a)$$

$$\frac{dI(t)}{dt} = -\langle k \rangle \ln[(1 - \beta)(1 - \Psi\beta)]I(t)S(t) - \mu I(t), \quad (4b)$$

$$\frac{dR(t)}{dt} = \mu I(t). \quad (4c)$$

Dividing Eq. (4a) by Eq. (4c) we arrive to the equation

$$\frac{dS}{dR} = \frac{\langle k \rangle}{\mu} \ln[(1 - \beta)(1 - \Psi\beta)]S, \quad (5)$$

which gives a relation between  $S$  and  $R$ . Then, the integration of Eq. (5) gives, after some algebra, the following equation that relates  $S$  and  $R$  at all times:

$$(1 - \beta)(1 - \Psi\beta) = \left(\frac{S}{S_0}\right)^{\frac{\mu}{\langle k \rangle \tilde{R}}} \quad (6)$$

where we have used the initial condition  $S_0 = S(t=0) = 1 - 1/N$  and  $R(t=0) = 0$ . At the stationary state ( $t = \infty$ ),  $R$  and  $S$  take the values  $R(t = \infty) = \tilde{R}$  and  $S(t = \infty) = \tilde{S} = 1 - \tilde{R}$ , respectively, since  $I(t = \infty) = 0$ . Then, the final density of stiflers obeys the equation

$$(1 - \beta)(1 - \Psi\beta) = \left(\frac{1 - R_\infty}{S_0}\right)^{\frac{\mu}{\langle k \rangle \tilde{R}}}. \quad (7)$$

Finally, solving for  $\beta$  leads to the expression

$$\beta = \frac{(1 + \Psi) - \sqrt{(1 + \Psi)^2 - 4\Psi \left[1 - \left(\frac{1 - \tilde{R}}{S_0}\right)^{\frac{\mu}{\langle k \rangle \tilde{R}}}\right]}}{2\Psi}, \quad (8)$$

which relates the stationary value of  $R$  with the propagation force  $\beta$ . For the extreme cases  $\Psi = 0$  and  $\Psi = 1$ , Eq. (8) is reduced to

$$\begin{aligned} \beta &= 1 - \left(\frac{1 - \tilde{R}}{S_0}\right)^{\frac{\mu}{\langle k \rangle \tilde{R}}} && \text{for } \Psi = 0 && \text{and} \\ \beta &= 1 - \left(\frac{1 - \tilde{R}}{S_0}\right)^{\frac{\mu}{2\langle k \rangle \tilde{R}}} && \text{for } \Psi = 1. \end{aligned} \quad (9)$$

In Fig. 5(a) we plot by solid lines the analytical solution  $\tilde{R}$  vs  $\lambda = \beta/\mu$  from Eqs. (9). We observe a very good agreement with the numerical integration of Eqs. (1) (empty symbols). An interesting relation between the two curves can be found from Eqs. (9), that is,  $\tilde{R}(\Psi = 1, \beta) = \tilde{R}(\Psi = 0, 1 - (1 - \beta)^2)$ . This means that the shape of  $\tilde{R}$  vs  $\lambda$  for  $\Psi = 1$  is the same as that for  $\Psi = 0$  with the x-axis shifted to the right by the factor  $[1 - (1 - \mu\lambda)^2]/\mu$ .

Another quantity of interest is the transition point  $\beta_c$  between an ‘‘endemic phase’’ where the information spreads to

a large fraction of the population for  $\beta > \beta_c$ , and a ‘‘healthy phase’’ where the information does not propagate for  $\beta < \beta_c$ . Starting from a situation with a very small fraction of spreaders  $I \ll 1$  and no stiflers  $R = 0$ , Eq. (4b) becomes

$$\frac{dI(t)}{dt} = \delta I(t), \quad (10)$$

with  $\delta \equiv -\langle k \rangle \ln[(1 - \beta)(1 - \Psi\beta)] - \mu$ , and where we have neglected terms of order  $I^2$ . For  $\delta > 0$  ( $\delta < 0$ ),  $I(t)$  grows (decays) exponentially fast. Then, at the transition point is  $\delta = -\langle k \rangle \ln[(1 - \beta_c)(1 - \Psi\beta_c)] - \mu = 0$ , from where arrive to the following expression for  $\lambda_c = \beta_c/\mu$ :

$$\lambda_c^\Psi = \frac{1 + \Psi - \sqrt{(1 + \Psi)^2 - 4\Psi(1 - e^{-\mu/\langle k \rangle})}}{2\Psi\mu}. \quad (11)$$

For  $\Psi = 0$  and  $\Psi = 1$  we obtain, respectively,  $\lambda_c^0 = [1 - e^{-\mu/\langle k \rangle}]/\mu$  and  $\lambda_c^1 = [1 - \sqrt{e^{-\mu/\langle k \rangle}}]/\mu$ . We can easily check that  $\lambda_c^0 = [1 - (1 - \mu\lambda_c^1)^2]/\mu$ , in agreement with the shift of the  $\tilde{R}$  vs  $\lambda$  curves mentioned above.

We now study the case where the curiosity changes over time according to  $\Psi(t) = I(t) + R(t)$ . Expanding  $\gamma$  from Eq. (2) to first order in  $I$  we obtain  $\gamma(t) \simeq -\langle k \rangle \ln[(1 - \beta)(1 - \beta R(t))]I(t)$ . Then, in the continuous time limit Eqs. (1) become

$$\frac{dS(t)}{dt} = \langle k \rangle \ln[(1 - \beta)(1 - \beta R(t))]I(t)S(t) \quad (12a)$$

$$\frac{dI(t)}{dt} = -\langle k \rangle \ln[(1 - \beta)(1 - \beta R(t))]I(t)S(t) - \mu I(t), \quad (12b)$$

$$\frac{dR(t)}{dt} = \mu I(t). \quad (12c)$$

Following an approach similar to the one above for constant  $\Psi$ , the transition point is obtained by linearizing Eq. (12b) around the fixed point  $I = 0$ ,  $R = 0$ . This leads to an equation like Eq. (10) with the same value of  $\delta$ . Therefore, the transition point with dynamical curiosity  $\Psi(t)$  is the same as that with  $\Psi = 0$ . This can be observed in Fig. 5(a) and also in Fig. 3 for all networks, except for the USA road network [panel (e)].

## VII. DISCUSSION AND FINAL REMARKS

The curiosity is an intrinsic feature that varies from person to person. The level of curiosity of people can be increased not only among the most curious subjects but also among the less curious individuals that are willing to adopt a new product or idea. More marketing campaigns are employing curiosity strategies to attract customers [27], [28]. For instance, the game Pokémon GO, the biggest mobile game in US history [28], introduces virtual reality and outdoor activities in public locations, which increase people’s curiosity. Another example is Philips RéAura, a skin rejuvenation brand where the customers can access an interactive site that shows diaries and trial results from real people [27]. The social platform for customer engagement and interaction promotes users’ curiosity about the truth behind the product claims. Stimulation of people’s curiosity can provide better results not only for the

marketing area but also for education and political or health campaigns, among others. In general, people's opinions are influenced by their close contacts, i. e., family and friends, and that are why they tend to spread these ideas more than those they receive from other sources outside of the circle of close relationships.

This work contributes to understanding how social curiosity can improve the potential diffusion of information on networks. We proposed an information spreading model based on the *SIR* dynamics and introduced the social curiosity as a heterogeneous dynamical parameter that evolves over time. Depending on the fraction of informed neighbors, susceptible individuals become curious and interact with the neighborhood to satisfy the curiosity about the new idea or information. We found that the effect of the social curiosity on the propagation of the information depends on the effective spreading rate  $\lambda$ . For values of  $\lambda$  above and close to the healthy-endemic transition point  $\lambda_c$ , the curiosity is able to increase the propagation level in more than twice compared to the classical approach without curiosity. The maximum number of new spreaders reached during the propagation process and the speed of propagation are also increased by the curiosity. However, as  $\lambda$  becomes smaller or larger than  $\lambda_c$ , the social curiosity has no significant effect. These results suggest that the curiosity mechanism reaches its maximum effectiveness when the system is just above the transition point, in the endemic phase. Another important point is the structure of the network. We found that the curiosity has a larger impact on homogeneous and spatial networks than on on-line social and scale-free networks.

The proposed curiosity mechanism might be applicable to other fields like biological networks and animal social behavior models. Another important issue to be studied is the interplay between the curiosity influence –like location-base strategies, the impact of network structure and dynamical parameters, and the personality homophily for the adoption/diffusion of ideas and publicity. More suitable models can be developed adopting the social curiosity and analyzing geolocated phenomena like the Pokémon GO game.

#### ACKNOWLEDGMENTS

This research employed the computing resources of the Center for Mathematical Sciences Applied to Industry (CeMEAI), financed by FAPESP. DAVO acknowledges CNPq (grant 140688/2013-7), FAPESP (grant 2016/23698-1) and the Banco Santander grant of RED Macro Universidades. LB acknowledges CAPES and FAPESP. FV acknowledges support from CONICET (Argentina). FAR acknowledges CNPq (grant 305940/2010-4) and FAPESP (grant 13/26416-9).

#### REFERENCES

- [1] N. Noë, R. M. Whitaker, and S. M. Allen, "Personality homophily and the local network characteristics of facebook," in *2016 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, Aug 2016, pp. 386–393.
- [2] M. E. Tabacchi, B. Caci, M. Cardaci, and V. Perticone, "Early usage of pokémon go and its personality correlates," *Computers in Human Behavior*, vol. 72, pp. 163 – 169, 2017.
- [3] A.-L. Barabási and R. Albert, "Emergence of scaling in random networks," *Science*, vol. 286, pp. 509–512, 1999.
- [4] M. Newman, *Networks: an introduction*. Oxford Uni. Press, Inc., 2010.
- [5] A. Guille, H. Hacid, C. Favre, and D. A. Zighed, "Information diffusion in online social networks: A survey," *ACM SIGMOD Record*, vol. 42, no. 1, p. 17, 5 2013.
- [6] R. Pastor-Satorras, C. Castellano, P. Van Mieghem, and A. Vespignani, "Epidemic processes in complex networks," *Reviews of Modern Physics*, vol. 87, no. 3, pp. 925–979, 8 2015.
- [7] D. Kempe, J. Kleinberg, and . Tardos, "Maximizing the Spread of Influence through a Social Network," *Theory of Computing*, vol. 11, no. 4, pp. 105 – 147, 2015.
- [8] M. A. Porter and J. P. Gleeson, *Dynamical Systems on Networks: A Tutorial*, 1st ed., ser. Frontiers in Applied Dynamical Systems: Reviews and Tutorials. Springer, 2016.
- [9] D. A. Vega-Oliveros, L. da F Costa, and F. A. Rodrigues, "Rumor propagation with heterogeneous transmission in social networks," *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2017, no. 2, p. 023401, 2017.
- [10] M. Zuckerman and P. Litle, "Personality and curiosity about morbid and sexual events," *Personality and Individual Differences*, vol. 7, no. 1, pp. 49 – 56, 1986.
- [11] J. A. Litman, "Curiosity and the pleasures of learning: Wanting and liking new information," *Cognition and Emotion*, vol. 19, no. 6, pp. 793–814, 2005.
- [12] Q. Wu and C. Miao, "Curiosity: From psychology to computation," *ACM Comput. Surv.*, vol. 46, no. 2, pp. 18:1–18:26, 2013.
- [13] M. I. Nick Wingfield, "Pokemon go brings augmented reality to a mass audience," July, 2016. [Online]. Available: <http://www.nytimes.com>
- [14] H. Zhang, A. Kuhnle, H. Zhang, and M. T. Thai, "Detecting misinformation in online social networks before it is too late," in *2016 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, Aug 2016, pp. 541–548.
- [15] L. Zhao, Q. Wang, J. Cheng, Y. Chen, J. Wang, and W. Huang, "Rumor spreading model with consideration of forgetting mechanism: A case of online blogging LiveJournal," *Physica A: Statistical Mechanics and its Applications*, vol. 390, no. 13, pp. 2619–2625, 2011.
- [16] M. Nekovee, Y. Moreno, G. Bianconi, and M. Marsili, "Theory of rumour spreading in complex social networks," *Physica A: Statistical Mechanics and its Applications*, vol. 374, no. 1, pp. 457–470, 1 2007.
- [17] D. Vega-Oliveros, L. Berton, A. Lopes, and F. Rodrigues, "Influence maximization based on the least influential spreaders," in *SocInf 2015, co-located with IJCAI 2015*, vol. 1398, 2015, pp. 3–8.
- [18] H. W. Hethcote, "The Mathematics of Infectious Diseases," *SIAM REVIEW*, 8 2006.
- [19] C. Buono, F. Vazquez, P. A. Macri, and L. A. Braunstein, "Slow epidemic extinction in populations with heterogeneous infection rates," *Physical Review E*, vol. 88, no. 2, p. 022813, 2013.
- [20] D. A. Vega-Oliveros, L. d. F. Costa, and F. A. Rodrigues, "Influence maximization on correlated networks through community identification," may 2017. [Online]. Available: <http://arxiv.org/abs/1705.00630>
- [21] B. Viswanath, A. Mislove, M. Cha, and K. P. Gummadi, "On the evolution of user interaction in Facebook," in *Proc. Workshop on Online Social Networks*, 2009, pp. 37–42.
- [22] D. Vega-Oliveros and L. Berton, "Spreader selection by community to maximize information diffusion in social networks," in *SIMBig 2015*, 2015, pp. 73–82.
- [23] A. Erdos P; Rényi, "On random graphs," *Publicationes Mathematicae*, vol. 6, no. 1, pp. 290–297, 1959.
- [24] M. Barthélemy, "Crossover from scale-free to spatial networks," *Europhysics Letters (EPL)*, vol. 63, no. 6, pp. 915–921, 2003.
- [25] J. McAuley and J. Leskovec, "Learning to discover social circles in ego networks," in *Advances in Neural Information Processing Systems*, 2012, pp. 548–556.
- [26] A. L. Barbieri, G. de Arruda, F. A. Rodrigues, O. M. Bruno, and L. da Fontoura Costa, "An entropy-based approach to automatic image segmentation of satellite images," *Physica A: Statistical Mechanics and its Applications*, vol. 390, no. 3, pp. 512 – 518, 2011.
- [27] G. Wheeler. (2013) Marketing's new mission: Creating a customer curiosity strategy. [Online]. Available: <http://www.mycustomer.com>
- [28] TEDxZurich. (2017) Pokémon go is now the biggest mobile game in u.s. history. [Online]. Available: <http://tedxzurich.com>