Information Is Not a Virus, and Other Consequences of Human Cognitive Limits

Kristina Lerman

Information Science Institute, University of Southern California, Marina del Rey, CA 90292, USA;
lerman@isi.edu; Tel.: +1-310-448-8714; Fax: +1-310-822-0751

Abstract: The many decisions that people make about what to pay attention to online shape the spread of information in online social networks. Due to the constraints of available time and cognitive resources, the ease of discovery strongly impacts how people allocate their attention to social media content. As a consequence, the position of information in an individual's social feed, as well as explicit social signals about its popularity, determine whether it will be seen, and the likelihood that it will be shared with followers. Accounting for these cognitive limits simplifies mechanics of information diffusion in online social networks and explains puzzling empirical observations: (*i*) information generally fails to spread in social media and (*ii*) highly connected people are less likely to re-share information. Studies of information diffusion on different social media platforms reviewed here suggest that the interplay between human cognitive limits and network structure differentiates the spread of information from other social contagions, such as the spread of a virus through a population.

Keywords: social contagion; information diffusion; social media

1. Introduction

The spread of information in online social networks is often likened to the spread of a contagious disease. According to this analogy, information—whether a trending topic, a news story, a song, or a video—behaves much like a virus, "infecting" individuals, who then "expose" their naive followers by mentioning the topic, sharing the video, or recommending the news story, *etc*. These followers may, in turn, become "infected" by sharing the information, "exposing" their own followers, and so on. If each person "infects" at least one other person, information will keep spreading through the network, resulting in a "viral" outbreak, similar to how a spreading virus can create an epidemic that sickens a large portion of the population. The analogy between the spread of a disease and information is the basis of computational methods that attempt to amplify the spread of information in networks by identifying influential "superspreaders" [1–5], and those that make inferences about the network from observations of how information spreads in it [6–10].

One of the simplest and most widely used models of the spread of epidemics in networks is the Independent Cascade Model (ICM) [1,6,11–13]. It describes a process wherein each exposure of a healthy but susceptible individual to a disease by an infected friend results in an independent chance of disease transmission: the more infected friends an individual has, the more likely he or she is to become infected. The model predicts the size of an outbreak (number of individuals infected) in any network for a given value of disease transmissibility (*i.e.*, how easily the disease is transmitted upon exposure). Figure 1 shows the size of outbreaks simulated using the ICM on a social media follower graph (red dots) [14]. The black symbols give the size of simulated outbreaks on a randomly-generated graph with the same degree distribution as the follower graph. The simulated outbreaks are close in size to the theoretically predicted values [15], given by the golden line in Figure 1. There exists a critical

value of transmissibility—the epidemic threshold [16]—below which the contagion dies out, but above which it spreads to a finite portion of the network. The epidemic threshold depends only on structural properties of the network, and not the details of the disease or its transmissibility [17]: specifically, the epidemic threshold is given by the inverse of the largest eigenvalue of the adjacency matrix representing the network [16,18]. Note that even above the epidemic threshold, contagions starting in isolated corners of the network may die out. However, in general, the higher the transmissibility, the farther the contagion spreads, reaching a non-negligible fraction of the network above the epidemic threshold, for example, 10%, 20%, *etc.* of the network.
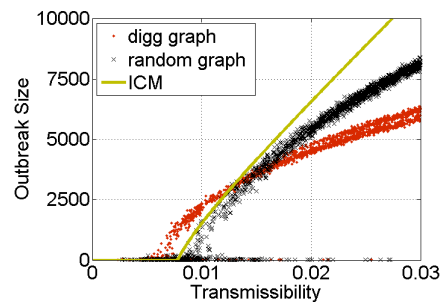
Figure 1. Size of simulated outbreaks on a real-world and random graphs as a function of transmissibility. Contagions are simulated using the independent cascade model (ICM) on the follower graph of the Digg social news platform (**red dots**) and a random graph with the same degree distribution (**black crosses**). The golden line gives theoretically predicted outbreak sizes.

How well does the *social contagion* analogy hold for social media? In this review of empirical studies of information diffusion in social media, I first present evidence that information fails to spread widely in online social networks. The vast majority of outbreaks are very small (see Figure 2), in stark contrast to the predictions of the epidemic model. To explain these findings, I present studies examining the mechanisms of information diffusion, specifically how people respond to multiple exposures to information. The key finding of these studies—that central individuals in an online social network are less susceptible to becoming "infected"—is sufficient to explain why social contagions fail to propagate. The studies also link the reduced susceptibility of central individuals to information overload and human reliance on cognitive heuristics to compensate for the brain's limited capacity to process information. Accounting for how people use cognitive heuristics to decide what information to pay attention to in social media dramatically simplifies dynamics of social contagion and allows for more accurate predictions of how far information will spread online.
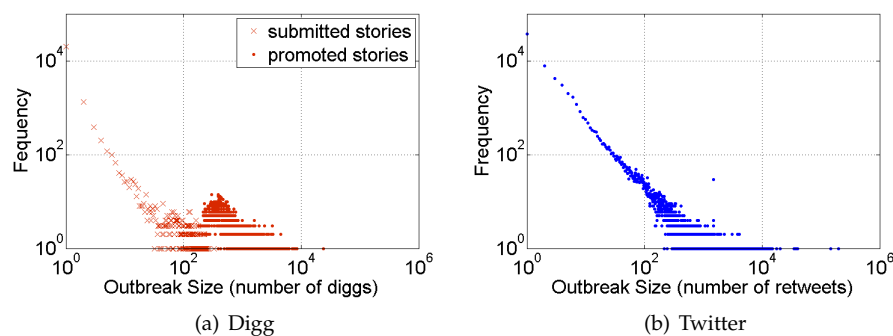
(a) Digg   (b) Twitter

Figure 2. Size of outbreaks in social media. Empirical measurements of the size of contagious outbreaks in social media sites (**a**) Digg and (**b**) Twitter. The *x*-axis represents the size of outbreak in terms of the number of people who dugg or retweeted specific information, and *y*-axis reports the frequency of events of that size.

2. Size of Social Contagions

Empirical studies of information spread in social media have failed to observe outbreaks as large as those predicted by the independent cascade model [14,19,20]. This review focuses on two widely-studied social platforms—Twitter and Digg—although similar behaviors were observed in a variety of other social platforms [19]. Twitter, a popular microblogging platform, allows registered users to broadcast short messages, or "tweets." These messages may contain URLs or descriptive labels, known as hashtags. In addition to composing original tweets, users can re-share, or "retweet," messages posted by others. In contrast to Twitter, Digg focuses solely on news. Digg users submit URLs to news stories they find on the web and vote for, *i.e.*, "digg," stories submitted by others. Both platforms include a social networking component: users can subscribe to the feeds of other users to see the tweets those users posted (on Twitter) or the news stories they submitted or voted for (on Digg). The follow relationship is asymmetric; hence, we refer to the subscribing users as *followers*, and the users they subscribe to as their *friends* (or followees).

To measure the size of outbreaks on Twitter, researchers used URLs to external web content embedded in tweets as unique markers of information [21]. They tracked these URLs as users shared or retweeted the messages with their followers. A similar strategy was used to track each news story on Digg. Thus, the number of times a message containing a URL was retweeted or a news story was "dugg" in their respective networks gave an estimate of the outbreak size in that network.

Figure 2 shows the distribution of outbreak sizes on the Twitter and Digg social platforms. Note that outbreaks have a long-tailed distribution, except for a bump on Digg that corresponds to promoted stories. When a newly submitted story accumulated enough votes, it was promoted to Digg's front page, where it was visible to everyone, not only to followers of voters [22]. The higher visibility of stories on the front page gave them a popularity boost, resulting in log-normally distributed popularity. However, even the most popular stories did not penetrate very far. Only one story, about Michael Jackson's death, could be said to have reached "viral" proportions, *i.e.*, reaching a non-negligible fraction of active Digg users (in this case, about 5%). The next most popular story reached fewer than 2% of Digg voters, and the vast majority of front page stories reached fewer than 0.1% of the voters. Similarly, very few of the outbreaks on Twitter reached more than 10,000 users, or less than 2% of the active user population. These findings are in line with other studies, including by Goel *et al.* [19], who analyzed seven online social networks, ranging from communication platforms to networked games, to reach the same conclusion: the vast majority of outbreaks in online social networks are small and terminate within one step of the source of information.

3. Mechanics of Contagion: Exposure Response

The observations above present a puzzle: what stops information from spreading widely on social media, and why is outbreak size so much smaller than predicted by the independent cascade model? A number of hypotheses could potentially explain the empirical findings:

Subcriticality The vast majority of information spread is sub-critical, with transmissibility below the epidemic threshold. As a result, information is unlikely to spread upon exposure, and can be considered uninteresting. This hypothesis is easy to dismiss, since it is difficult to imagine that all the information shared on many different social media platforms is uninteresting.

Load balancing Social media users may modulate transmissibility of information to prevent too many pieces of information from spreading and creating information overload. This hypothesis is difficult to evaluate, though it is not very credible, since such wide-scale coordination would be difficult to achieve. Moreover, it would require users to correctly estimate the popularity of different pieces of information in their local neighborhood, a measurement that is easily skewed in networks [23].

Novelty decay Transmissibility of information could diminish over time as information loses novelty. A study [24] explicitly addressed this hypothesis, and found that the probability to retweet

information on Twitter does not depend on its absolute age, but only the time it first appeared in a user's social feed.

Network structure Although it is conceivable that network structure (e.g., clustering or communities) could limit the spread of information, this hypothesis was ruled out [14]. As can be seen in Figure 1, the structure of the actual Digg follower graph somewhat reduces the size of outbreaks, but not nearly enough to explain empirical observations.

Contagion mechanism The decisions people make to vote for a story on Digg or retweet a URL on Twitter, once their friends have shared, could differ substantially from the ICM. These differences could prevent information from spreading [14].

To characterize the mechanisms of contagion, researchers use the "exposure response function." Since a person may be exposed to some information (or disease) by several friends, exposure response function gives the probability of an infection as a function of the number of exposures. Under the independent cascade model, infection probability rises monotonically with the number of infected friends as: $p_{\text{ICM}}(\text{infection}|k \text{ exposures}) = 1 - (1 - \mu)^k$, where μ is the transmissibility. Using social media data, researchers empirically measured the exposure response function for Twitter and Digg users. To do this, they found all users who became "infected" (e.g., retweeted a URL [24] or adopted a hashtag [25] on Twitter, or "dugg" a story on Digg [14]) after k of their friends (*i.e.*, the users he or she follows) became "infected." The exposure response function is the ratio of the number of users who became "infected" to the number who did not become "infected" for different values of k.

Figure 3 shows the exposure response functions for Digg and Twitter, averaged over all users. The shape and magnitude of the exposure response functions are fundamentally different from that of the ICM. The form of the exposure response indicates that, while initial exposures increase infection probability, additional exposures suppress new infections. According to Romero *et al.* [25], such responses are suggestive of complex contagion, another popular model for describing social contagions, where "infection" does not occur until exposure by some specified fraction of friends [26–29].

(a) Digg

(b) Twitter
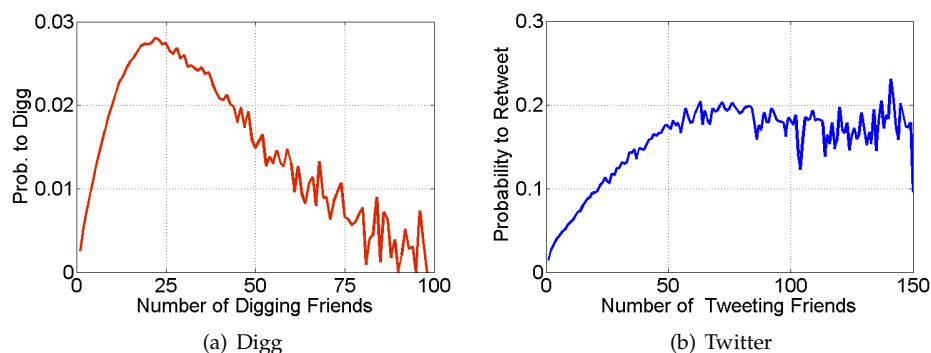
Figure 3. The exposure response function for social media users. The figures report the probability (averaged over all users) to respond to information, *i.e.*, (**a**) to digg a news story or (**b**) retweet a URL, as a function of number of friends who previously did so.

Does the suppressed response to multiple exposures inhibit the spread of information online? Ver Steeg *et al.* [14] simulated the spread of social contagions on the Digg follower graph with the suppressed exposure response function suggested by Figure 3a. In the simulation, exposure response was approximated as follows. If a node had any infected friends, it became infected with probability μ. However, if it did not get infected (with probability $1 - \mu$), it was forever immune to new infections. Figure 4 shows the size of the resulting outbreaks (black crosses) as a function of transmissibility μ. The outbreaks are an order of magnitude smaller than those predicted by the independent cascade model, and in line with empirically observed outbreaks (red dots). This suggests that online contagions fail to spread due to the reduced susceptibility of social media users to multiple exposures.
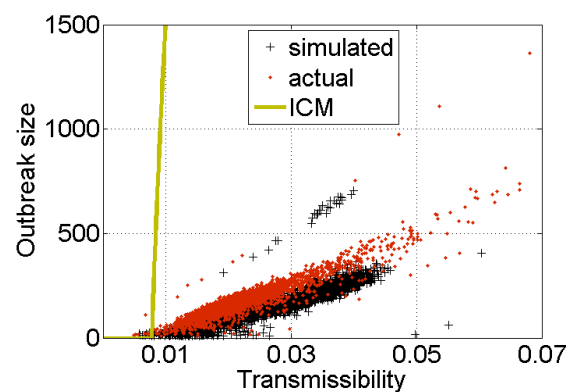
Figure 4. Size of simulated outbreaks on Digg as a function of transmissibility. Simulations of social contagion on the Digg follower graph (**black crosses**) using suppressed exposure response function suggested by Figure 3a. Actual outbreaks on Digg are shown as red dots, while theoretically predicted (**gold**) line is the same as in Figure 1. Suppressed response to repeated exposures vastly decreases the size of outbreaks as compared to prediction of the ICM (Figure 1).

4. Limited Attention and Cognitive Heuristics

Why do social media users not respond to multiple exposures to information? In contrast to viral infection, social media users must actively seek out information and decide to share it before becoming "infected." The enormous flux of information on social media often saturates human ability to process information [30]. Faced with an over-abundance of stimuli, humans evolved mechanisms to parsimoniously direct their attention to the most salient stimuli. What is salient depends on context: color, contrast, and motion help guide visual attention to important features of the environment, such as a predator. Social stimuli are also salient, as they aid coordination and help people avoid conflict. A variety of other cognitive heuristics are used to quickly (and unconsciously) focus attention on salient information [31,32]. In the context of social media, information that appears at the top of the web page or user's social feed is salient. As a result of this cognitive heuristic, known as "position bias" [33], people pay more attention to items at the top of a list than those in lower positions. Social influence bias, communicated through social signals, helps direct attention to online content that has been liked, shared or approved by many others [34,35].

Cognitive heuristics interact with how a web site displays information to users to alter the dynamics of social contagion. Twitter presents friends' messages in a user's feed as a chronologically ordered queue, with the most recently tweeted messages at the top. Similarly, Digg orders news stories submitted by friends in a reverse chronological order. Due to position bias, a user is more likely to see newest messages at the top of the feed than older messages in lower positions. Researchers conducted controlled experiments on Amazon Mechanical Turk to quantify position bias [36]. They presented study participants with a list of 100 items and asked them to recommend those they found interesting. Figure 5 shows the relative decrease in recommendations received at each list position, compared to what the items shown in those positions are expected to receive. Items in top list positions (0–5) receive three to five times as much attention as those in lower positions (40–75), purely by virtue of being in those positions.

However, in the observational data from social media, we do not know the feed position of a message at the time the user responds to it. Instead, we know that its position should be proportional to its age, *i.e.*, time since its arrival in the user's feed, when the latter is a queue ordered by time of item's arrival. Figure 6 confirms the effect. It shows the probability to digg (or retweet) an item as a function of the time since item's arrival. Though Twitter and Digg differ substantially in their functionality and user interface, they behave very similarly. The probability on both sites drops precipitously with time,

which suggests that social media users are far less likely to see—and retweet—older messages in lower feed positions than newer messages in top feed positions.
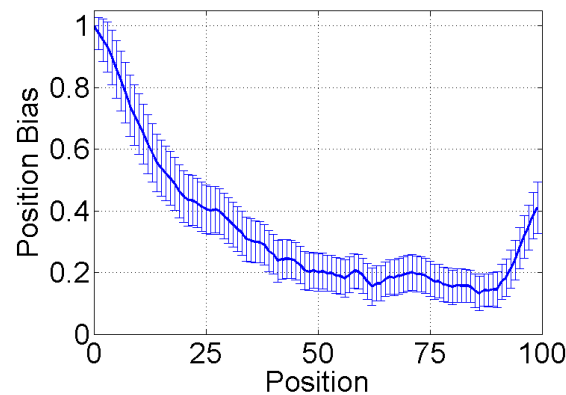
Figure 5. Position bias. The relative decrease in recommendations received by items at different position within a list, compared to expected recommendations. Items in top positions (1–5) receive four to five times as much attention as items in lower positions (50–75). The increase near the end of the list is created by people who start inspecting the list from the end.


(a) Digg

(b) Twitter
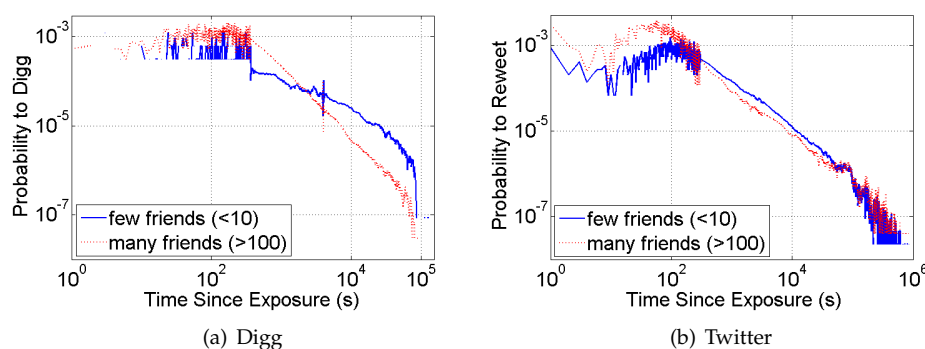
Figure 6. Time response function for social media. The probability to (**a**) digg a news story and (**b**) retweet a URL, as a function of the time since exposure, *i.e.*, message's arrival in the user's feed. To remove the confounding effects of multiple exposures, the figure only considers single exposure items. Digg stories were only followed until promotion (the first 24 hours) during which time they were only visible to the followers. The data are smoothed using progressively wider smoothing windows, as in [24].

Cognitive heuristics also interact with network structure, altering the dynamics of social contagion. The visibility of an item in a feed of a well-connected user (who follows many others) decreases faster in time than the visibility of an item in the feed of a poorly-connected user (who follows few friends). As a result, well-connected users with more friends rarely retweet old content. This is because these users receive many newer messages from their multitude of friends, which quickly push a given item further down the queue, where it is less likely to be seen. In contrast, poorly-connected users receive few messages, so that the visibility of an item does not decay as quickly. This effect is evident in Figure 6, where the probability to retweet (or digg) an item decreases faster for well-connected users (with more than 250 friends) than for the poorly-connected users (with fewer than 10 friends).

The quickly decaying visibility of information in the feeds of well-connected social media users reduces their susceptibility to becoming "infected" by that information. Figure 7 shows the exposure response function for two classes of Twitter users: those with few friends and those with many friends.

The response probability of well-connected users with many friends is much lower compared to poorly-connected users, consistent with the argument that they have a harder time finding specific messages in their long feeds. Note also that both response functions increase monotonically, in line with simple contagion models, such as the ICM. The non-monotonic behavior observed for adoption of hashtags on Twitter [25] and in Figure 3 does not represent complex contaction, but is simply an artifact of averaging over heterogeneous populations of users with different cognitive load, *i.e.*, different volumes of information in their feeds. Averaging the curves in Figure 7, we will observe an exposure response that initially increases, since both classes of users contribute; however, as the number of "infected" friends increases further, only users with many friends contribute to the response, bringing the average response function down. This is an illustration of "heterogeneity's ruses" [37]: averaging over heterogeneous populations, each with its own behavior, can produce nonsensical behavioral patterns. When studying social systems, one needs to isolate the more homogeneous populations and carry out analysis within each population [38].
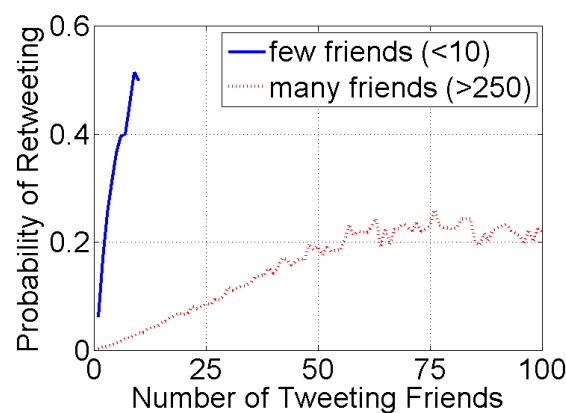
Figure 7. Exposure response function for Twitter. The figure shows the average probability to retweet a message as a function of the number of friends who previously tweeted it for two classes of users, separated according to the number of friends they follow. The well-connected users with many friends are less likely to retweet; hence, they are less susceptible than users with few friends.

5. Predicting Social Contagions

Knowing how cognitive heuristics constrain user behavior enables us to more accurately predict social contagions. To become "infected," a user must first discover at least one message containing the information. We approximate a message's visibility using the time response function, of the kind shown in Figure 6, that gives the probability that a user with n_f friends retweets or votes at a time Δt after the exposure [24].

To understand the dynamics of social contagion, we must also specify how users respond to multiple exposures to information. Here, the details of how the web site presents information matter. Twitter puts each newly retweeted item—the new exposure—at a top of the followers' feeds, creating a new opportunity for the followers to discover the item. Thus, if *k* friends tweet some information, it will appear in a user's feed *k* times in different positions. In contrast, Digg does not change the news story's relative position after a friend's digg, but increments the number of recommendations shown next to the story: after *k* friends digg a story, it still appears only once in the user's feed, but with the number *k* next to it. This number serves as a social signal that changes a user's response. The effect of social signals on user's likelihood to become "infected" can be measured experimentally [35] or estimated from observational data [39].

Putting these factors together, [39] proposed a simple model of social contagion where each exposure can independently cause an "infection" (*i.e.*, a retweet). In contrast to the plain ICM, "infection" probability depends on the visibility of the exposures, which is related to the time of the

exposures on Twitter or the time of the first exposure on Digg. A social signal, if present, will amplify "infection" probability. To validate this model, the authors used it to forecast "infections" and compared them to observed "infections." Specifically, they calculated on a minute-by-minute basis the observed frequency that a user with some number of exposures in their feed retweeted some specific information on Twitter or dugg a story on Digg in the subsequent 30 seconds. Then, they calculated the theoretical probability that the same user would act in those 30 seconds, given the same exposures. Figure 8 shows the observed *vs.* predicted probability of those infections, for different numbers of exposures in the users' feeds. For reference, perfect forecasts lie along the $y = x$ line. The unbiased fidelity of the proposed model suggests that once visibility of the exposures is taken into account, social contagion operates as a simple contagion, *i.e.*, with infection probability increasing monotonically with the number of exposures. Other works incorporated visibility into models of user behavior that account for user interests [40] and sentiment [41] about topics, their limited attention [42], and the multiple channels for finding information, such as on Digg's front page [22].

(a) Digg (b) Twitter
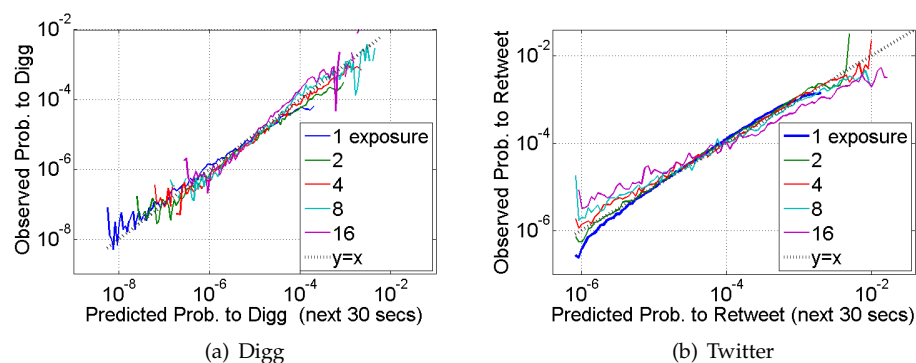
Figure 8. Predicting response. Observed probability to (**a**) digg or (**b**) retweet an item as a function of predicted probability, for different numbers of exposures to the item.

6. Discussion

The notion that networks amplify the flow of information has ignited the imagination of researchers and public alike. The few success stories—songs and videos that have spread in a chain reaction from person to person to reach millions—keep marketers searching for formulas for creating viral campaigns. Success, however, is rare. Empirical studies of the spread of information in online social networks revealed that information rarely spreads beyond the source. The search for answers as to why information fails to spread in social media has uncovered the vital role of brain's cognitive limits in social media interactions.

These cognitive limits are what differentiates the spread of information from the spread of a virus, and they must be accounted for in models of information diffusion. Specifically, in order to spread some information on social media, a person first has to discover it in his or her social feed. Discovery depends sensitively on how the web site arranges the feed, the flux of incoming information, and the effort the person is willing and able to expend on the discovery process. Moreover, as people add more friends, the volume of information they receive may grow superlinearly due to the friendship paradox in social networks [43] and its generalizations [44,45]: a person's friends are more active and post more messages on average then the person himself or herself does. As a result, the volume of information may inevitably exceed an individual's cognitive capacity, creating conditions for information overload [30]. To deal with information overload, people rely on cognitive heuristics to focus only on salient information. In the context of social media, this means paying attention to the most recent messages at the top of their feed, and ignoring the rest. This reduces the probability that highly connected people will see and spread any given piece of information in their feed, making them less susceptible to becoming "infected." The reduced susceptibility of central users suppresses the

spread of social contagions in social media. Accounting for these phenomena in models of information diffusion allows us to more accurately predict how far information will spread online.

The interplay between networks and human cognitive limits may have other non-trivial consequences. Potentially, people who have higher capacity to process information may put themselves in network positions allowing them greater access to information [40,46], which they may then leverage for personal gain [47,48]. Understanding the role of social networks and cognitive heuristics and biases in individual and collective behavior remains an open research area.

Conflicts of Interest: The author declares no conflict of interest.

References

1. Kempe, D.; Kleinberg, J.; Tardos, E. Maximizing the spread of influence through a social network. In Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '03), Washington, DC, USA, 24–27 August 2003; ACM Press: New York, NY, USA, 2003; pp. 137–146.

2. Leskovec, J.; Adamic, L.A.; Huberman, B.A. The dynamics of viral marketing. In Proceedings of the 7th ACM Conference on Electronic Commerce, EC '06, Ann Arbor, MI, USA, 11–15 June 2006; ACM: New York, NY, USA, 2006; pp. 228–237.

3. Leskovec, J.; Krause, A.; Guestrin, C.; Faloutsos, C.; VanBriesen, J.; Glance, N. Cost-effective Outbreak Detection in Networks. In Proceedings of the 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '07, San Diego, CA, USA, 21–24 August 2011; ACM: New York, NY, USA, 2007; pp. 420–429.

4. Gruhl, D.; Guha, R.; Nowell, D.L.; Tomkins, A. Information diffusion through blogspace. In Proceedings of the 13th International Conference on World Wide Web, WWW '04, New York, NY, USA, 17–22 May 2004; ACM: New York, NY, USA, 2004; pp. 491–501.

5. Castellano, C.; Fortunato, S.; Loreto, V. Statistical physics of social dynamics. *Rev. Mod. Phys.* **2009**, *81*, 591–646.

6. Newman, M.E.J. Spread of epidemic disease on networks. *Phys. Rev. E* **2002**, *66*, doi:10.1103/PhysRevE.66.016128.

7. Anagnostopoulos, A.; Kumar, R.; Mahdian, M. Influence and correlation in social networks. In Proceeding of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '08, Las Vegas, NV, USA, 24–27 August 2008; ACM: New York, NY, USA, 2008; pp. 7–15.

8. Gomez Rodriguez, M.; Leskovec, J.; Krause, A. Inferring Networks of Diffusion and Influence. In Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '10, Washington, DC, USA, 25–28 July 2010; ACM: New York, NY, USA, 2010; pp. 1019–1028.

9. Rodriguez, M.G.; Leskovec, J.; Balduzzi, D.; Schölkopf, B. Uncovering the structure and temporal dynamics of information propagation. *Netw. Sci.* **2014**, *2*, 26–65.

10. Myers, S.; Zhu, C.; Leskovec, J. Information Diffusion and External Influence in Networks. In Proceedings of the KDD, Beijing, China, 12–16 August 2012.

11. Hethcote, H.W. The Mathematics of Infectious Diseases. *SIAM Rev.* **2000**, *42*, 599–653.

12. Goldenberg, J.; Libai, B.; Muller, E. Talk of the Network: A Complex Systems Look at the Underlying Process of Word-of-Mouth. *Mark. Lett.* **2001**, *12*, 211–223.

13. Gruhl, D.; Liben-Nowell, D.; Guha, R.; Tomkins, A. Information diffusion through blogspace. *SIGKDD Explor. Newsl.* **2004**, *6*, 43–52.

14. Steeg, G.V.; Ghosh, R.; Lerman, K. What stops social epidemics? In Proceedings of the 5th International Conference on Weblogs and Social Media, Catalonia, Spain, 17–21 July 2011.

15. Moreno, Y.; Pastor-Satorras, R.; Vespignani, A. Epidemic outbreaks in complex heterogeneous networks. *Eur. Phys. J. B* **2002**, *26*, 521–529.

16.   Wang, Y.; Chakrabarti, D.; Wang, C.; Faloutsos, C. Epidemic Spreading in Real Networks: An Eigenvalue Viewpoint. In Proceedings of the 22nd International Symposium on Reliable Distributed Systems, Florence, Italy, 6–18 October 2003; pp. 25–34.

17.   Chakrabarti, D.; Wang, Y.; Wang, C.; Leskovec, J.; Faloutsos, C. Epidemic Thresholds in Real Networks. *ACM Trans. Inf. Syst. Secur.* **2008**, *10*, 13.

18.   Ghosh, R.; Lerman, K. A Parameterized Centrality Metric for Network Analysis. *Phys. Rev.* **2011**, *E 83*, doi:10.1103/PhysRevE.83.066118.

19.   Goel, S.; Watts, D.J.; Goldstein, D.G. The Structure of Online Diffusion Networks. In Proceedings of the 13th ACM Conference on Electronic Commerce, Valencia, Spain, 4–8 June 2012.

20.   Bakshy, E.; Hofman, J.M.; Mason, W.A.; Watts, D.J. Everyone's an influencer: Quantifying influence on twitter. In Proceedings of the Fourth ACM International Conference on Web Search and Data Mining, Hong Kong, China, 9–12 February 2011.

21.   Lerman, K.; Ghosh, R. Information Contagion: An Empirical Study of Spread of News on Digg and Twitter Social Networks. In Proceedings of the 4th International Conference on Weblogs and Social Media (ICWSM), Washington, DC, USA, 23–26 May 2010.

22.   Hogg, T.; Lerman, K. Social Dynamics of Digg. *EPJ Data Sci.* **2012**, *1*, doi:10.1140/epjds5.

23.   Lerman, K.; Yan, X.; Wu, X.Z. The "Majority Illusion" in Social Networks. *PLoS ONE* **2016**, *11*, 1–13.

24.   Hodas, N.O.; Lerman, K. How Limited Visibility and Divided Attention Constrain Social Contagion. In Proceedings of the ASE/IEEE International Conference on Social Computing, Amsterdam, The Netherlands, 11 May 2012.

25.   Romero, D.M.; Meeder, B.; Kleinberg, J. Differences in the Mechanics of Information Diffusion Across Topics: Idioms, Political Hashtags, and Complex Contagion on Twitter. In Proceedings of the 20th International Conference on World Wide Web, Hyderabad, India, 28 March–1 April 2011.

26.   Granovetter, M. Threshold Models of Collective Behavior. *Am. J. Sociol.* **1978**, *83*, 1420–1443.

27.   Watts, D.J. A simple model of global cascades on random networks. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 5766–5771.

28.   Centola, D.; Macy, M. Complex contagions and the weakness of long ties1. *Am. J. Sociol.* **2007**, *113*, 702–734.

29.   Centola, D. The spread of behavior in an online social network experiment. *Science* **2010**, *329*, 1194–1197.

30.   Rodriguez, M.G.; Gummadi, K.; Schoelkopf, B. Quantifying Information Overload in Social Media and its Impact on Social Contagions. In Proceedings of the Eighth International AAAI Conference on Weblogs and Social Media, Ann Arbor, MI, USA, 1–4 June 2014.

31.   Kahneman, D. *Attention and Effort*; Prentice Hall: Englewood Cliffs, NJ, USA, 1973.

32.   Kahneman, D. *Thinking, Fast and Slow*; Farrar, Straus and Giroux: New York, NY, USA, 2011.

33.   Payne, S.L. *The Art of Asking Questions*; Princeton University Press: Princeton, NJ, USA, 1951.

34.   Salganik, M.J.; Dodds, P.S.; Watts, D.J. Experimental Study of Inequality and Unpredictability in an Artificial Cultural Market. *Science* **2006**, *311*, 854–856.

35.   Hogg, T.; Lerman, K. Disentangling the Effects of Social Signals. *Hum. Comput. J.* **2015**, *2*, 189–208.

36.   Lerman, K.; Hogg, T. Leveraging position bias to improve peer recommendation. *PLoS ONE* **2014**, *9*, doi:10.1371/journal.pone.0098914.

37.   Vaupel, J.W.; Yashin, A.I. Heterogeneity's ruses: Some surprising effects of selection on population dynamics. *Am. Stat.* **1985**, *39*, 176–185.

38.   Kooti, F.; Lerman, K.; Aiello, L.M.; Grbovic, M.; Djuric, N.; Radosavljevic, V. Portrait of an Online Shopper: Understanding and Predicting Consumer Behavior. In Proceedings of the 9th ACM International Conference on Web Search and Data Mining, San Francisco, CA, USA, 22–25 February 2016.

39.   Hodas, N.O.; Lerman, K. The Simple Rules of Social Contagion. *Sci. Rep.* **2014**, *4*, doi:10.1038/srep04343.

40.   Kang, J.; Lerman, K. User Effort and Network Structure Mediate Access to Information in Networks. In Proceedings of the 9th International AAAI Conference on Weblogs and Social Media (ICWSM), Oxford, UK, 26–29 May 2015.

41.   Hogg, T.; Lerman, K.; Smith, L.M. Stochastic Models Predict User Behavior in Social Media. *ASE Hum.* **2013**, *2*. Available online: http://arxiv.org/abs/1308.2705 (accessed on 8 May 2016).

42.   Weng, L.; Flammini, A.; Vespignani, A.; Menczer, F. Competition among memes in a world with limited attention. *Sci. Rep.* **2012**, *2*, doi:10.1038/srep00335.

43.   Feld, S.L. Why Your Friends Have More Friends than You Do. *Am. J. Sociol.* **1991**, *96*, 1464–1477.

44. Hodas, N.O.; Kooti, F.; Lerman, K. Friendship Paradox Redux: Your Friends Are More Interesting Than You. In Proceedings of the 7th International AAAI Conference On Weblogs And Social Media (ICWSM), Cambridge, MA, USA, 8–11 July 2013.

45. Kooti, F.; Hodas, N.O.; Lerman, K. Network Weirdness: Exploring the Origins of Network Paradoxes. In Proceedings of the International Conference on Weblogs and Social Media (ICWSM), Ann Arbor, MI, USA, 1–4 June 2014.

46. Aral, S.; Van Alstyne, M.W. The Diversity-Bandwidth Tradeoff. *Am. J. Sociol.* **2011**, *117*, 90–171.

47. Granovetter, M.S. The Strength of Weak Ties. *Am. J. Sociol.* **1973**, *78*, 1360–1380.

48. Burt, R. *Structural Holes: The Social Structure of Competition*; Harvard University Press: Cambridge, MA, USA, 1995.