Evidence of complex contagion of information in social media: An experiment using Twitter bots

Bjarke Mønsted[1]☯, Piotr Sapieżyński[1]☯, Emilio Ferrara[2,3]☯, Sune Lehmann[1]☯*

1 Technical University of Denmark, Applied Mathematics and Computer Science, 2800 Lyngby, Denmark, **2** Information Sciences Institute, University of Southern California, Marina del Rey, CA 90292, United States of America, **3** Indiana University, Network Science Institute, Bloomington, IN, United States of America

☯ These authors contributed equally to this work.
* sljo@dtu.dk

Abstract

It has recently become possible to study the dynamics of information diffusion in techno-social systems at scale, due to the emergence of online platforms, such as Twitter, with millions of users. One question that systematically recurs is whether information spreads according to simple or complex dynamics: does each exposure to a piece of information have an independent probability of a user adopting it (simple contagion), or does this probability depend instead on the number of sources of exposure, increasing above some threshold (complex contagion)? Most studies to date are observational and, therefore, unable to disentangle the effects of confounding factors such as social reinforcement, homophily, limited attention, or network community structure. Here we describe a novel controlled experiment that we performed on Twitter using 'social bots' deployed to carry out coordinated attempts at spreading information. We propose two Bayesian statistical models describing simple and complex contagion dynamics, and test the competing hypotheses. We provide experimental evidence that the complex contagion model describes the observed information diffusion behavior more accurately than simple contagion. Future applications of our results include more effective defenses against malicious propaganda campaigns on social media, improved marketing and advertisement strategies, and design of effective network intervention techniques.

Introduction

The diffusion of information and ideas in complex social systems has fascinated the research community for decades [1]. The first proposal to use epidemiological models for the analysis of the spreading of ideas was put forth more than fifty years ago [2]. Such models, where each exposure results in the same adoption probability, are referred to as *simple contagion* models.

It was subsequently suggested, however, that more complex effects might come into play when considering the spread of ideas rather than diseases. For example, some people tend to stop sharing information they consider "old news", while others refuse to engage in

discussions or sharing certain opinions they do not agree with [3–5]. Such models, in which adoption probabilities instead depend strongly on the number of adopters in a person's social vicinity in a way where exposure attempts cannot be viewed as independent, are referred to as *complex contagion* [6] models. Concretely, we use a threshold complex contagion model, in which the adoption probability is assumed to increase slowly for low number of unique exposure sources, then increase relatively quickly when the number of sources approaches some threshold level (see 'Models' for full details).

The role of contagion in the spreading of information and behaviors in (techno-)social networks is now widely studied in computational social science [7–19], with applications ranging from public health [20] to national security [21]. The vast majority of these studies are, however, either observational, and therefore prone to biases introduced by confounding factors (network effects, cognitive limits, etc.), or entail controlled experiments conducted only on small populations of a few dozens individuals [6, 7]. To date, these limitations have prevented the research community from drawing a conclusive answer as to the role of simple and complex information contagion dynamics at scale.

In this paper we shed new light on the nature of information diffusion using a large-scale experiment on Twitter, in which we study the spreading of hashtags within a controlled environment. Creating a controlled environment for experiments within online platforms is especially challenging for researchers that do not have access to the system's design itself, as traditional techniques such as A/B testing cannot be employed. Even for service providers like Facebook, ethical concerns emerged when random control trials were carried out without review board approval [15].

For this experiment, we leveraged algorithm-driven Twitter accounts (social bots) [22]. We had previously shown that a coordinated network of Twitter bots can be effective in influencing trending topics on Twitter [23]. This study is a follow-up experiment designed to quantitatively investigate how users react to information stimuli presented by single or multiple sources. In particular, for this experiment, teams of students from the Technical University of Denmark (DTU) worked together to create a network of Twitter bots (a botnet) designed to attract a large number of human followers. We programmed the bots to spread Twitter hashtags (see Table 1) in a synchronized manner among a set of real Twitter users from a selected geographical area. A large number of users in our target dataset followed one or multiple bots (See Fig 1B), which allowed us to study the effect of multiple exposures from distinct sources on information contagion.

The decision to use Twitter bots to perform coordinated interventions has several advantages: first, we are able to ensure that the hashtags we introduce are new to Twitter, and therefore that they are seen by the target users for the first time when we perform experiments.

Table 1. List of interventions.

Hashtag	Message
#getyourflushot	Encouraging Twitter users to vaccinate.
#highfiveastranger	Encouraging users to engage in positive human interactions.
#somethinggood	Sharing a recent positive experience.
#HowManyPushups	Encouraging healthy behaviors and fitness.
#turkeyface	Photoshopping a celebrity's face onto a turkey.
#SFThanks	Hashtag for Thanksgiving in San Fransisco.
#blackfridaystories	Sharing Black Friday shopping stories.
#BanksySF	Rumor that Banksy, the street artist, was in San Fransisco.

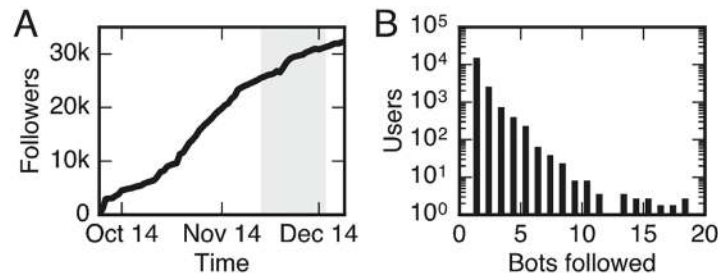https://doi.org/10.1371/journal.pone.0184148.t001

Fig 1. Illustration of the status of our botnet at the time of the interventions. The bots had accumulated a large number (∼ 25000) of followers (A) at the time of the interventions (shaded region), and many of the target users followed several distinct bots (B).

Second, it enables the bots to work together to expose users to each intervention multiple times. Finally, the Twitter botnet mitigates the confounding effects of homophily [24–26]. For example, when conducting a purely observational study, it is a fundamental problem to distinguish whether a user is more likely to adopt information shared by many of their friends because they are influenced by their friends sharing the content, or simply because friends tend to be similar, so anything tweeted shared by the user's friends is more likely to be of interest to the user.

In the remainder of the paper we will discuss the experimental framework design in detail, then present two statistical models for simple and complex contagion, developed in order to evaluate the two competing hypotheses, and finally show the results of the experimental evaluation.

Results

Deploying the botnet. Creating a botnet with a large number of followers with a network structure suited for testing our hypotheses presented several challenges which are described below.

We began by ensuring that the bots would appear to be human-like if subjected to a cursory inspection. We achieved this goal by having the bots generate content using simple natural language processing rules as well as 'recycling' popular content from other Twitter users. We also had the bots tweet at irregular intervals, but with frequencies set according to a circadian pattern. Finally, we used some Twitter users' tendency to reciprocate friendships to ensure that the bots were followed by a large number of accounts while themselves following only a few; a following/follower ratio much smaller than one is unusual in typical twitter bots. The full botnet consisted of 39 algorithmically driven Twitter accounts. See 'Materials and Methods' for full details on botnet-creation.

Once we had established the botnet, we focused on establishing a network structure that would allow for investigating the mechanism driving contagion processes. Our strategy was simple: Whenever a user followed one of our bots, the ID of this user was automatically communicated to the remaining bots, which then also attempted to get that user to follow them. This strategy resulted in a botnet followed by a large number of human users (around 25 000 total followers a the time of interventions), in which a large users followed multiple bots, allowing us to test the effects of multiple exposures to information. Fig 1A shows the total number of followers as a function of time, while Fig 1B displays the distribution of users following *n* bots. Having obtained a botnet with a large number of followers and a desirable network structure, the bots performed a series of coordinated interventions, described in the following.

Fig 2. Bots in a botnet can work together to provide users with multiple exposures to an intervention.
(A) user U only follows bot B_1. Bot (B_1) acts as a proxy and exposes the user not only to its own content, but also to content from two other bots (B_2 and B_3), that the user does not follow. (B) Twitter feed from the perspective of user U.

The general intervention strategy implemented by the bots follows:

1. Each bot tweets 2 original tweets about a given #hashtag;

2. Each bot retweets the first 4 tweets about that #hashtag;

3. Each bot retweets 15 tweets containing that #hashtag that do not originate from other bots;

4. Each bot favorites all tweets about the given #hashtag.

Step 1 of this protocol was based on human-generated tweets; this allowed students to create content designed to increased the likelihood of adoption. In order to avoid the confounding effect of users adopting hashtags they encountered from other sources than our botnet, we based the study on hashtags (see Table 1) that had not previously been observed on Twitter. Steps 2–4 were instead automated. By retweeting each other's content, the bots provided a higher exposure to the target users with respect to what would have been possible if bots could only have targeted their mutual friends, as illustrated in Fig 2. An overview of the hashtags that we introduced is shown in Table 1. The hashtags we introduced support positive behaviors (e.g., encouraging vaccinations or positive human interactions, sharing something good, etc.) and in some cases are contextualized with the time period of the intervention (e.g., fostering stories about Thanksgiving and Black Friday).

To track exposures and contagions, each bot automatically recorded when a target user retweeted intervention-related content, and also each exposure that had taken place prior to the retweeting. It is important to remark that users cannot be expected to consume the entirety of content generated by those they follow: the probability of seeing a tweet can depend on many factors including the total number of accounts a user follows, the activity level of each of those accounts, and the amount of time that user spends on Twitter. The two contagion models we created, described in the following, model this uncertainty explicitly.

Models

In the following, we propose two contagion models, namely *simple contagion* model (SC), in which all exposure attempts are considered to be independent, and *complex contagion* threshold model (CC), and derive quantitative predictions for them. Both models take into account the uncertainty regarding the target users observing a given tweet. Specifically, we do not have direct access to a user's *actual exposures*, to an intervention, but only to the *attempted*

exposures, N, meaning simply the number of times a bot followed by the user published a tweet containing the hashtag in question. The simple contagion model employs only the total number of attempted exposures, which we denote k. The complex contagion model, however, is only concerned with the number of unique sources κ from which one or more exposures have succeeded. This is because a central idea in threshold models is the reluctance to partake in activities until a number of individuals in one's social group have already done so [3]. To avoid cluttered notation, we write k and 'number of exposures' in descriptions relevant for both models, although these should be replaced with κ and 'number of unique exposure sources' for the CC case.

Going forward, we separate the two factors that enter into a users tendency to adopt a behavior. Firstly, the probability of the user experiencing k exposures, and, secondly, the probability $P(\mathrm{RT}|k)$ of the user deciding to retweet content after experiencing k exposures. Thus, we model the probability of a user retweeting content from an intervention, given bot activity A as

$$P(\mathrm{RT}|A) = \sum_k P(k|A)P(\mathrm{RT}|k).$$

(1)

In the following, $A = [a_1, a_2, \ldots]$ denotes a list of the number of times a user has received attempted exposures from each bot (disregarding those with zero attempts). For example, $A = [2, 1, 3, 1]$ means that a given user has been subjected to 2 attempted exposure from one bot, 3 from another bot, and 1 from two bots. In the case of SC, where only the total number of exposures is of interest, we will use $N = \sum_i a_i$ to denote the total number of attempted exposures.

Simple contagion. We model the number of exposures by assuming that a user sees a given tweet with some independent probability q. Thus, the number of actual exposures follows a binomial distribution $B(k; N, q)$ given by q and the number of attempted exposures N,

$$P(k|N, q) = \binom{N}{k}q^k(1-q)^{N-k} = \frac{N!}{k!(N-k)!}q^k(1-q)^{N-k}.$$

(2)

In SC, each actual exposure has some probability ρ of 'infecting' the user, which is independent of other exposures. Hence the probability for an infection after k exposures is simply

$$P(\mathrm{RT}|k)_{\mathrm{SC}} = 1 - (1-\rho)^k,$$

(3)

which is almost linear in k for small values of ρ. Inserting this expression into Eq (1) we get

$$P(\mathrm{RT}|N)_{\mathrm{SC}} = \sum_{k=0}^{N} \frac{N!}{k!(N-k)!}q^k(1-q)^{N-k}(1-(1-\rho)^k),$$

(4)

which is equivalent to the simpler expression

$$P(\mathrm{RT}|N)_{\mathrm{SC}} = 1 - (1-\rho q)^N.$$

(5)

Under *results*, we fit the parameters in Eq (5) to the data obtained by our experiment. Next, we derive an expression for the retweet probability of the complex contagion model.

Complex contagion. When quantifying the predictions of CC, we face two obstacles: *(i)* redefining the conditional retweet probability $P(\mathrm{RT}|A)$ in order to incorporate the threshold effect of CC; and, *(ii)* obtaining an expression of the probability distribution for κ given the previous activity A.

Let us first derive the probability distribution for κ given previous activity A. The probability p_i of source i resulting in one or more actual exposures is given by a binomial distribution

using similar considerations as those leading to Eq (5):

$$p_i = 1 - (1 - q)^{a_i}. \tag{6}$$

Hence, the distribution of unique exposures is the result of independent draws from $|A|$ Bernoulli trials with a_i draws from each, with individual success probabilities p_i, also known as Poisson's Binomial [27]. For example, the probability of $\kappa = 1$ given a list of attempted exposures A is obtained by summing over the different ways we may achieve success in only a single Bernoulli trials:

$$P(\kappa = 1 | A) = \sum_{j=1}^{|A|} \overbrace{(1 - (1-q)^{a_j})}^{p_j} \prod_{i \neq j} \overbrace{(1-q)^{a_i}}^{1-p_i}. \tag{7}$$

Generalizing this to any $\kappa \leq |A|$, we sum over every unique combination of κ successful trials. Denoting the set of sets of κ integers between 1 and $|A|$ by S_κ, we get

$$P(\kappa | A) = \sum_{s \in S_\kappa} \prod_{i \in s} (1 - (1-q)^{a_i}) \prod_{i \notin s} (1-q)^{a_i}, \tag{8}$$

$$S_\kappa = \{s \subseteq \{1, 2, \ldots, |A|\}, |s| = \kappa\}. \tag{9}$$

We include a note in the SI on how to efficiently compute Eq (8), as this expression becomes infeasible to compute using a brute force approach when $|A| > 25$. As in Eq (1), we sum over positive κ to obtain a final expression for the retweet probability given a list A of exposure attempts, by computing the sum $\sum_\kappa P(\kappa|A)P(\text{RT}|\kappa)$ over the probabilities given by Eq (8).

$$P(\text{RT}|A) = \sum_{\kappa=1}^{|A|} \sum_{s \in S_\kappa} \prod_{i \in s} (1 - (1-q)^{a_i}) \prod_{i \notin s} (1-q)^{a_i} P(\text{RT}|\kappa). \tag{10}$$

Now we select a threshold function for $P(RT|\kappa)$. We choose a Sigmoid function,

$$P(\text{RT}|\kappa)_{\text{CC}} = \rho_l + \frac{\rho_h - \rho_l}{1 + e^{-w(\kappa - \kappa_0)}}, \tag{11}$$

as it employs both a threshold κ_0, steepness w and the lower and upper limits, ρ_l and ρ_h. Sigmoids are commonly used to model soft thresholds, for example as activation functions in neural networks [28], or as fuzzy membership functions [29]. Combining Eqs (10) and (11), the expression for $P(\text{RT}|A)$ becomes

$$P(\text{RT}|A)_{\text{CC}} = \sum_{\kappa=1}^{|A|} \sum_{s \in S_\kappa} \prod_{i \in s} (1 - (1-q)^{a_i}) \prod_{i \notin s} (1-q)^{a_i} \left(\rho_l + \frac{\rho_h - \rho_l}{1 + e^{-w(\kappa - \kappa_0)}} \right). \tag{12}$$

Having derived expressions for the retweet probability for a user given previous exposure activity for both the SC and CC hypotheses, we proceed to fit the models to our experimental data.

Analysis. We now use these two contagion models to investigate how the adoption probability $P(\text{RT})$ varies as a function of the exposure numbers in our dataset. By studying how well each model fits the observed data, we can determine which model is the most appropriate description of the contagion processes measured in the experiment.

An example of the distributions for $q = 0.2$ and the best fits of the SC and CC models are shown in Fig 3.
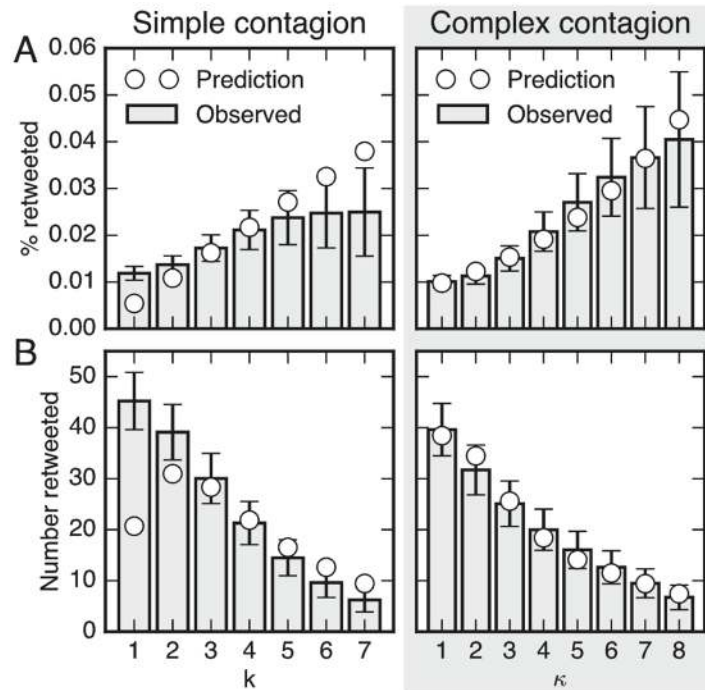
Fig 3. Simple contagion (SC, left) does not adequately describe the contagion dynamics: The best fit underestimates the probability of retweeting after a low number of exposures and overestimates the probability with a large number of exposures. The best fit of complex contagion (CC, right) dynamics correctly estimates the probability of retweeting across the number of sources of exposure. A. Percentage of tweets that were retweeted after k successful exposures (SC) or after exposures from κ sources (CC). B. Number of tweets retweeted following k successful exposures (SC), or after exposures from κ sources (CC). Best fit of SC model (Eq (4)) and CC model (Eq (10)) to the data using $q = 0.20$, plotted up to $k = 7$ (and $\kappa = 8$) to avoid plotting noisy data for large values of k (and κ).

Fig 3 suggests that the SC model from Eq (4) is not an adequate fit to the observed data, whereas the CC model from Eq (10) provides an excellent fit. The figure indicates that the CC model, which models contagion as a function of the number of distinct sources provides a better explanation for the user behavior on Twitter.

In order to compare the models in a way that takes into account different model complexities, we use the Bayesian Information Criterion (BIC) score [30] on simulations using the probabilities provided by the two models (see Methods for details). The results, displayed in Fig 4, show that the CC model results in better BIC scores for any value of q. In general, a difference in BIC scores larger than 10 points is considered a very strong evidence in support of the model with the lower score [31]. Fig 4 shows gaps between the average BIC scores of the two models that are substantially larger than 10 points throughout the entire range of values of q, supporting the hypothesis that the CC model is the best explanation for the dynamics of information diffusion on Twitter.

For very small values of q ($q < 0.1$) the gap between the BIC scores of the two models is small, and as q grows the gap increases to reach its maximum for values of q around 0.5. The reason for the low BIC scores in the case of very low values of q is that the estimates of exposure numbers from Eqs (2) (SC) and (8) depend on q and yield a very low number of estimated successful exposures for low values of q, which causes the error bars on the number of estimated retweets (such as those from Fig 3B) to grow large.
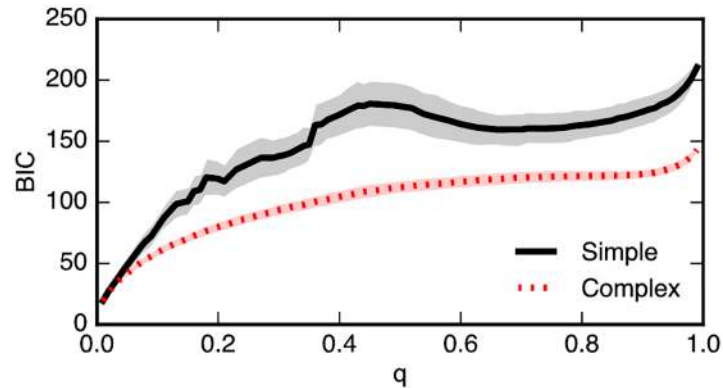
Fig 4. BIC scores for both SC and CC models for a range of values of q, the lower the score the better. Across the values of the q parameter, complex contagion model achieves lower BIC scores than simple contagion. The thick lines are the mean values of the simulations, and the shaded regions are the percentiles corresponding to one standard deviation, i.e. they contain 68% of the simulation results.

Methods

Data. All data was collected in accordance with the Danish regulations for personal data; additionally the study has been subject to Institutional Review Board (IRB) approval. The IRB grantee is Indiana University (protocol number 1410501891), which was the hosting institution of the only U.S.-based author (Emilio Ferrara) at the time when this experiment was performed. All co-authors aligned to the requirements imposed by Indiana University's approved protocol. For data access please contact Copenhagen Center for Social Data Science (http://sodas.ku.dk/contact/), or the corresponding author: Sune Lehmann.

Botnet creation. We designed the Twitter bots as part of a graduate course on social networks. The goal was to create bots which appear, at a cursory glance, to be human-operated Twitter accounts, but in reality are algorithmically driven (by means of Python scripts). The bot creation was divided into two phases: first, the goal was to build convincing accounts that real users might want to follow. Second, we worked to infiltrate a set of geographically co-located real users and spread new hashtags among them.

In phase 1, each group of 2-4 students manually created 1-3 personas (with interests, music taste, favorite sports team, etc.) and corresponding Twitter profiles, each with a profile picture, profile description, background picture, etc., resulting in a total of 39 bots. Each group also manually posted a number of initial tweets for each bot.

One of the key objectives was to achieve a large follower base while maintaining a low following/follower ratio. A low following/follower ratio is unusual among bots [21] and signals popularity on Twitter. Our bots achieved a low ratio by capitalizing on the fact that many new users with relatively few followers (and other Twitter bots) tend to reciprocate the link when they gain a new follower. Therefore, we used the following strategy: Every day, each bot automatically followed approximately 100-200 randomly selected accounts with a low follower count or the string 'followback' in the description. After 24 hours, the bots unfollowed the accounts that failed to reciprocate their follow. This routine was repeated every subsequent day. Using this strategy, the bots were able to maintain a following/follower ratio close to 1, while gaining large amounts of followers. The bots avoided automatic detection by limiting the churn among their followers, since performing too many (un)follow operations in a day leads

to a suspension of the account. As a whole, the botnet was successful in gaining a large group of followers which grew steadily throughout the duration of the experiment, as shown in Fig 1A.

While attracting followers, the bots gradually assumed a number of behaviors designed to emulate human behavior:

Geographical patterns. All bots' self-reported location in their Twitter profile was set to the San Francisco Bay Area. In addition, all bots tweeted with geo-tagged tweets, set to originate from a random location within the Bay Area bounding box. This allowed our bots to target a geographically-confined region.

Temporal patterns. Bots also timed their tweets to match typical diurnal patterns corresponding to the pacific time zone, and produce content that reflected circadian patterns of activity commonly observed online [32].

Content. Finally, based on simple natural language processing rules, the bots automated tweeting and re-tweeting of content that matched the persona developed above.

As final step of phase 1, the bots unfollowed users which were obviously spam/bot accounts in order to decrease their following/followed ratio. To investigate the quality of each bot, we routinely used the online service *Bot or Not* API [33] (http://truthy.indiana.edu/botornot/) to ensure that the bots appeared human to state-of-the-art bot-detection-software.

In phase 2, the bots began following non-bot Twitter accounts within the target area (San Francisco/Bay Area), leveraging the information users self-reported in their Twitter profiles (location string). To achieve the goal of having individuals in the target area following multiple bots, the bots maintained a shared list of Twitter accounts that followed-back any of the bots— and all bots followed those real accounts over the following days. As a result, many Twitter users in the target set ended up following multiple bots by the time when the interventions occurred during the period between November 15th to December 2nd, 2014. The distribution of the number of bots followed by other Twitter users during the intervention period is shown in Fig 1B.

Statistics of observed data. The following shows how the observations, including the error bars, in Fig 3 were obtained. For both SC and CC, we investigate how $P(\text{RT})$ changes as a function of k, then iterate over each of the interventions and for each target user we compute the distribution of exposure numbers, according to Eq (2) for SC, and according to the Poisson binomial distribution shown in Eq (8) for CC. These distributions allow us to estimate the number of retweets after k exposures in the following way: Consider a series of events S_1, S_2, . . ., S_n, each representing a user retweeting an intervention-related tweet. For an event S_i, we have probabilities $p_{i,1}, p_{i,2}, . . ., p_{i, n}$ of the event representing $k = 1, k = 2, . . ., k = n$ true exposures. Hence, considering a discrete value $k = j$, the event can belong to bin j with a probability $p_{i,j}$, and it can belong in another bin with probability $1 - p_{i,j}$; i.e., it is drawn drawn from a Bernoulli distribution with $p_i = p_{i,j}$ and $\sigma_i^2 = p_{i,j}(1 - p_{i,j})$. Similarly, the following event is drawn from another Bernoulli distribution independent of the first, and so the distribution of each bin follows another Poisson binomial distribution with $\mu = \sum_i p_i$ and $\sigma^2 = \sum_i \sigma_i^2$. This process approaches the normal distribution $\mathcal{N}(\mu, \sigma^2)$, when the number of Bernoulli draws becomes large due to the central limit theorem (see SI Appendix for details). Thus, can we obtain an approximate distribution for the number of observed retweets for each value of k.

Bayesian information criterion. The Bayesian information criterion (BIC) score is defined as

$$\text{BIC} = -2\ln(L) + k\ln(n), \tag{13}$$

where L is the likelihood of the data given the model, k is the number of model parameters, and n is the number of data points. We compute the likelihood based on the fits to the number of retweets, i.e. fits like those shown in Fig 3B: For each exposure number k, we have (from our previous analysis) an estimate of the number of times, N_k, a user has experienced k exposures. To ensure a discrete number of retweets, we run a series of simulations, computing $P(k|A)$ for each retweeting user and adding 1 to a bin k, which is selected using that probability distribution. We denote the number of retweets in bin k by n_k, and discard bins in which $n_k < 5$. As our models provide the probability $P(\text{RT}|k)$ of each exposure succeeding in eliciting a response from the exposed user, the likelihood of each bin in one such simulation is given by a binomial distribution, and the total likelihood is simply the product of those, i.e.

$$L = \prod_k \binom{N_k}{n_k} P(\text{RT}|k)^{n_k} (1 - P(\text{RT}|k))^{N_k - n_k}. \tag{14}$$

We repeat this simulation 10^3 times for both SC and CC for the full range of values of q.

Conclusion

Diffusion phenomena in social and techno-social systems have attracted much attention due to the importance of understanding dynamics such as disease propagation, adoption of behaviors, emergence of consensus and influence, and information spreading [1, 6–8]. In contrast to modeling epidemics, for which clear laws have been mathematically formulated and empirically validated [2, 4], modeling and understanding information diffusion has proved challenging, in part due to the inability to perform controlled experiments at scale and due to the abundance of confounding factors that bias observational studies [24–26]. Two competing hypothesis have been debated, namely that information spreads according to simple or complex contagion. In this work we test the two hypotheses by creating a controlled experimental framework on Twitter: we deployed 39 coordinated social bots [22] that interacted with a selected cohort of participants (our target population), and carried out a variety of interventions, in the form of attempts to spread new positive messages (i.e., memes for social good). The bots recorded the behavior of the target users and all their interactions with the bots and with other users, while tracking the number of exposures to each message over a period of more than one month. The data we collected allowed us to test two Bayesian models that we derived to capture the diffusion dynamics of simple and complex information contagion. Specifically, in our complex contagion model, we assume that the probability of adoption depends on the number of unique sources of information, rather than the number of exposures.

The statistical evidence clearly shows that the complex contagion model is a better explanation for the observed data than the simple contagion model. This implies that exposures from multiple sources impacts the probability of spreading a given piece of information. This threshold mechanism differs significantly from, say, the spreading of a virus, where many exposures from a single source are sufficient to increase probability of infection. A variety of explanations for the complex contagion hypothesis have been proposed in social theory, including social reinforcement and social influence, echo chambers, human cognitive limits, etc. [1, 3, 9–11, 13, 19]. While our work identifies the type of mechanism according to which information spreads from person to person, much work is still needed to discriminate which factors drive this phenomenon. We expect that future work will explore these factors and further disentangle and explain the dynamics of human communication in social networks.

Supporting information

S1 Appendix. Efficiently generating unique unordered subsets. Explains and evaluates the procedure used to efficiently generate unordered subsets.
(PDF)

S2 Appendix. Gaussian approximation of the poisson binomial. Demonstrates the validity of a Gaussian approximation to the Poisson binomial distribution.
(PDF)

S3 Appendix. Note on data collection. Describes the procedure for recording the data.
(PDF)

Author Contributions

Conceptualization: Piotr Sapieżyński, Emilio Ferrara, Sune Lehmann.

Data curation: Piotr Sapieżyński, Emilio Ferrara, Sune Lehmann.

Formal analysis: Bjarke Mønsted, Emilio Ferrara, Sune Lehmann.

Investigation: Bjarke Mønsted, Piotr Sapieżyński, Emilio Ferrara, Sune Lehmann.

Methodology: Bjarke Mønsted, Emilio Ferrara, Sune Lehmann.

Project administration: Piotr Sapieżyński, Emilio Ferrara, Sune Lehmann.

Software: Bjarke Mønsted, Piotr Sapieżyński, Emilio Ferrara.

Supervision: Piotr Sapieżyński, Emilio Ferrara, Sune Lehmann.

Visualization: Bjarke Mønsted, Piotr Sapieżyński.

Writing – original draft: Bjarke Mønsted, Piotr Sapieżyński, Emilio Ferrara, Sune Lehmann.

Writing – review & editing: Bjarke Mønsted, Piotr Sapieżyński, Emilio Ferrara, Sune Lehmann.

References

1. Castellano C, Fortunato S, Loreto V. Statistical physics of social dynamics. Reviews of modern physics. 2009; 81(2):591. https://doi.org/10.1103/RevModPhys.81.591

2. Goffman W, Newill V. Generalization of epidemic theory. Nature. 1964; 204(4955):225–228. https://doi.org/10.1038/204225a0 PMID: 14212412

3. Granovetter M. Threshold models of collective behavior. American journal of sociology. 1978; p. 1420–1443. https://doi.org/10.1086/226707

4. Daley DJ, Kendall DG. Epidemics and rumours. Nature. 1964; 204(4963):1118–1118. https://doi.org/10.1038/2041118a0 PMID: 14243408

5. Daley D, Kendall DG. Stochastic rumours. IMA Journal of Applied Mathematics. 1965; 1(1):42–55. https://doi.org/10.1093/imamat/1.1.42

6. Centola D, Macy M. Complex contagions and the weakness of long ties. American journal of Sociology. 2007; 113(3):702–734. https://doi.org/10.1086/521848

7. Centola D. The spread of behavior in an online social network experiment. Science. 2010; 329 (5996):1194–1197. https://doi.org/10.1126/science.1185231 PMID: 20813952

8. Onnela JP, Reed-Tsochas F. Spontaneous emergence of social influence in online systems. Proceedings of the National Academy of Sciences. 2010; 107(43):18375–18380. https://doi.org/10.1073/pnas.0914572107

9. Romero DM, 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. ACM; 2011. p. 695–704.

10. Bond RM, Fariss CJ, Jones JJ, Kramer AD, Marlow C, Settle JE, et al. A 61-million-person experiment in social influence and political mobilization. Nature. 2012; 489(7415):295–298. https://doi.org/10.1038/nature11421 PMID: 22972300

11. Ugander J, Backstrom L, Marlow C, Kleinberg J. Structural diversity in social contagion. Proceedings of the National Academy of Sciences. 2012; 109(16):5962–5966. https://doi.org/10.1073/pnas.1116502109

12. Weng L, Menczer F, Ahn YY. Virality prediction and community structure in social networks. Scientific reports. 2013; 3. https://doi.org/10.1038/srep02522

13. Muchnik L, Aral S, Taylor SJ. Social influence bias: A randomized experiment. Science. 2013; 341 (6146):647–651. https://doi.org/10.1126/science.1240466 PMID: 23929980

14. Karsai M, Iñiguez G, Kaski K, Kertész J. Complex contagion process in spreading of online innovation. Journal of The Royal Society Interface. 2014; 11(101):20140694. https://doi.org/10.1098/rsif.2014.0694

15. Kramer AD, Guillory JE, Hancock JT. Experimental evidence of massive-scale emotional contagion through social networks. Proceedings of the National Academy of Sciences. 2014; 111(24):8788–8790. https://doi.org/10.1073/pnas.1320040111

16. Hodas NO, Lerman K. The Simple Rules of Social Contagion. Scientific Reports. 2014; 4:4343. https://doi.org/10.1038/srep04343 PMID: 24614301

17. Ferrara E, Yang Z. Measuring emotional contagion in social media. PloS one. 2015; 10(11):e0142390. https://doi.org/10.1371/journal.pone.0142390 PMID: 26544688

18. Goel S, Anderson A, Hofman J, Watts DJ. The structural virality of online diffusion. Management Science. 2015; 62(1):180–196.

19. Lerman K. Information Is Not a Virus, and Other Consequences of Human Cognitive Limits. Future Internet. 2016; 8(2):21. https://doi.org/10.3390/fi8020021

20. Campbell E, Salathé M. Complex social contagion makes networks more vulnerable to disease outbreaks. Scientific reports. 2013; 3. https://doi.org/10.1038/srep01905 PMID: 23712758

21. Subrahmanian V, Azaria A, Durst S, Kagan V, Galstyan A, Lerman K, et al. The DARPA Twitter Bot Challenge. Computer. 2016; 49(6):38–46. https://doi.org/10.1109/MC.2016.183

22. Ferrara E, Varol O, Davis C, Menczer F, Flammini A. The rise of social bots. Communications of the ACM. 2016; 59(7):96–104. https://doi.org/10.1145/2818717

23. Lehman S, Sapiezynski P. #BanksyInBoston or #BotsInCopenhagen?; 2014. Available from: http://www.aaai.org/Symposia/Spring/sss14symposia.php#ss08

24. McPherson M, Smith-Lovin L, Cook JM. Birds of a feather: Homophily in social networks. Annual review of sociology. 2001; p. 415–444.

25. Aral S, Muchnik L, Sundararajan A. Distinguishing influence-based contagion from homophily-driven diffusion in dynamic networks. Proceedings of the National Academy of Sciences. 2009; 106 (51):21544–21549. https://doi.org/10.1073/pnas.0908800106

26. Shalizi CR, Thomas AC. Homophily and contagion are generically confounded in observational social network studies. Sociological methods & research. 2011; 40(2):211–239. https://doi.org/10.1177/0049124111404820

27. Wang YH. On the number of successes in independent trials. Statistica Sinica. 1993; p. 295–312.

28. Basheer I, Hajmeer M. Artificial neural networks: fundamentals, computing, design, and application. Journal of microbiological methods. 2000; 43(1):3–31. https://doi.org/10.1016/S0167-7012(00)00201-3 PMID: 11084225

29. Shi Y, Eberhart R, Chen Y. Implementation of evolutionary fuzzy systems. IEEE Transactions on Fuzzy Systems. 1999; 7(2):109–119. https://doi.org/10.1109/91.755393

30. Schwarz G. Estimating the dimension of a model. The annals of statistics. 1978;.

31. Raftery A. Bayesian model selection in social research. Sociological methodology. 1995;.

32. Golder SA, Macy MW. Diurnal and seasonal mood vary with work, sleep, and daylength across diverse cultures. Science. 2011; 333(6051):1878–1881. https://doi.org/10.1126/science.1202775 PMID: 21960633

33. Davis CA, Varol O, Ferrara E, Flammini A, Menczer F. Botornot: A system to evaluate social bots. In: Proceedings of the 25th International Conference Companion on World Wide Web. International World Wide Web Conferences Steering Committee; 2016. p. 273–274.