**ORIGINAL ARTICLE**

# Modeling the popularity of twitter hashtags with master equations

Oscar Fontanelli[1] 🔘 · Demian Hernández[2] · Ricardo Mansilla[1]

## Abstract

In this work we introduce a simple mathematical model, based on master equations, to describe the time evolution of the popularity of hashtags on the Twitter social network. Specifically, we model the total number of times a certain hashtag appears on user's timelines as a function of time. Our model considers two kinds of components: those that are internal to the network (degree distribution) as well as external factors, such as the external popularity of the hashtag. From the master equation, we are able to obtain explicit solutions for the mean and variance and construct confidence regions. We propose a gamma kernel function to model the hashtag popularity, which is quite simple and yields reasonable results. We validate the plausibility of the model by contrasting it with actual Twitter data obtained through the public API. Our findings confirm that relatively simple semi-deterministic models are able to capture the essentials of this very complex phenomenon for a wide variety of cases. The model we present distinguishes from other existing models in its focus on the time evolution of the total number of times a particular hashtag has been seen by Twitter users and the consideration of both internal and external components.

**Keywords**  Social network modeling · Hashtag propagation · Master equations

## 1 Introduction

The emergence and popularization of social networking services constitutes an unprecedented social phenomenon that has transformed the way people communicate, get access to different kinds of information, establish communities and many other things. These novel communication channels allow for the fast and massive diffusion of both information and disinformation, a feature that has been well exploited by marketing agencies, social movements, political parties and government agencies, among others. It is, therefore, relevant to understand the process of information diffusion over this kind of networks.

Among the most popular social networking services, such as Facebook, YouTube or Instagram, the microblogging site Twitter stands as particularly effective for information diffusion purposes. According to 2016 data (about.twitter.com), Twitter has approximately 320 million active users (accounts that show activity at least once a month), which represent approximately 9% of total Internet users worldwide (www.itu.int). According to these same sources, approximately 500 million messages are sent over this network everyday.

The growing interest in modeling and understanding different dynamical processes that occur on this social network is manifested in the large number of studies on this matter in recent years. Kawamoto et al. have proposed a multiplicative process model for information spread (Kawamoto 2013; Kawamoto and Hatano 2014). Kwon et al. (2012) and (2013) have proposed models for the evolution of the number of messages, the propensity to send or resend messages and have categorized messages according to predictability and sustainability (Kwon et al. 2012, 2013; Ko et al. 2014). Weng et al. (2012) and (2013) have elaborated an agent-based model for information overflow and have discovered similarities between images diffusion over Twitter and epidemic spreads (Weng et al. 2012, 2013). Mathiesen et al. (2013) have studied scaling laws of big brands tweet-rates, which have been modeled through classic stochastic equations (Mathiesen et al. 2013; Mollgaard and Mathiesen 2015). Sutton et al. have made statistical analysis for the diffusion of official warnings during disasters and have identified some factors that contribute to information diffusion

✉ Oscar Fontanelli
  oscarfontanelli@ciencias.unam.mx

[1]  Centro de Investigaciones Interdisciplinarias en Ciencias y Humanidades, UNAM, Mexico City, Mexico

[2]  Facultad de Ciencias, UNAM, Mexico City, Mexico

(Jeannette Sutton et al. 2015). There are also some works that model topic popularity and information spread with SIR or SIRI-like equations (Xiong et al. 2012; Jin et al. 2013; Skaza and Blais 2017). Bao et al. (2019) have studied the predictability of the number of times a message will be shared or resent (Bao et al. 2019). Bauman et al. have modeled community polarization on social networks and specifically analyzed this with Twitter data (Fabian et al. 2020). Yook et al. (2020) have developed models to account for the observed probability distributions and scaling laws of images and topics popularity (Yook and Kim 2020). There are as well many other studies for different kinds of phenomena that occur on this social network, other than dynamical process, see for example (Gonçalves et al. 2011; Alexandre et al. 2018; Alexandre and Makse Hernán 2019; Zhang et al. 2018). Finally, there are many other studies for this kind of phenomena on other social networks, see for example (Crane and Sornette 2008; Hogg and Lerman 2009; Fang and Huberman 2007; Miotto and Altmann 2014; Miotto et al. 2017; Wang et al. 2018).

In this work, we propose and validate a model, based on semi-deterministic master equations, for the temporal evolution of the number of times a certain label appears on the Twitter network (these labels are called *hashtags*, as we explain in the next section). Notice this is not a model for the number of times a message is sent or shared, but for the number of times it can be seen on the network, which depends on the number of links the nodes that are sending this message have (the degree distribution of the network).

Clearly, a label being shared by nodes with just a few links will behave differently, on a global scale, than a label being shared by nodes with many links. We use this as measure of popularity for the label or hashtag and construct our model under the hypotheses that this popularity is influenced by the degree distribution (a feature that is intrinsic to the network) and also by the extrinsic popularity of the topic (see Bandari et al. (2012) for a discussion on this subject). Data obtained through the Twitter API show that our model is indeed plausible. As far as we know, this is the first attempt to model this quantity with semi-deterministic equations. Even though information diffusion and hashtag propagation is a very complex phenomenon and the model we propose is relatively simple, we see that it is consistent with observed data for an ample diversity of hashtags.

## 1.1 Relation with other models

Modeling and prediction for popularity of tweets, retweets, trends and hashtags in Twitter is a topic that has received a lot of attention in recent years, and there are several models for this phenomena. We show in Table 1 a selection of some of the most relevant models that have been developed in the last ten years. Broadly speaking, we can group these models according to their methodology or according to their main task. For example, works of Ma et al. (2012), (2013), Kupavskii et al. (2012), Kong et al. (2014), Doong (2016), Nadia et al. (2019), Zhang et al. (2016), Hai et al. (2020) are all based on supervised machine learning, deep learning

**Table 1** Schematic view of some relevant models in the last ten years for tweet, retweet, trend and hashtag prediction in Twitter

| References | Methodology | Main purpose |
| --- | --- | --- |
| Ma (2012) and (2013) | Supervised Machine Learning | Prediction of hashtag popularity |
| Kupavski (2012) | Supervised Machine Learning | Retweet prediction |
| Kong (2014) | Supervised Machine Learning | Hashtag bursting prediction |
| Doong (2016) | Supervised Machine Learning | Prediction of hashtag popularity |
| Firdaus (2019) | Supervised Machine Learning | Retweet prediction |
| Zhang (2016) | Deep Learning and Neural networks | Retweet prediction |
| Yu (2020) | Deep Learning | Prediction of peak time for hashtag popularity |
| Pervin (2015) | Statistical analysis | Analysis of hashtag co-occurrence |
| Pancer and Poole (2016) | Hierarchical regression analysis | Prediction of links and retweets |
| Zhang (2013) | Non-linear auto-regression models | Trend prediction |
| Xiong (2012) | Epidemiological models | Modeling of information propagation |
| Jin et al. (2013) | Epidemiological models | Modeling of information cascades |
| Skaza and Blais (2017) | Epidemiological models | Modeling of hashtag dynamics |
| Kawamoto (2013) | Stochastic models and random processes | Modeling the dynamics of retweet activities |
| Ko et al. (2014) | Mathematical modeling | Modeling the propensity to tweet and retweet |
| Mollgaard (2015) | Stochastic and mathematical modeling | Modeling of tweet rates |
| Zhao (2015) | Mathematical and statistical modeling | Prediction of tweet popularity |
| Rizoiu et al. (2017) | Stochastic and mathematical modeling | Prediction of popularity for tweeted videos |
| Bao et al. (2019) | Mathematical modeling and supervised ML | Prediction of tweet popularity and retweets |

and neural networks; models of Pervin et al. (2015), Pancer and Poole (2016), Zhang et al. (2013) are based on statistical analysis and regression techniques; works of Xiong et al. (2012), Jin et al. (2013, Skaza and Blais (2017) propose different variations of epidemiological models for information spread; finally, models of Kawamoto (2013), Ko et al. (2014), Mollgaard and Mathiesen (2015), Qingyuan et al. (2015), Rizoiu et al. (2017), Bao et al. (2019) are either mathematical or stochastic models. Our present work is in line with this last group.

However, none of the models from this last group addresses the topic of hashtag popularity evolution. Works that have addressed this particular phenomenon, such as Ma et al. (2012), (2013), Kong et al. (2014), Doong (2016) are not mathematical-dynamical models but machine learning models. Perhaps the model of Skaza and Blais (2017) is closest to ours, in the sense that it is a mathematical model for hashtag dynamics, but it does not consider stochastic elements. As far as we know, the model that we present in this work is the first mathematical semi-deterministic (stochastic) model for the popularity evolution of hashtags on Twitter.

Our model also integrates internal and external factors that contribute to the propagation of the hashtag. Other models that integrate these two kinds of factors are those of Rizoiu et al. (2017), Ko et al. (2014), but they are not addressing the phenomenon of hashtag propagation. In short, we present here a simple semi-deterministic mathematical model, based on master equations, for hashtag popularity evolution on Twitter that integrates both internal and external factors.

This paper is organized as follows: in Sect. 2 we describe the phenomenon we want to study on Twitter in terms of network theory; in Sect. 3 we develop the model, based on semi-deterministic master equations; we show in Sect. 4 how to obtain solutions for the mean number of messages and its variance; in Sect. 5, we explain how we modeled the extrinsic topic-popularity function; in Sect. 6, we show how we calibrated the model parameters with data and demonstrate that the model is consistent with empirical data from Twitter; finally, Sects. 7 and 8 discuss implications and limitations of the model, as well as future research paths.

## 2 Twitter as a directed network

From a network perspective, Twitter is a directed network where nodes are Twitter users and links represent a follower/friend relationship between them. Users interact on the network by sending messages called *tweets*. Not every user on the network receives all messages. A *follower* of user $i$ is a user that receives all messages sent by $i$. If $j$ is a follower of $i$, then $j$ receives messages sent by $i$ but not the other way

around. If $j$ is a follower of $i$, then we say that $i$ is a *friend* of $j$. In this way, there is a directed link in the network from node $i$ to node $j$, through which a message can flow. Every time $i$ sends a message, all of its friends receive it. If a user receives a message and decides to resend it to his of her followers, we say that this user *retweets* the message. We say that the original message is a tweet and the resent message is a *retweet*. Similarly, a user can *like* a message from its friends, in which case its followers will also see this message. In this way, a specific message can propagate through the network via retweets and likes.

A *hashtag* is a keyword or phrase used to describe a certain topic or theme. Hashtags are preceded by the hash sign (#) and they are widely used because they categorize tweets in a way that is easy for other users to find. Many different messages can be categorized by a common hashtag; if this is the case, all these messages usually speak about a common topic or theme. A certain hashtag propagates through the network if users retweet messages that contain it or if they send new messages categorized by the same hashtag. A hashtag propagates and popularizes when many users are sending messages about a topic of current interest.

A word, phrase, topic or hashtag that is mentioned at a greater rate than others is said to be a *trending topic*. Trending topics become popular either through a concerted effort by users or because of an event that prompts people to talk about a specific topic. We recall that the purpose of this work is to model, by way of master equations, the popularity evolution of hashtags. We develop this model in the next section.

## 3 The model

There are three ways in which messages arrive to a user's timeline (leaving promoted content aside): when a friend of the user tweets, retweets or likes a post with the desired hashtag. We will call a *read* of a hashtag to the event where this hashtag appears on a user's timeline because a friend of this user has tweeted, retweeted or liked a message which includes this hashtag (actually these reads are potential reads, considering that users may never read their entire timelines). Therefore, if a user with $n$ followers sends, resends or likes a message with a certain hashtag, we say that this message has $n$ new reads, indicating that $n$ users have received it (consequently, the hashtag also has $n$ new reads). We want to model the time evolution for the number of reads of all messages categorized by a specific hashtag.

We denote by $X(t)$ the total number of times a certain hashtag have appeared on user's timelines at time $t$. We are not looking at a specific tweet or message, but at all messages categorized by the hashtag. This quantity $X(t)$ does not include the number of times the hashtag has been seen as a result of a

specific search, but only the number of times it has been seen (or potentially seen) because users have tweeted, retweeted or liked a message categorized by the hashtag. In this work, we propose a model for this quantity $X(t)$.

We see that $X(t)$ is a measure of the popularity of a certain topic, phrase or news on the network at time $t$. At any fixed time, we consider $X(t)$ to be a random variable; our goal is to find an equation for the probability of having exactly $x$ reads of a certain hashtag at time $t$, which we denote $P(X = x, t)$. In this way, $X(t)$ is the total number of potential reads at time $t$ of all messages labeled by a specific hashtag and not the number of reads of one particular tweet or message. We intend to model the spread of the hashtag, not the spread of a specific message.

We say that a user *shoots* every time he or she sends or re-sends a message with the hashtag of interest. Users may shoot a message with the desired hashtag because they saw the hashtag on their timeline or because this hashtag has some external popularity (for example, they saw the hashtag on television, other media or on the trending topic list). Let $N$ be the total number of users in the community and let $w(t)$ be the average rate at which users shoot (messages with the desired hashtag). This means that the average probability density for every user to shoot in the time interval $(t, t + dt)$ is $w(t)dt$. Finally, let $f(y)$ be the out-degree distribution of the network, so the probability of a randomly picked user to have $y$ followers is $f(y)$. The contributions to $P(X = x, t)$ are the following:

- There were $x$ reads at time $t$ and nobody shot (which happens with probability $1 - Nw(t)dt$),
- there were $x - 1$ reads at time $t$ and exactly one user with $y = 1$ follower shot (which happens with probability $Nw(t)dtf(1)$),
  ⋮
- there were 0 messages at time $t$ and exactly one user with $y = x$ followers shot (which happens with probability $Nw(t)dtf(x)$).

Since we will consider the limit of very short time intervals, $dt \to 0$, other possible contributions, such as more than one user shooting during the interval $(t, t + dt)$, do not need to be included, as their contribution will be of higher order in $dt$. Summing up all contributions, we get the equation, from the law of total probability,

$$P(x, t + dt) = P(x, t)[1 - Nw(t)dt] +$$

$$Nw(t)dt \sum_{i=1}^{x} P(x - i, t)f(i) + O(dt^2).$$

Rearranging terms and taking the continuous-time limit $dt \to 0$, we obtain the partial differential equation for $P(x, t)$,

$$\frac{\partial P(x, t)}{\partial t} = -Nw(t) \left[ P(x, t) - \sum_{i=1}^{x} P(x - i, t)f(i) \right].$$

We can further approximate the out-degree distribution $f(y)$ to be a continuous distribution with support $[m, \infty)$ so there is a minimum of (possibly zero) $m$ followers per user. With this approximation, we get the equation

$$\frac{\partial P(x, t)}{\partial t} = -Nw(t) \left[ P(x, t) - \int_{m}^{x} P(x - y, t)f(y)dy \right].$$

After a change of variable and rearranging terms, we finally get the equation

$$\frac{\partial P(x, t)}{\partial t} = -Nw(t)P(x, t) + $$
$$Nw(t) \int_{0}^{x-m} P(y, t)f(x - y)dy. \tag{1}$$

This equation, along with the initial condition of zero reads at time $t = 0$,

$$P(x, 0) = \delta(x) \tag{2}$$

constitutes a master equation for the evolution of the number of reads containing a certain hashtag in the network. In a mean-field framework, $w(t)dt$ is the probability density of an average user in the network to send o resend a hashtag in the time interval $(t, t + dt)$; therefore, this function represents a measure of the popularity of the hashtag at time $t$. If the hashtag under consideration is very popular, then it has a high probability of being mentioned in new messages and the messages that contain it have a high probability of being resent. We will refer to this probability rate $w(t)$ as the *hashtag-popularity function*.

## 4 Solutions for the mean and variance

Explicit solutions for Eq.(1) will depend on the forms of the popularity function $w(t)$ and the out-degree or followers distribution $f(y)$ and will be generally not available. However, we can get an equivalent equation for the moment generating function (mgf) of $X(t)$, which we will denote $M_X(s, t)$ and we will be able to utilize it to derive equations for the mean and variance of $X(t)$.

Consider the Laplace transform with respect to $x$,

$$L_s^{(x)}[g(x)] = \int_{0}^{\infty} e^{-sx} g(x)dx.$$

Direct integration shows that the Laplace transform of the integral on the right-size of Eq. (1) is

$$L_s^{(x)}\left[\int_0^{x-m} P(y,t)f(x-y)dy\right] =$$

$$\int_0^\infty e^{-sy}P(y,t)dt \int_m^\infty e^{-sy}f(y)dy =$$

$$L_s^{(x)}[P(x,t)]E_f[e^{-sx}].$$

From the relationship between the moment generating function and the Laplace transform $L_{-s}^{(x)}[P(x,t)] = M_X(s,t)$ we can derive an equation for $M_X(s,t)$ by taking the Laplace transform of Eq. (1),

$$\frac{\partial M_X(s,t)}{\partial t} = N(M_f(s) - 1)w(t)M_X(s,t). \tag{3}$$

Here, $M_f(s)$ is the mgf of the out-degree or followers distribution $f(y)$. Taking the Laplace transform of the initial condition Eq. (2) we get

$$M_X(s,0) = 1. \tag{4}$$

Because of the popularity function $w(t)$, Eq. (3) will be in general a non-linear differential equation for $M_X(s,t)$ and we cannot give a general explicit solution. We can, however, utilize the fact that the n-th moment of a distribution, if it exists, is given by the n-th derivative of the mgf evaluated at zero,

$$E[X(t)^n] = \frac{\partial^n M_X(s,t)}{\partial s^n}\bigg|_{s=0}.$$

For $n = 1$, we obtain a very simple equation for the expectation of $X(t)$,

$$\frac{dE[X(t)]}{dt} = Nw(t)\langle f\rangle, \qquad E[X(0)] = 0,$$

where $\langle f\rangle$ is the first moment of the out-degree distribution, i.e. the mean number of followers of users in the community. This equation has the solution

$$E[X(t)] = N\langle f\rangle \int_0^t w(s)ds. \tag{5}$$

In a similar way, we can get an initial value problem for the second moment,

$$\frac{dE[X^2(t)]}{dt} = Nw(t)[2\langle f\rangle E[X(t)] + \langle f^2\rangle], \quad E[X^2(0)] = 0,$$

where $\langle f^2\rangle$ is the second moment of the followers distribution. Thus,

$$E[X^2(t)] = N\int_0^t w(s)[2\langle f\rangle E[X(s)] + \langle f^2\rangle]ds.$$

Finally, we can have an expression for the variance of $X(t)$,

$$Var[X(t)] = N\int_0^t w(s)\left[2E[X(s)]\langle f\rangle + \langle f^2\rangle\right]ds - (E[X(t)])^2.$$

Integrating by parts the first term of the variance, rearranging terms and simplifying, we get

$$Var[X(t)] = N\langle f^2\rangle \int_0^t w(s)ds = \frac{\langle f^2\rangle}{\langle f\rangle}E[X(t)]. \tag{6}$$

## 5 Modeling the popularity function

Consider the simplest possible case, where the interest a hashtag produces remains constant over time, thus the probability rate $w(t)$ is a constant function. By using $w(t) = c$ where $c$ is a constant, we obtain from Eqs. (5) and (6)

$$E[X(t)] = Nc\langle f\rangle t, \qquad Var[X(t)] = Nc\langle f^2\rangle t.$$

A more realistic consideration is that the interest grows until it reaches a peak, then decays and vanishes for very large times. This behavior can be represented in several ways. Here we will examine one simple possibility, which is a function proportional to a gamma distribution kernel,

$$w(t) = c\frac{e^a}{(ab)^a}t^a e^{-t/b}, \tag{7}$$

where $a, b > 0$ are parameters that control the shape of the interest function and $c$ is the value of $w(t)$ at its peak. Notice that $w(t)$ reaches its maximum value $w_{max} = c$ at $t_{max} = a \cdot b$ and has an inflection point at $t_{inf} = a \cdot b + b\sqrt{a}$. With this popularity function, we get from Eqs.(5) and (6)

$$E[X(t)] = \frac{Ncbe^a\langle f\rangle}{a^a}\gamma(t/b, a+1),$$

$$Var[X(t)] = \frac{Ncbe^a\langle f^2\rangle}{a^a}\gamma(t/b, a+1). \tag{8}$$

Here, $\gamma(x,s)$ is the lower incomplete gamma function, $\gamma(x,s) = \int_0^x e^{-t}t^{s-1}ds$. By utilizing the Stirling approximation for the gamma function

$$\Gamma(z) = \sqrt{\frac{2\pi}{z}}\left(\frac{z}{e}\right)^z\left(1 + O\left(\frac{1}{z}\right)\right),$$

we can approximate the limits for the expectation and variance for very large times,

$$E[X(t)] \longrightarrow \frac{Ncbe^a \langle f \rangle}{a^a} \Gamma(a+1)$$

$$\simeq Nbc\langle f \rangle \sqrt{2\pi(a+1)},$$

$$Var[X(t)] \longrightarrow \frac{Ncbe^a \langle f^2 \rangle}{a^a} \Gamma(a+1)$$

$$\simeq Nbc\langle f^2 \rangle \sqrt{2\pi(a+1)}$$

for large values of the parameter $a$.

Notice that this is not the only way in which we can model the popularity function, but it constitutes a relatively simple function that yields acceptable fits, as we will see in the following section.

## 6 Model calibration and validation

In order to corroborate the validity of the model, we analyzed the time evolution of popular hashtags on Twitter during a two-month period, between May and June 2021. We obtained data through the public Twitter API with the *rtweet* library for the statistical software R (Kearney 2019). We looked for worldwide trends (top 3 trending topics) including hashtags and downloaded statuses with these hashtags at the maximum rate allowed by the public API (18,000 tweets every 15 min), thus collecting an approximate of 100 million statuses (tweets, retweets and replies). After deletion of duplicates, we ended with a dataset of approximately 4 million statuses sent by 1.6 million different users.

We implemented the following pipeline to contrast empirical observations with model predictions:

1. From the sample of tweets, directly compute number of different users $N$, mean number of followers $\langle f \rangle$ and mean square number of followers $\langle f^2 \rangle$.
2. Divide time interval of the sample into $n$ equal length sub-intervals, then compute fraction of different users that sent a message within each sub-interval. This gives us the empirical popularity function $w(t)$.
3. Empirical $w(t)$ is usually very noisy, so we smooth this time series with a simple $k-$point moving average filter. This gives us a smoothed empirical popularity function.
4. Fit parameters for theoretical $w(t)$ with Levenberg-Marquardt non-linear least squares.
5. Utilize the cumulative sum of followers as an empirical approximation for the time evolution of $X(t)$.
6. With fitted $w(t)$ parameters, and knowing theoretical $E[X(t)]$ and $Var[X(t)]$, construct 95% approximate confidence regions for $X(t)$ and contrast with empiric observations. This requires that the variance of the followers distribution is finite.

Model constraints are $a, b, c > 0$ for the popularity function and $\langle f^2 \rangle < \infty$ for the followers distribution (the latter so we can construct finite confidence regions for $X(t)$). We illustrate this procedure in detail in Fig. 1. In this example, we analyzed messages that included #MasterChef, which was the world leading trend on Twitter on June 1st, 2021 (this hashtag was utilized to promote and comment a popular TV show in Argentina). This particular sample of tweets
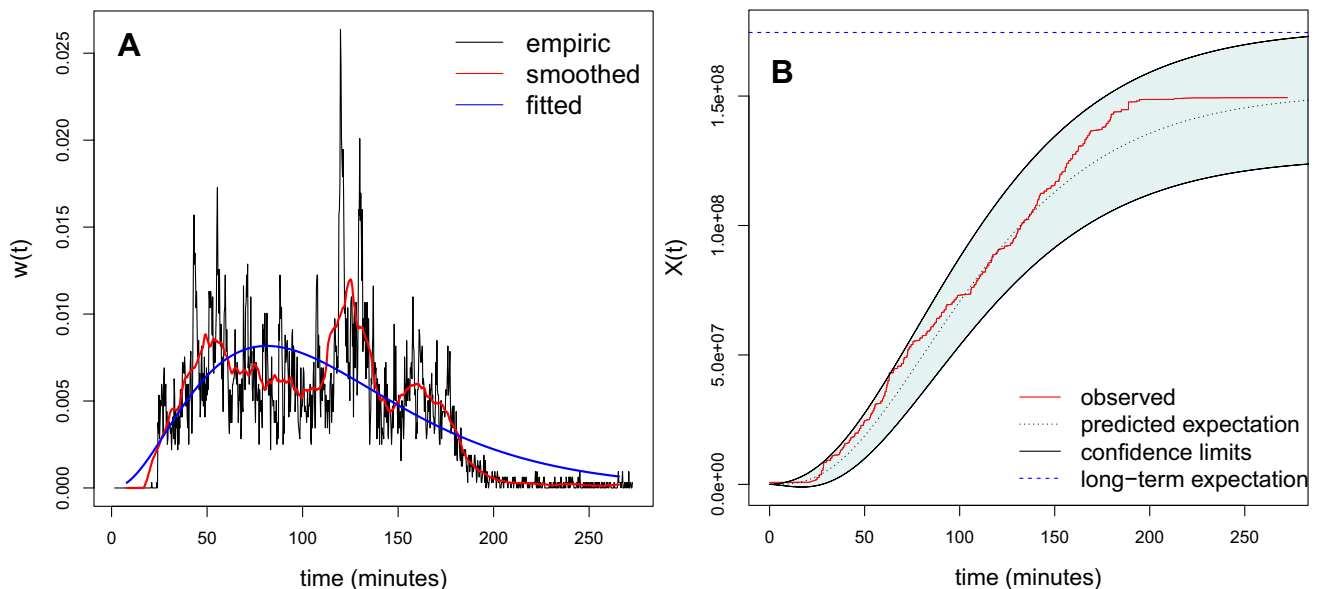


**Fig. 1** Time activity analysis for #MasterChef. On panel (**A**), we see the procedure for getting the empirical $w(t)$ and fitting the theoretical one. On panel (**B**), we compare the observed $X(t)$ with the model's predictions

spanned a time interval of 272 minutes, from 2021-06-01 20:23 CDT (central daylight time) to 2021-06-02 00:55 CDT. On panel A, we see the empirical $w(t)$ (normalized number of messages per time interval) which, as we mentioned, is very noisy; we see in red a smoothed $w(t)$ (with a $k-$ point moving average) and in blue the best fit to the red one. We can see in this plot with some clarity the moment in which people started using this hashtag, the moment of maximum popularity and the moment in which this activity ceased. On panel B, we show in red time vs cumulative sum of followers (the observed $X(t)$); the black dotted line indicates the predicted $E[X(t)]$ (according to our model) and the blue region is the 95% approximate confidence regions for $X(t)$. Finally, the blue dotted line shows the theoretical maximum (limit when $t \longrightarrow \infty$) according to the model. Panels where we do not show this theoretical maximum occur because it falls outside our plotting region (recall this is the limit for very large times).

We repeated this analysis for several hashtags that were world trends between May and June 2021. We show in Fig. 2 a variety of examples.

As an additional test to asses the accuracy of our model, we divided each dataset into a training set and a test set. We show an example of this procedure in Fig. 3. For the #TeacherAppreciationWeek (49,985 statuses), we took a random sample of 30% of the data and utilized this set (this is called the train set) to fit the parameters of the popularity function $w(t)$. In a machine learning approach, this would be the equivalent of the model training stage. With this fitted parameters, we constructed the solution of our model (Eq. 8) and compared this predictions with the remaining 70% of observations (the test set); the latter step would be the equivalent of the model testing stage. In this way, we test the accuracy of our model on previously unseen data, thus getting an idea of how well the model would perform in other situations. For the testing stage, we count the fraction of times the observed $X(t)$ fell inside the predicted confidence region. Since we are constructing approximate 95% confidence regions, this should be a number close to 0.95. We call thus the quantity the "precision" of the model on the test set, which was equal to 0.9452, for the example, we showed in Fig. 3. We got very similar results for all other hashtags we show on this work.

## 6.1 Counterexamples

The model that we introduce in this work is very simple, yet very general and flexible, depending on the form we propose for the popularity function $w(t)$. We have seen that even with a simple form for this function, the model accurately captures the behavior for several different hashtags and trends. It is also important to notice that the spread of a hashtag on Twitter is a very complex phenomenon, and

that it is possible to find examples where even the best fit of our model may not be accurate enough. For illustration purposes, we showed in Fig. 4 three such examples, each one with a different behavior.

See the case of #GreysAnatomy (sample of 63 thousand messages). The time series for this hashtag (empirical $w(t)$) has a very pronounced peak at the very beginning, then abruptly falls. Even though we can fit an empirical popularity function that does not look so bad, the model underestimates the actual spread, as we see on the panel below. This kind of underestimation arises frequently when there are very narrow activity peaks. For a second example, consider the #LetsGoPens. Here we see two activity peaks, which the theoretical popularity function we propose fails to capture. This is reflected on the approximate confidence regions, which does not entirely contain the observed $X(t)$. Another interesting example is that of #RealMadrid. Here we observe that the activity did not start slowly, but rapidly grew from zero to its peak, then slowly fell down. This happened because this particular activity started when an account with disproportionately many followers (that of the Real Madrid Soccer Team) started the hashtag, then it was continued by not-so-popular users. This very popular account enlarges the follower-distribution variance and we see the consequence on the lower panel, where the confidence region is very (perhaps defectively) large.

## 7 Discussion

The mathematical model that we constructed (Eq. 1) for the time evolution of a hashtag popularity is very general and flexible. Depending on the specific forms of the popularity function $w(t)$ and the followers distribution $f(y)$ this equation can represent a large diversity of situations. It is in part because of this generality that we are not able to provide a closed-form solution to this equation; in spite of this, we can find solutions for the mean and variance, therefore, we can construct approximate confidence regions for the quantity of interest $X(t)$.

The particular popularity function that we utilized in this paper is one of the simplest functions that starts at the origin, has a peak, then decays to zero, thus reflecting the grow and decay of the interest in a certain topic or hashtag. In spite of this simplicity, we see with the data we analyzed that the model accurate captures the evolution of $X(t)$ for a wide variety of hashtags, from which some examples are shown in this paper. There are of course some situations where the behavior is more complex and this particular form of $w(t)$ fails to give accurate, quantitative predictions. We showed three examples of this, which represent three archetypal behaviors we observed during our study. It may be that the same model is able to describe these kinds of situations by
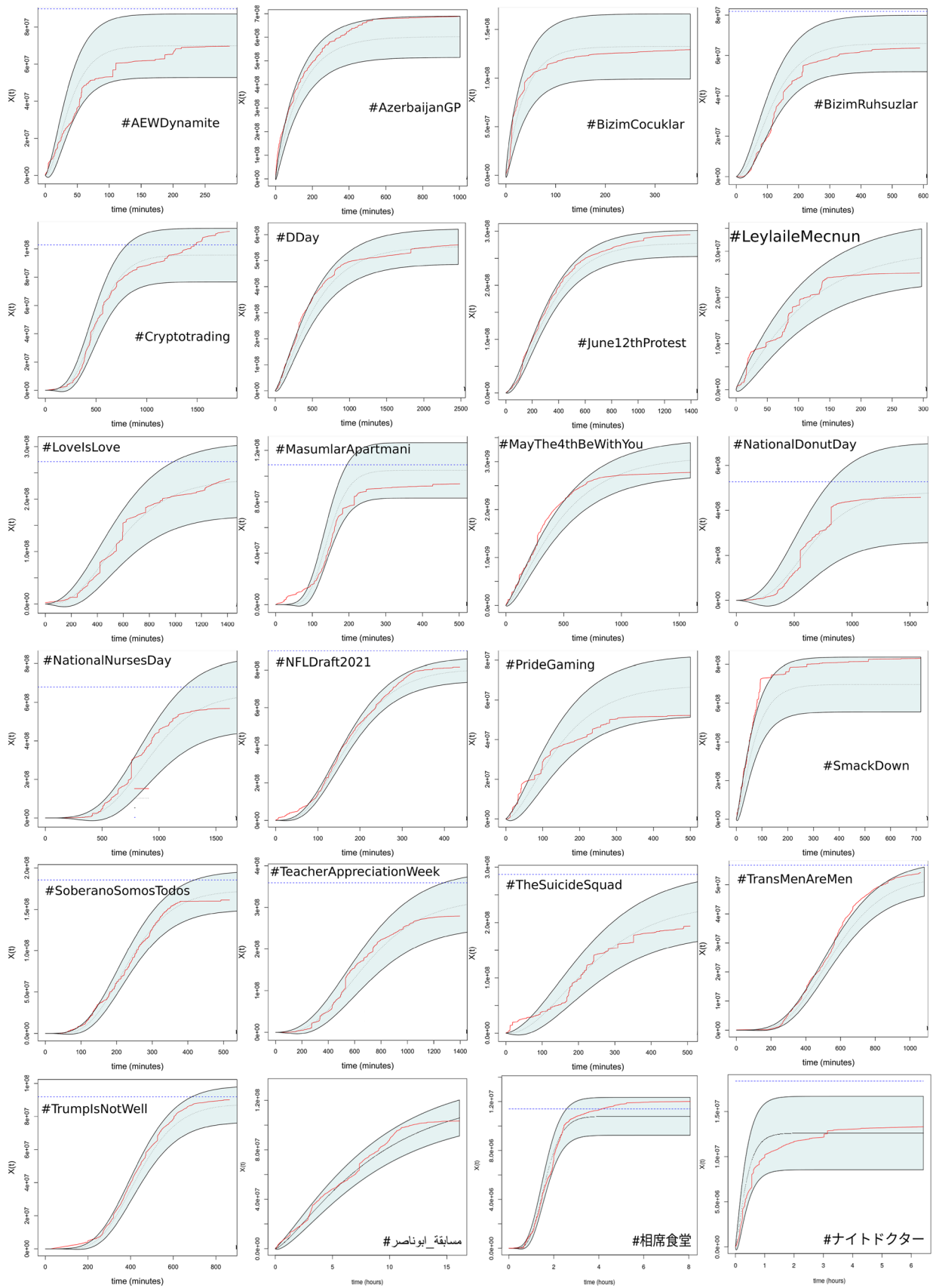
**Fig. 2** Observed time evolution of $X(t)$ (red) and model predictions for a variety of hasthags that were trending topics between May and June 2021
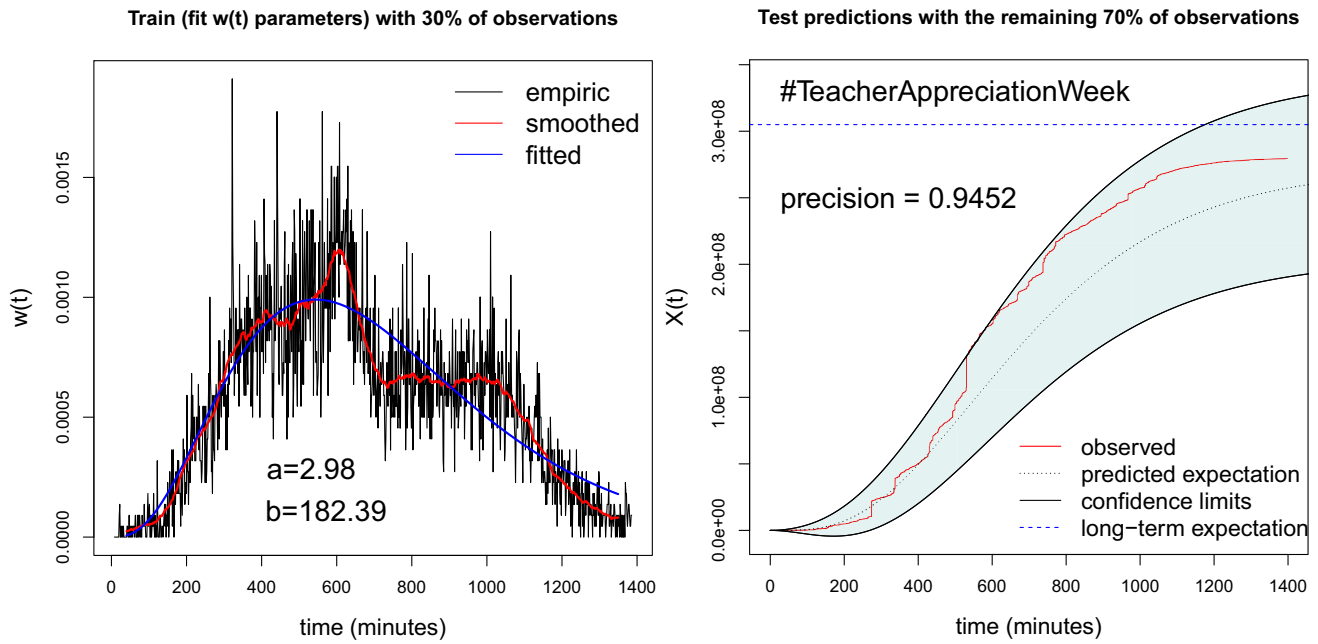
**Fig. 3** In the first stage, we took a train set with 30% of observations to fit parameters, and in a second stage, we compared the model predictions against a test set with the remaining 70% of data. Precision refers to the fraction of times the predicted confidence region contained the actual observation
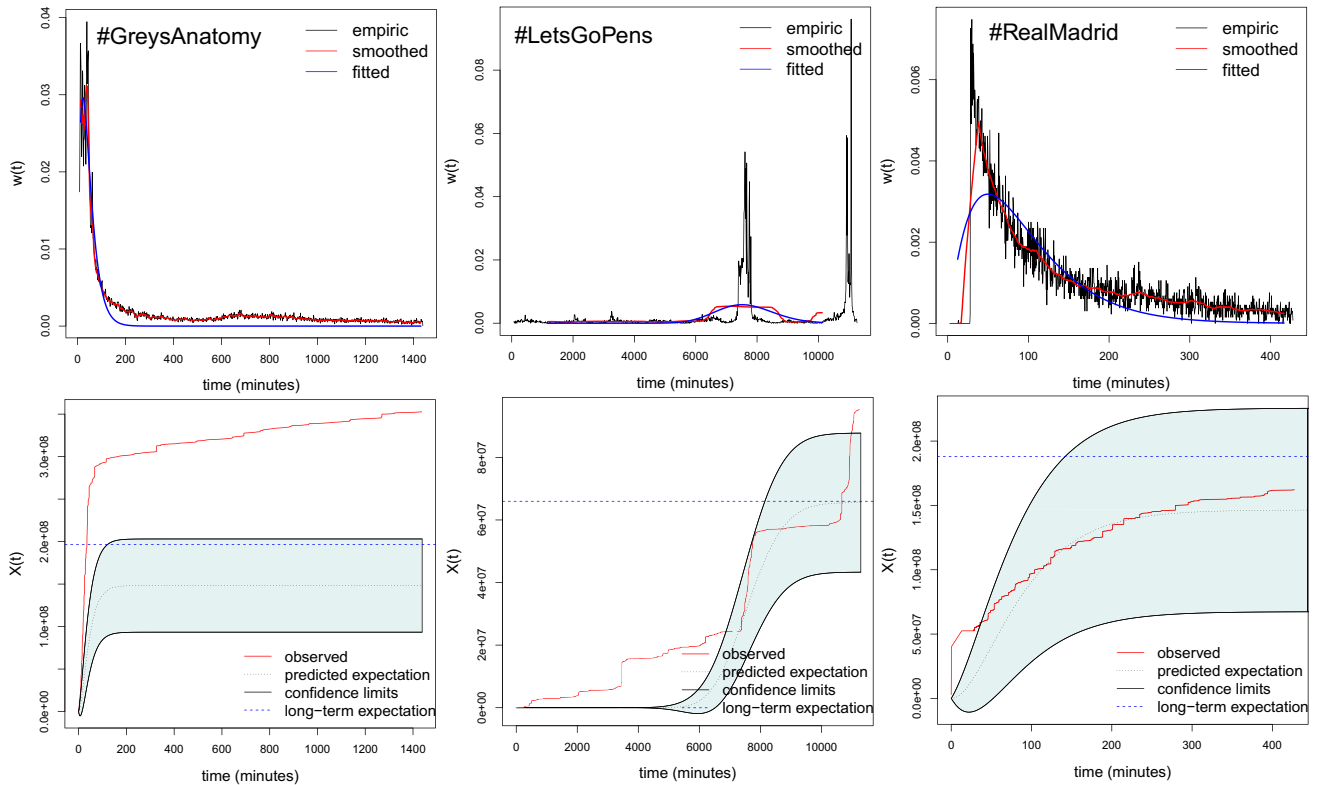


**Fig. 4** Three examples where our model does not entirely capture the popularity evolution of the hashtag, either because a very narrow popularity function (#GreysAnatomy), a popularity function with two local maxima (#LetsGoPens) or a followers distribution with very large variance (#RealMadrid)

modifying the popularity function $w(t)$, which is a matter of future study. Notice that the parameters of the popularity function are fixed, ignoring the possibility that the shape of the popularity function varies with time, for example through a back-feeding process (a popular hashtag gets more and more popular over time). The possibility of a popularity function that updates and that is itself an unknown function is also a matter of future study.

Regarding this popularity function, notice that it may be quantitatively and qualitatively different in different places and situations, therefore requiring a careful tuning and estimation process for each case. The model in Eq. 1 does not say anything about this function $w(t)$ and in this sense, we say that the model is general and flexible, however, we would need to know as much as possible about $w(t)$ for modeling and predicting in specific situations. For example, our entire study was performed during the COVID-19 pandemic (May and June 2021), when Twitter users were likely to behave differently than during normal situations (for example, some studies report a significant increase in social media consumption during lockdowns, see for example (Lemenager et al. 2021; Cellini et al. 2020)).

The measure we utilize for the popularity of a hashtag is the total number of times it appears on user's timelines, which depends on how many users have posted it and how many followers these users have. There are other measures as well, such as the number of likes, retweets and the number of impressions (total number of times a tweet has been seen). The quantity that we studied in this work is very closely related to the number of impressions, except that it does not take into account the number of times a message has been seen because of a direct search. How to extend our model to some other metrics, which are indeed more common and well-known, is an interesting matter. We believe that the consistency between our model and the data is encouraging to study these possible extensions.

The model that we constructed is a model with no compartments, in the sense that nodes (users) are not distinguished between those that have seen the hashtag at time $t$ and those who have not. There are some very interesting and useful models that utilize this approach (Jin et al. 2013; Skaza and Blais 2017; Xiong et al. 2012). However, in this model, we are considering the propagation of a hashtag, not of a particular message. A Twitter user may utilize a particular hashtag because its friends utilized it, hence the user saw it on its timeline (intrinsic factor), or because the hashtag appeared on other media outside Twitter (for example, the #MasterChef that was promoted from a very popular TV show). See (Kwon et al. 2013) for a discussion on how the unpredictability of Twitter is caused by the exposure of users to external environments). Here we decided to model this external exposure to the hashtag via the popularity function $w(t)$ in a way that users may utilize the hashtag even if they have not seen it inside the network. The agreement between our model predictions and the data we analyzed suggests that our approach is also plausible and capable of yielding satisfactory quantitative results. The approach we utilized is in line with and extends the results in the works of Hogg and Lerman (2009), Kawamoto (2013), Kawamoto and Hatano (2014), Mollgaard and Mathiesen (2015) in the sense that we propose a stochastic one-compartment model for information diffusion over a network. Our approach is distinguished from these other approaches in the quantity that we model (popularity of a hashtag), the consideration of intrinsic and extrinsic factors, the construction of approximate confidence regions and the validation with several different hashtags.

The data that we analyzed to calibrate and validate our model was obtained through the public Twitter API. This data collection tool has some limitations: we can only make 18 thousand requests every 15 minutes and we can only access tweets that are 10 days old or newer. We believe that a more comprehensive data base would be helpful and illustrating to see the performance of our model on a more global scale. In spite of these limitations, we observed that our model is consistent with the observations.

Although there are several other models for describing and predicting hashtag propagation, this is, to the best of our knowledge, the first mathematical-dynamical model for this phenomenon. Being a dynamical model, we are in position not only of telling whether a certain hashtag will go viral or not, but also to give some description about the speed at which it spreads over the network. This is a feature that machine learning models do not have, since they are not dynamical models.

This work contributes to the application of stochastic and semi-deterministic equations for the modeling of information diffusion on digital social networks. We extend the above-mentioned lines of work, and we show that even simple models like ours, because of their flexibility and generality, are capable of achieving quantitatively correct predictions for a large number and variety of situations. Future sophistication of this kind of models may prove useful not only for our understating of information propagation but also for early prediction of trends.

## 8 Conclusions

We have presented a mathematical model, based on master equations, for the temporal evolution of the popularity of a certain hashtag or topic on the Twitter network. Unlike other existing approaches, this model focuses on the potential total number of times a certain hashtag has been seen by Twitter users as a function of time. For the construction of this model, we considered two main components that influence this dynamics: on one side, characteristics intrinsic to

the community and the network such as number of nodes and mean and variance of the degree distribution; these are underlying components of the network. On the other hand, we have the time evolution of the interest users have on the hashtag we are modeling, which we quantify as the probability rate for each user in the community of sending a message as a function of time. This popularity function is an extrinsic component influencing this dynamics.

The mathematical model that we present in this paper supposes indeed a great simplification of the problem of message and hashtag propagation on social networks, yet it shows consistency and agreement with the data we observed and analyzed. We believe these findings are encouraging for the further development of more sophisticated semi-deterministic models for complex diffusion processes on digital social networks.

Accurately predicting the evolution and impact a certain tweet or hashtag will have on the network is a difficult task, and it is currently a matter of great interest. With this model, we hope to contribute to the understanding of this phenomenon and to the argumentation for the adequacy and applicability of stochastic and semi-deterministic mathematical models in the study of diffusion processes on social networks. Finally, in this work, we analyzed only the Twitter social network, but this activity may not be completely different from dynamics on other social networks, online or offline; we believe that the present model, though very simple, can give interesting insights into the behavior of other networks.

## Declarations

**Conflict of interest**  The authors declare that they have no conflict of interest.

## References

Alexandre B, Flaviano M, Makse Hernán A (2018) Validation of twitter opinion trends with national polling aggregates: Hillary clinton vs donald trump. Sci Rep 8(1):1–16

Alexandre B, Makse Hernán A (2019) Influence of fake news in twitter during the 2016 us presidential election. Nat Commun 10(1):1–14

Anders Mollgaard, Joachim Mathiesen (2015) Emergent user behavior on twitter modelled by a Stochastic differential equation. PloS One 10(5):10196

Bandari R, Asur S, Huberman BA (2012) The pulse of news in social media: forecasting popularity. In: Sixth international AAAI conference on weblogs and social media

Bao Z, Liu Y, Zhang Z, Liu H, Cheng J (2019) Predicting popularity via a generative model with adaptive peeking window. Physica A 522:54–68

Bruno Gonçalves, Nicola Perra, Alessandro Vespignani (2011) Modeling users' activity on twitter networks: validation of Dunbar's number. PloS One 6(8):1708

Cellini N, Canale N, Mioni G, Costa S (2020) Changes in sleep pattern, sense of time and digital media use during covid-19 lockdown in italy. J Sleep Res 29(4):e13074

Crane R, Sornette D (2008) Robust dynamic classes revealed by measuring the response function of a social system. Proc Natl Acad Sci 105(41):15649–15653

Doong SH (2016) Predicting twitter hashtags popularity level. In: 2016 49th Hawaii international conference on system sciences (HICSS), pp 1959–1968. IEEE

Fabian B, Philipp L-S, Sokolov Igor M, Michele S (2020) Modeling echo chambers and polarization dynamics in social networks. Phys Rev Lett 124(4):048301

Fang W, Huberman BA (2007) Novelty and collective attention. Proc Natl Acad Sci 104(45):17599–17601

Firdaus SN, Ding C, Sadeghian A (2019) Topic specific emotion detection for retweet prediction. Int J Mach Learn Cyber 10(8):2071–2083

Hai Yu, Ying H, Shi P (2020) A prediction method of peak time popularity based on twitter hashtags. IEEE Access 8:61453–61461

Hogg T, Lerman K (2009) Stochastic models of user-contributory web sites. In: Third international AAAI conference on weblogs and social media

Jeannette Sutton C, Gibson B, Phillips NE, Spiro ES, League C, Johnson B, Fitzhugh SM, Butts CT (2015) A cross-hazard analysis of terse message retransmission on twitter. Proc Natl Acad Sci 112(48):14793–14798

Jin F, Dougherty E, Saraf P, Cao Y, Ramakrishnan N (2013) Epidemiological modeling of news and rumors on twitter. In: Proceedings of the 7th workshop on social network mining and analysis, pp 1–9

Kawamoto T (2013) A stochastic model of tweet diffusion on the twitter network. Physica A 392(16):3470–3475

Kawamoto T, Hatano N (2014) Viral spreading of daily information in online social networks. Physica A 406:34–41

Kearney MW (2019) rtweet: collecting and analyzing twitter data. J Open Source Softw 4(42):1829. https://doi.org/10.21105/joss.01829

Ko J, Kwon HW, Kim HS, Lee K, Choi MY (2014) Model for twitter dynamics: public attention and time series of tweeting. Physica A 404:142–149

Kong Shoubin, Mei Qiaozhu, Feng Ling, Ye Fei, Zhao Zhe (2014) Predicting bursts and popularity of hashtags in real-time. In: Proceedings of the 37th international ACM SIGIR conference on research and development in information retrieval, pp 927–930

Kupavskii A, Ostroumova L, Umnov A, Usachev S, Serdyukov P, Gusev G, Kustarev A (2012) Prediction of retweet cascade size over time. In: Proceedings of the 21st ACM international conference on information and knowledge management, pp 2335–2338

Kwon HW, Choi MY, Kim HS, Lee K (2012) Dynamic characteristics of tweeting and tweet topics. J Korean Phys Soc 60(4):590–594

Kwon HW, Kim HS, Lee K, Choi MY (2013) Information-sharing tendency on twitter and time evolution of tweeting. EPL (Europhysics Letters) 101(5):58004

Lemenager T, Neissner M, Koopmann A, Reinhard I, Georgiadou E, Müller A, Kiefer F, Hillemacher T (2021) Covid-19 lockdown restrictions and online media consumption in germany. Int J Environ Res Public Health 18(1):14

Ma Z, Sun A, Cong G (2013) On predicting the popularity of newly emerging hashtags in twitter. J Am Soc Inform Sci Technol 64(7):1399–1410

Ma Z, Sun A, Cong G (2012) Will this# hashtag be popular tomorrow? In: Proceedings of the 35th international ACM SIGIR conference on research and development in information retrieval, pp 1173–1174

Mathiesen J, Angheluta L, Ahlgren PTH, Jensen MH (2013) Excitable human dynamics driven by extrinsic events in massive communities. Proc Natl Acad Sci 110(43):17259–17262

Miotto MJ, Altmann Eduardo G (2014) Predictability of extreme events in social media. PloS One 9(11):1005

Miotto JM, Kantz H, Altmann EG (2017) Stochastic dynamics and the predictability of big hits in online videos. Phys Rev E 95(3):032311

Nadia FS, Chen D, Alireza S (2019) Topic specific emotion detection for retweet prediction. Int J Mach Learn Cybern 10(8):2071–2083

Pancer E, Poole M (2016) The popularity and virality of political social media: hashtags, mentions, and links predict likes and retweets of 2016 us presidential nominees' tweets. Soc Influ 11(4):259–270

Pervin N, Phan TQ, Datta A, Takeda H, Toriumi F (2015) Hashtag popularity on twitter: analyzing co-occurrence of multiple hashtags. In: International conference on social computing and social media. Springer, pp 169–182

Rizoiu M-A, Xie L, Sanner S, Cebrian M, Yu H, Van H (2017) Pascal Expecting to be hip: Hawkes intensity processes for social media popularity. In: Proceedings of the 26th international conference on world wide web, pp 735–744

Skaza J, Blais B (2017) Modeling the infectiousness of twitter hashtags. Physica A: stat mech appl 465:289–296

Wang W, Ming Tang H, Stanley E, Braunstein LA (2018) Social contagions with communication channel alternation on multiplex networks. Phys Rev E 98(6):062320

Weng L, Flammini A, Vespignani A, Menczer F (2012) Competition among memes in a world with limited attention. Sci Rep 2:335

Weng L, Menczer F, Ahn Y-Y (2013) Virality prediction and community structure in social networks. Sci Rep 3:2522

Xiong F, Liu Y, Zhang Z, Zhu J, Zhang Y (2012) An information diffusion model based on retweeting mechanism for online social media. Phys Lett A 376(30–31):2103–2108

Yook S-H, Kim Y (2020) Origin of the log-normal popularity distribution of trending memes in social networks. Phys Rev E 101(1):012312

Zhang A, Zheng M, Pang B (2018) Structural diversity effect on hashtag adoption in twitter. Physica A 493:267–275

Zhang P, Wang X, Li B (2013) On predicting twitter trend: factors and models. In: Proceedings of the 2013 IEEE/ACM international conference on advances in social networks analysis and mining, pp 1427–1429

Zhang Q, Gong Y, Wu J, Huang H, Huang X (2016) Retweet prediction with attention-based deep neural network. In: Proceedings of the 25th ACM international on conference on information and knowledge management, pp 75–84

Zhao Q, Erdogdu MA, He HY, Rajaraman A, Leskovec J (2015) Seismic: a self-exciting point process model for predicting tweet popularity. In: Proceedings of the 21th ACM SIGKDD international conference on knowledge discovery and data mining, pp 1513–1522