A model for meme popularity growth in social networking systems based on biological principle and human interest dynamics

Cite as: Chaos **29**, 023136 (2019); https://doi.org/10.1063/1.5085009
Submitted: 07 December 2018 . Accepted: 29 January 2019 . Published Online: 26 February 2019

Le-Zhi Wang, Zhi-Dan Zhao, Junjie Jiang, Bing-Hui Guo, Xiao Wang, Zi-Gang Huang, and Ying-Cheng Lai

A model for meme popularity growth in social networking systems based on biological principle and human interest dynamics

Cite as: Chaos **29**, 023136 (2019); doi: 10.1063/1.5085009
Submitted: 7 December 2018 · Accepted: 29 January 2019 ·
Published Online: 26 February 2019

View Online Export Citation CrossMark

Le-Zhi Wang,[1,a)] Zhi-Dan Zhao,[1,2,a)] Junjie Jiang,[1] Bing-Hui Guo,[3] Xiao Wang,[4] Zi-Gang Huang,[5,6,b)] and Ying-Cheng Lai[1,7,c)]

AFFILIATIONS

[1]School of Electrical, Computer and Energy Engineering, Arizona State University, Tempe, Arizona 85287, USA
[2]Web Sciences Center, Big Data Research Center, University of Electronic Science and Technology of China, Chengdu 610054, China
[3]School of Mathematics, Beihang University, Beijing 100191, China
[4]School of Biological and Health Systems Engineering, Arizona State University, Tempe, Arizona 85287, USA
[5]School of Life Science and Technology, Xi'an Jiaotong University, Xi'an 710049, China
[6]Institute of Computational Physics and Complex Systems, Lanzhou University, Lanzhou, Gansu 730000, China
[7]Department of Physics, Arizona State University, Tempe, Arizona 85287, USA

[a)]**Contributions:** L.-Z. Wang and Z.-D. Zhao contributed equally to this work.
[b)]**Electronic mail:** huangzg@xjtu.edu.cn
[c)]**Electronic mail:** Ying-Cheng.Lai@asu.edu

ABSTRACT

We analyze five big data sets from a variety of online social networking (OSN) systems and find that the growth dynamics of meme popularity exhibit characteristically different behaviors. For example, there is linear growth associated with online recommendation and sharing platforms, a plateaued (or an "S"-shape) type of growth behavior in a web service devoted to helping users to collect bookmarks, and an exponential increase on the largest and most popular microblogging website in China. Does a universal mechanism with a common set of dynamical rules exist, which can explain these empirically observed, distinct growth behaviors? We provide an affirmative answer in this paper. In particular, inspired by biomimicry to take advantage of cell population growth dynamics in microbial ecology, we construct a base growth model for meme popularity in OSNs. We then take into account human factors by incorporating a general model of human interest dynamics into the base model. The final hybrid model contains a small number of free parameters that can be estimated purely from data. We demonstrate that our model is universal in the sense that, with a few parameters estimated from data, it can successfully predict the distinct meme growth dynamics. Our study represents a successful effort to exploit principles in biology to understand online social behaviors by incorporating the traditional microbial growth model into meme popularity. Our model can be used to gain insights into critical issues such as classification, robustness, optimization, and control of OSN systems.

With advances in information technologies, a novel class of complex dynamical systems has emerged: online social networking (OSN) systems. The complexity of OSN systems is enormous: posting and sharing of messages by users, sudden occurrence of breaking news events, random drifts in user interests, etc., all leading to drastic variations of the network structure and dynamics with time and making (big) data analysis an essential approach to uncovering the inner dynamical working of these systems. A phenomenon that has attracted recent attention is growth dynamics of memes such as news, ideas, knowledge, or rumors in OSN systems. Previous models focusing on the individual level were unable

to account for the common phenomenon of group popularity, raising the need to develop a comprehensive model that incorporates heterogeneity of users and memes to quantitatively describe the collective dynamics of meme popularity. Another challenge in the construction of a model for meme popularity lies in its distinct growth behaviors in different OSN systems. Our analysis of five big data sets from diverse social networking platforms has revealed at least three characteristically different types of behaviors: linear, plateaued (or "S" shaped), and exponential growth in time. Is it possible to construct a single model that can explain the distinct growth behaviors? This paper provides an affirmative answer. The general principle underlying our work is that, while OSN systems are human-engineered with vast complexity, nature has solved difficult problems in complex systems. Animals, plants, microbes, and even cells are extremely well self-organized natural systems with superior functions and efficiency. The first ingredient of our model is then an approximate equivalence between meme evolution in OSN systems and microbial cell population growth. This leads to a probabilistic, population-level base model, where at any given time, a cell can experience one of the three possible events: division (generation), death, and survival. Regarding memes as the "microscopic" elements of OSN systems, the possible events that can happen to a meme are similar: posting/forwarding, being overwritten (exclusion), or simple survival, which are equivalent, respectively, to cell division, death, and survival in a microbial system. Because meme growth is a kind of human behavior, it is also necessary to consider additional model ingredients beyond the biological equivalence. The second ingredient of our model is then to incorporate human-interest dynamics into the bio-inspired base growth model. The outcome is a hybrid model for meme popularity dynamics, which contains four free parameters that can be determined from data. The striking result is that the model can predict the detailed meme popularity growth behaviors in all real OSN systems studied, regardless of their characteristically distinct origins, thereby providing a solid ground for its validity and universal applicability. To be able to predict the dynamical evolution of memes is of great social, economic, and political interest. What we have achieved in this paper is a universal model for this task with minimally required information.

I. INTRODUCTION

Online social networking (OSN) systems are now ubiquitous and play an increasingly important role in the modern society, as they provide unprecedented platforms supporting communications among a vast number of users all over the world. Due to the availability of massive data sets from OSN systems, quantitative system analyses have become possible.[1–24] Previous efforts focused on issues such as network and opinion co-evolution,[1] user behavior modeling on networks,[2,7] the dynamics of users' activity across topics and time,[3,10] human interest dynamics in e-commerce and communication,[14] evolutionary dynamics of forwarding network in the Weibo platform,[15] competition among different Twitter topics,[8,16,20,25] popular topic-style analyses in the Twitter-like social media,[11–13] information diffusion patterns in different domains,[5,9] and the effect of coexistence of different OSN network services.[19] These studies mainly considered two issues common to various OSN systems: the collective behaviors of users and the dynamics of posts or memes. Based on empirical findings, e.g., power-law scaling relations, mathematical, and/or physical models have been developed to predict the scaling laws. For example, a two-layer model has been proposed to characterize the viral dynamics and media influence,[26] a branching process has been used to explain the power-law distribution of meme popularity,[16,20] and a Bayesian probabilistic model has been developed to characterize the evolution of tweets.[27] While these models are able to simulate or predict certain aspects of meme popularity in real OSN systems, they are often dependent upon the structure of the underlying social network, limiting their applicability to specific types of social networking platforms with specialized functions. Due to the vast complexity and diversity of the OSN systems, a quantitative, generally applicable model for the dynamical evolution of key variables of empirical interest is lacking. In this paper, exploiting biological principles, we develop a universal model to explain the characteristically distinct behaviors of meme growth observed from diverse OSN systems.

Previous efforts in this field are briefly summarized as follows. We define memes broadly as some items that serve to attract user attention and induce heterogeneous dynamical behaviors in the OSN. Especially, memes are referred to not only as news pieces, ideas, certain information pieces, knowledge items, or rumors but also as bookmarks, movies, Weibo messages, music pieces, etc. The network to characterize the user-item relationship typically possesses a bipartite structure.[28] Some memes can go viral, some might receive constant attention, and some simply get ignored. To uncover the mechanisms that drive the fates of different types of memes on a microscopic scale is a challenging task. In this regard, in a previous work,[8] it was found that in OSN systems, the distribution of meme popularity is typically heterogeneous as a result of the mutual "competition" among different coexisting memes for users' attention. This observation provides the base for the proposal of a theoretical model to describe the dynamical evolution of memes with a particular focus on the influence of user actions on information diffusion.[16,20] The mathematical backbone of the theory is branching processes, and it has successfully explained certain empirical observations such as the distribution of meme popularity growth associated with the *Twitter* data sets. A key assumption of the theory, which is somewhat ideal and thus makes feasible an analytic treatment, is that users have constant activity rates, and memes are equipped with the same "fitness." While the idealization of identical users and memes

enables a mathematical analysis, the key ingredient of meme dynamics in real world OSN systems is heterogeneity in user and meme behaviors. An alternative modeling approach was to apply a self-exciting point process (e.g., the Hawkes process[29]) to predict the popularity of tweets based on partial information about the network structure and observations of the retweeting times.[27,30] This type of models can successfully predict the total final number of retweets but, because of the focus on information diffusion at the individual level, they are unable to account for the common phenomenon of group popularity. It is also worth noting that memes popularity dynamics are distinct from epidemic spreading dynamics on complex networks[31–33] in that, in the former, individuals receive and spread a large number of memes while in the latter, the type of viruses is typically one or two. The theories and computational methods developed in the past on network spreading dynamics[33] are generally *not* applicable to meme popularity dynamics.

At present, a comprehensive model that incorporates the heterogeneity of users and memes to quantitatively describe the collective dynamics of meme popularity is lacking. A more significant challenge in the construction of a model for meme popularity lies in its distinct growth behaviors in different OSN systems. In particular, by analyzing five big data sets from diverse social networking platforms, we find three characteristically distinct types of behaviors: linear, plateaued (or "S" shaped), and exponential growth in time. Is it possible to construct a single model that can explain these distinct growth behaviors? Naturally, such a model will contain a small number of free parameters whose values depend on the specific OSN system and can be estimated from data. Except for the differences in the values of the free parameters, the basic elements of the model are identical for OSN systems from diverse contexts. In this sense, the model can be regarded as universal. Our main idea is to exploit biological principles (biomimicry) to develop such a model. The guiding principle is that, while OSN systems are man-made with vast complexity, nature has solved difficult problems in more complex systems, especially in biology. Nevertheless, because meme growth dynamics are driven by human behaviors, it is also necessary to incorporate human aspects into the model. For this purpose, we exploit a model for human interest dynamics.[14] The final outcome is a hybrid model for meme growth based on the combination of biomimicry inspired by cell growth in microbial ecology, human behaviors, and empirical laws extracted from big data sets. The model can accurately predict characteristically distinct growth behaviors in a diverse array of OSN systems.

In Sec. II, we identify the similarities between cell growth in microbial ecology and meme growth in OSNs to establish our biomimicry principle. In Sec. III, we develop a hybrid model incorporating human interest dynamics into our cell-growth model for meme evolution. In Sec. IV, we present numerical support for our hybrid model based on five big data sets from OSN systems and provide a mathematical analysis. In Sec. V, we summarize the main results and discuss possible model generalization.

II. BIOMIMICRY PRINCIPLE AND EMPIRICAL SUPPORT

A. Biomimicry: From microbial ecology to social network

In microbial ecology, quantitative analysis of cell population is fundamental.[34–46] Due to computational constraints, most cell models are at the population level because it is impractical to monitor the state of each individual cell and count the number of living cells at every time step. A typical cell population model contains three ingredients: cell division (birth), death, and survival. Likewise, in an OSN system, it is unrealistic to study the behavior of every single post, but it is feasible to obtain data reflecting the collective behaviors of these posts. Especially, a meme can be generated from one user and passed onto another in response to certain social event (birth), it can disappear if there is no or no longer any interest in it (death), or it can simply be associated with the same user without any change (survival). This analog suggests that, dynamically, a meme associated with an individual is equivalent to a cell in a microbial system.

To place the cell-meme correspondence on a quantitative footing, we consider the probabilistic model of cell growth and mortality.[40] At each basic time step, three possible events can occur to an individual cell: it can divide, can die, or remain alive. The probabilities of the respective events are represented as functions of time with parameters estimated from experimental data. Considerations of the events and the associated probabilities provide a base for us to formulate a meme probabilistic model. An illustrative example is presented in Fig. 1. In this toy OSN system, there are seven users and four types of memes that constitute a bipartite network, as shown in Fig. 1(a), where the former and the latter are connected with each other by posting or forwarding actions. Especially, two users are directionally connected if a meme is forwarded from one to another. Figure 1(b) shows an example of the posting or forwarding action for each type of meme from the users point of view, which is essentially a space-time representation of how the memes are created and evolved. Here, $t = 0$ specifies the initial time of the observational window of the system. For example, users U1, U4, U5, and U7 post meme types M1, M2, M3, and M4 at time $t = 0$, $t = 1$, $t = 2$, and $t = 2$, respectively, while users U2, U3, and U4 forward the M1 type at $t = 3$, $t = 4$, and $t = 5$, respectively. As a result, four users (U1, U2, U3, and U4) are linked to the meme type M1, as indicated in the bipartite network in Fig. 1(a). Similarly, user U5 forwards meme type M4 at $t = 3$, M3 at $t = 4$, and M2 at $t = 5$, and user U6 forwards type M4 at $t = 6$. For a meme associated with a user at a given time, it can be forwarded, be overwritten (excluded), or simply survive, as indicated in Fig. 1(c), which corresponds to the three possible events that can occur to a cell in microbial ecology: division, death, or remaining alive. Figure 1(d) shows the same example as in Fig. 1(b), but from meme's point of view to illustrate the three events for each of the four meme types. From a global perspective, a population of a specific meme starts to grow when it is first posted. The population size increases as users continue to forward this meme. Between each pair of consecutive forwarding events, the meme is in

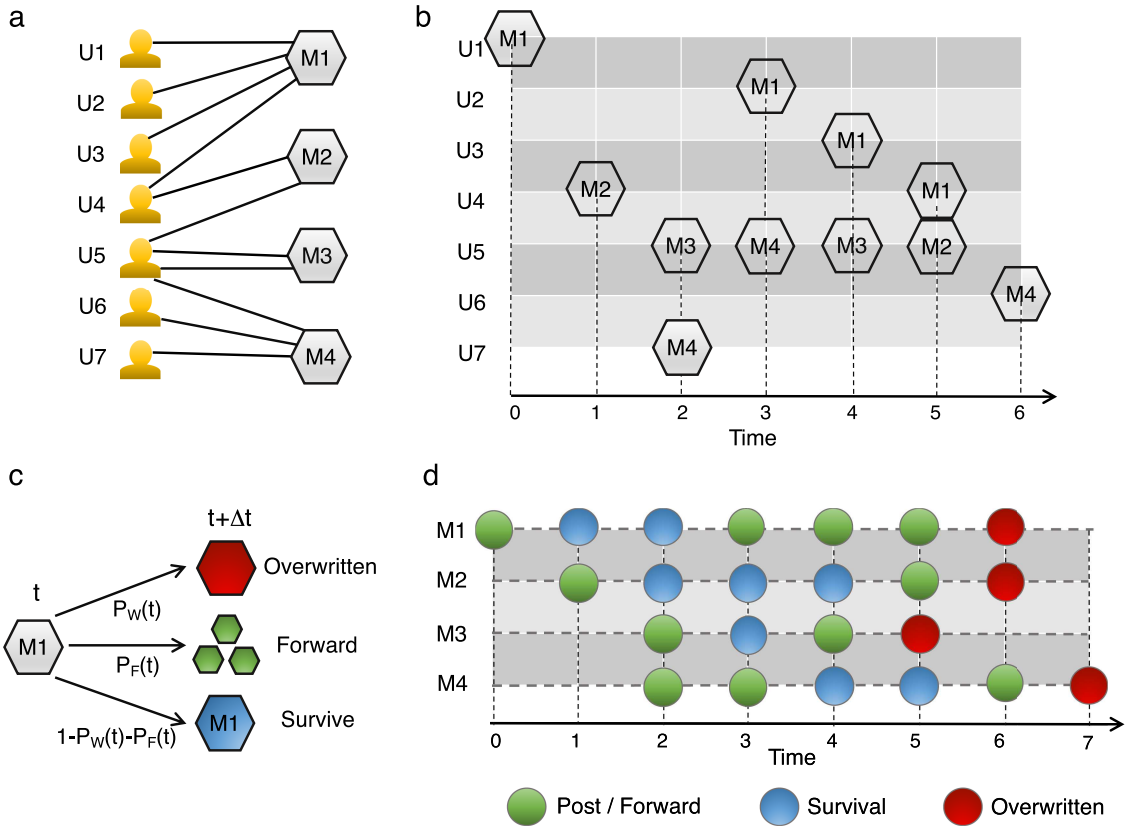

FIG. 1. Schematic illustration of evolutionary dynamics of meme popularity. The toy system has seven users and four types of memes, where a meme associated with an individual user corresponds to a cell in microbial ecology. (a) Bipartite graph representation of users and memes, which are connected with each other by posting or forwarding actions. (b) Representation of forwarding actions at different time. Each horizontal line is associated with an individual user and each hexagon represents a meme. For a horizontal line, a hexagon appearing at a time indicates that a meme (regardless of its type) has been created or forwarded at this time. (c) The three events that can occur to a meme at a given time, together with the respective probabilities: $P_F(t)$—the probability of being forwarded, $P_W(t)$—the probability of being overwritten (exclusion probability), and the survival probability $1 - P_F(t) - P_W(t)$. (d) An example of the time evolution of meme popularity, where the green circles represent posting or forwarding events, the blue circles define survival events, and the red circles correspond to exclusion events. A meme is regarded as dead (excluded) after the last forwarding event.

the survival state. After the last forwarding event, the meme is regarded as being excluded or dead.

The population of a type of memes at any time is determined by the numbers of newly forwarded, excluded, and survived posts. Let $F(t)$, $W(t)$, and $S(t)$ be the numbers of forwarded, excluded, and survived memes, respectively, at time t. The meme population at this time can be written as

$$N(t) = S(t) + F(t). \tag{1}$$

The fraction of meme population is defined as

$$P_N(t) = \left(\frac{S(t)+F(t)}{F(t)+S(t)+F(t)}\right) \Big/ \max\left(\frac{S(t)+F(t)}{F(t)+S(t)+F(t)}\right),$$

and the interactions among memes are described by the following normalized forwarding and exclusion probability functions:

$$P_F(t) = \left(\frac{F(t)}{F(t)+S(t)+W(t)}\right) \Big/ \max\left(\frac{F(t)}{F(t)+S(t)+W(t)}\right),$$

$$P_W(t) = \left(\frac{W(t)}{F(t)+S(t)+W(t)}\right) \Big/ \max\left(\frac{W(t)}{F(t)+S(t)+W(t)}\right),$$

respectively, which determine the values of $F(t)$ and $W(t)$.

B. Validation of biomimicry principle with empirical online data sets

The massive empirical data sets analyzed in this article are from large-scale online systems: Delicious, *Douban*, and *Weibo*. The basic statistical properties of four data sets are listed in Table I, where the term "Records" represents the number of records in each raw data set, "Memes" denotes the total number of memes in each raw data set, "Users" indicates the total number of users involved in each raw data set,

TABLE I. Basic properties of four data sets studied in this paper. The resolution unit of time is Days.

Data sets	Records	Memes	Users	Duration (months)	References
Delicious	361,928,091	886,405	43,968,955	42	47
Douban Book	20,199,759	455,177	557,879	53	14
Douban Movie	65,205,220	504,066	86,503	52	14
Douban Music	25,596,271	403,835	395,035	50	14

and "Duration" is the duration of each processed data set. The details of these data sets are described as follows.

Delicious is a web service focusing on helping users collect bookmarks. Each record consists of the operation time, user's ID, Universal Resource Locator (URL), and the tag of URL. In this data set, a meme is defined as a bookmark, and users' collections of bookmarks are regarded as forwarding actions. An "excluded" (or "overwritten") bookmark at time t means that this bookmark no longer appears in the system after time t.

Douban is a major Chinese Social Networking Service website. It allows users to record information and make recommendations related to books, movies, music, etc. Each record contains users id, time stamps, and item rating actions. We analyze three Douban data sets: *Douban Book*, *Douban Movie*, and *Douban Music*. We define each rated item as a meme and treat each rating action as a forwarding event. A book (movie or a piece of music) not recurring at time t is regarded as "excluded" (or "overwritten").

Sina Weibo is by far the largest and most popular microblogging website in China: it is a widely used twitter-like microblogging social network medium with more than 500 million registered users in China.[48] The appealing features of the data include wide publicity, real-time availability of information, and message compactness. Similar to Twitter, Weibo attracts users through all kinds of breaking news and spotlight topics. All users can see messages, called Weibo in Chinese, published by concerned users. Given a specific topic of interest, an individual can participate by retweeting (forwarding) or tweeting (posting) any interesting Weibo.[49] In this data set, for each message with the forwarding information, we record the original Weibo id, user id who forwards this Weibo, and the time of creation. Each message represents a meme with a possible forwarding action.

The probabilities $P_F(t)$ and $P_W(t)$ are key quantities in our meme popularity model, which can be estimated from the empirical data. Figure 2 shows the forwarding and exclusion probabilities for the five data sets: *Delicious*, *Douban Book*,

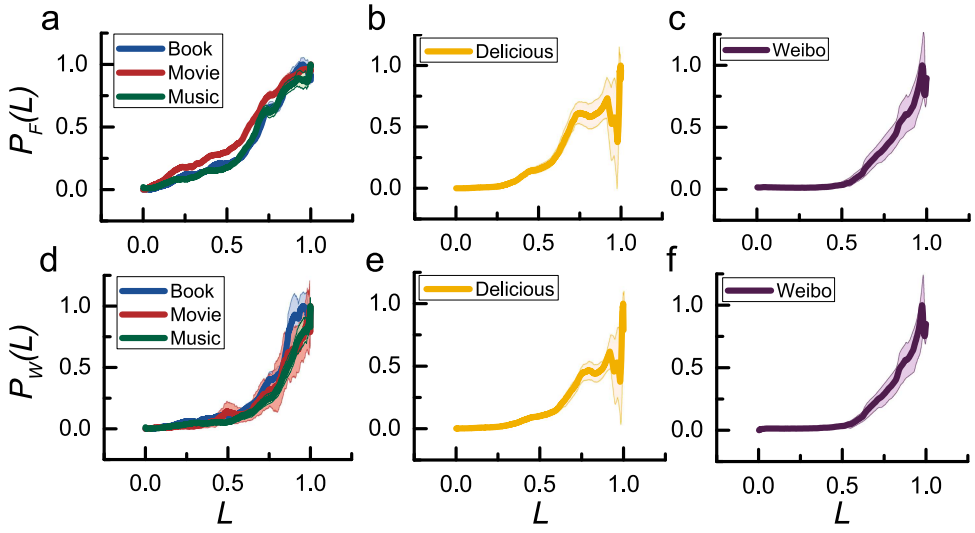

FIG. 2. Time evolution of forwarding and exclusion probabilities estimated from empirical data sets. In each panel, lines and the corresponding shadow areas indicate the average forwarding or exclusion probabilities and the error bars, respectively. The time and the probabilities have been normalized to facilitate a quantitative comparison among the five data sets that arise from different social networking contexts. The probabilities are for the group of data sets *Douban Book*, *Movie*, and *Music* [(a) and (d)], *Delicious* [(b) and (e)], and *Weibo* [(c) and (f)]. Note that it is not necessary that the probabilities $P_A(L)$ and $P_W(L)$ are increasing functions of time. However, since the focus is on the meme growth dynamics in the normalized time interval considered ($L \in [0, 1]$), these probabilities exhibit an increasing trend. Also, since the probabilities are normalized by their respective maximum values at $L = 1$, their values being unity does not mean that all users are active at $L = 1$.

Douban Movie, Douban Music, and *Sina Weibo* (or *Weibo*). We find it convenient to organize the five data sets into three groups in terms of their nature: *Douban Book*, *Movie*, *and Music* (group 1), *Delicious* (group 2), and *Weibo* (group 3). To facilitate a comparison of the dynamical behaviors of different systems, we define the normalized time $t \to L$. For convenience, we choose the range of L to be $[0,1]$, where $L = 1$ is defined as the time when the fraction $P_N(L)$ of meme population reaches the maximum. The probabilities $P_F(t) \to P_F(L)$ and $P_W(t) \to P_W(L)$ are also normalized by their respective maximum values at $L = 1$. Initially, the two normalized probabilities have relatively low values, but they begin to increase after certain time. This initial "silent" phase corresponds to the lag phase that occurs before the accumulation phase for cell growth in microbial ecology,[34–46] where the growth and mortality rates are expected to be low at the beginning but increase with time. The striking phenomenon is that, for the five data sets arising from diverse social networking contexts, the time evolution of the probabilities exhibits quite similar features, suggesting a universal mechanism underlying the dynamical evolution of meme popularity.

III. MODEL CONSTRUCTION

A. Basic principles underlying the construction of a universal model for meme popularity dynamics

Our first step is to hypothesize the equivalence between meme evolution in OSN systems and microbial cell population growth so as to develop a probabilistic, population-level base model. In such a dynamical evolution model of cell population,[34–46] at any given time, a cell can experience one of the three possible events: division (generation), death, and survival. Likewise, memes are the "microscopic" elements of OSN systems. At a given time, the possible events that can happen to a meme are similar: posting/forwarding, being overwritten (exclusion), or simple survival, which are equivalent, respectively, to cell division, death, and survival in a microbial system, as schematically illustrated on the left side of Fig. 3. This equivalence, or biomimicry, leads to a base model for meme population growth. Additional model ingredients beyond the biological equivalence must be sought. This is reasonable as OSN systems are man-made and, as such, human factors can play a significant role in the dynamics. The second step in our model building is then to incorporate human interest dynamics that, intuitively, are correlated with meme popularity. Incorporating a general model for the dynamical evolution of human interest into the bio-inspired base growth model, we arrive at a hybrid model for meme popularity dynamics, as shown on the right side of Fig. 3. The final model contains four free parameters that can be determined from data, as we demonstrate using empirical big data sets from diverse OSN systems in interest sharing (e.g., various *Douban* platforms and *Delicious*) and the online OSN platform *Weibo*. The striking result is that the model can predict the detailed meme popularity growth behaviors in all real OSN systems studied, regardless of their characteristically distinct origins, thereby providing a solid ground for its validity and universal applicability. While our model predicts three distinct meme popularity growth behaviors from five data sets, with the values of the four free parameters determined from data, the model has the capability to predict growth behaviors beyond the three types.

It should be emphasized that, while the cell growth model captures certain features of the meme dynamics, our data analysis indicates that large errors can arise when attempting to predict the detailed growth dynamics of meme population

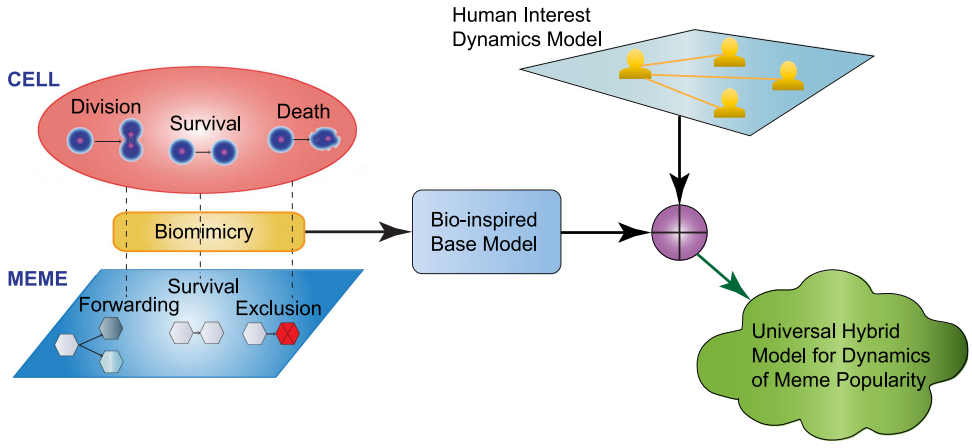

FIG. 3. Illustration of the basic principles underlying the construction of a universal model for meme popularity dynamics: biomimicry and human interest dynamics. Cell population dynamics in microbial ecology contain three basic elements: cell division, death, and survival. The dynamical evolution of memes in OSN systems has three corresponding elements: forwarding, exclusion, and survival. A base model for meme popularity dynamics can then be constructed according to the cell growth model in biology. Human factors, however, play a significant role in the meme evolution and, therefore, must be taken into account. A combination of the biology inspired base model and human interest dynamics model leads to a universal hybrid model for meme popularity dynamics.

in terms of the time evolution of the forwarding and exclusion probabilities estimated directly from the empirical data, as shown in Fig. 2. Especially, while the biologically motivated model is able to provide an overall picture (statistical behaviors) about the meme growth dynamics, it is not possible for it to capture the minuscule details in each data set.

B. Beyond biomimicry: Incorporation of human interest dynamics

While biomimicry can provide pivotal insights into uncovering the laws governing human social networking systems, certain aspects of the human behaviors are fundamentally absent in any biological system. In particular, in the context of our present study, while microbial cell evolution provides a base for constructing models for online social behaviors, it has a purely biological origin while OSN systems are man-made. There ought to be some differences between the two types of systems. Indeed, in spite of the remarkable agreement among the forwarding and exclusion probability functions across OSN systems of different nature (Fig. 2), a detailed examination of the evolution of the meme population N(t) for the three groups of data sets reveals some discrepancies, indicating the need to include additional factors that are not present in the microbial cell evolution model. In fact, it is necessary to incorporate human factors into models of OSN systems. For the dynamical evolution of meme popularity, the most pertinent factor is human interest.

Meme growth in OSN systems is a human behavior. In general, human behaviors are driven by human interest. In the past decade or so, there has been a great deal of effort in modeling and understanding human behavior and interest dynamics.[14,51–57] Of particular interest are characteristics such as the distribution of the interevent time of human behaviors, the distribution of the return time to revisit a particular interest, interest ranking and transition, and the distribution of the time for which an interest lasts. Because of the sensitive dependence of human interest on environment factors, it was previously speculated that the dynamical underpinnings of human interest are random,[8,58,59] and this led to the development of the Markovian type of models for human interest where an individual's history of interests (except those in the immediate past) plays no role in his/her present action.[60–62] Deviations from the Markovian dynamics were reported.[14,63,64] For example, a systematic analysis of a number of big online data sets revealed that an algebraic (power-law) scaling behavior, which is characteristic of non-equilibrium complex systems, governs both the interevent time and event determination statistics associated with human interest dynamics.[14,51] This implies that there are intrinsic dynamical rules underlying the human interest dynamics. Three such rules were hypothesized: preferential return, inertial effect, and exploration of new interests, and a mathematical model was developed to explain the empirically uncovered algebraic scaling laws.[14]

To take into account human interest dynamics in constructing a meme popularity model, we exploit previously studied scaling laws associated with models:[14,56,57] the bursting characteristic of human behaviors, the algebraic distribution of the time required to revisit an interest, and exploration of new interest. In particular, we first assume that the time interval τ for an individual to forward the same meme to other individuals in the OSN system obeys the distribution $p(\tau) \sim \tau^{-\alpha}$, where $\alpha > 0$ is the algebraic scaling exponent that is effectively a parameter in our model. Note that $p(\tau)$ actually represents the bursting characteristic of human behaviors. Next, we assume that an individual has the probabilities ρ and $(1-\rho)$ to forward a new and an old meme, respectively, where ρ can be regarded as an event determination probability. The parameters α and ρ can be estimated from data.

Figure 4 displays the user behaviors extracted from the empirical data sets based on consideration of human interest dynamics. The particular quantity that we examine is the activation rate or probability, the ratio of the number of activated users at time t to the number of all users involved by this time. Figures 4(a)–4(c) show, for the five data sets, the user activation probabilities versus the normalized time L. We see that, associated with the data sets *Douban Book*, *Douban Movie* and *Douban Music*, users exhibit a similar behavior in the activation rates and the three curves can be well fit by a *sigmoid function* [Fig. 4(a)]. For the data set *Delicious*, whose activation rate curve is shown in Fig. 4(b), there exists a "step." We thus divide the curve into two parts and fit each with a sigmoid function with different parameter values. For the data set *Weibo*, as shown in Fig. 4(c), a single sigmoid function fits the activation curve well. Figure 4(d) shows the distributions of the individual interevent time for the five data sets, which are approximately algebraic with the value of the exponent α ranging from one to three, where the interevent time τ is defined as the time intervals between two consecutive actions by the same user.[14,51–57] Figures 4(e) and 4(f) display the distributions of the event determination probability ρ associated with data sets *Delicious* and *Weibo*. Note that in the data sets *Douban Book*, *Movie*, and *Music*, users tend to visit or comment on an item only once, leading to a high peak in the distribution of ρ near the unity value. Table II summarizes the parameters of the sigmoid functions in Figs. 4(a)–4(c), and the parameters α and ρ are estimated directly from the data sets. Note that, for the *Delicious* data set, there are two sets of values for the parameters B and C: one for each sigmoid function.

C. Construction of model for dynamical evolution of meme popularity

Our considerations of both microbial cell evolution and human interest dynamics in combination with the empirical observation enable a formal construction of a model to describe the dynamical evolution of meme popularity. We start from the basic model that includes meme diffusion and user activities. For a fixed group of users, at each time step, a random fraction of the users become active (enabled), and each enabled user can post memes or forward some existing ones. In accordance with the quantitative behaviors extracted from

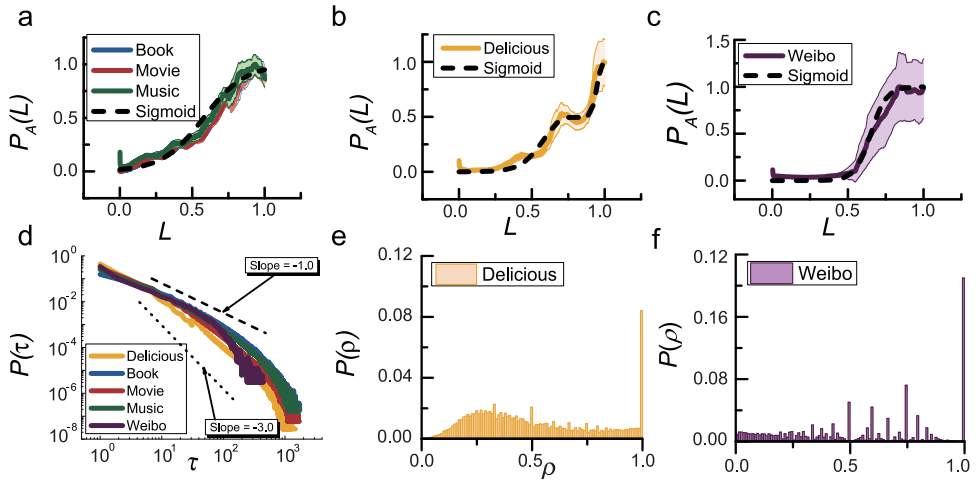

FIG. 4. Users behaviors extracted from the empirical data sets with respect to human interest dynamics. (a)–(c) Activation curves (the fraction of users activated at normalized time L) together with the error bars for the three groups of empirical data, where the yellow, blue, red, green, and purple solid lines and the corresponding shadow areas represent the average activation rates and error bars for *Delicious*, *Douban Book*, *Douban Movie*, *Douban Music*, and *Weibo*, respectively. The black dashed curves are the fitted sigmoid functions. For (a), the user activation curves are from the three data sets: *Douban Book*, *Movie*, and *Music*, which exhibit a striking agreement at a detailed level and can be fitted by a single sigmoid function. For the data set *Delicious* in (b), a plateau in the user activation curve emerges, which can be fitted by two distinct sigmoid functions. For the data set *Weibo* in (c), the activation curve can be fitted by a single sigmoid function. (d) Algebraic distribution of the individual interevent time for the five data sets, where the algebraic scaling exponent α has values ranging from one to three. (e) and (f) Distributions of parameter ρ from the data sets *Delicious* and *Weibo*, respectively. See Table II for the fitting values of all model parameters from data, which are estimated from least-squares fitting subject to the standard Kolmogorov-Smirnov (KS) test[50] with $D = 0.1$.

the empirical data sets, we assume that the fraction of active users, $P_A(L)$, follows a sigmoid function with parameters B and C:

$$P_A(L) \sim \frac{1}{1 + e^{-B\cdot(L-C)}}. \quad (2)$$

For an active user, the interevent time τ, the time interval when this user decides to forward (post) a meme, follows a power law distribution with the parameter α. The user has probability $(1-\rho)$ to forward an old meme that he/she has forwarded before and probability ρ to forward a new meme. Once a meme is created and alive on the network, its evolution is determined by the microbial cell diffusion model. Our hybrid meme diffusion model is illustrated in Fig. 5.

TABLE II. Estimated parameters for empirical data sets. Parameters associated with the sigmoid functions are B and C that are normalized through $B = \tilde{B}/\log(T)$ and $C = \tilde{C}/\log(T)$, where $\tilde{B}$ and $\tilde{C}$ are obtained from empirical data sets and T is the length of the time series. The parameters α and ρ characterize the interevent time and event determination activities.

Data sets	B	C	α	ρ
Delicious	0.24 and 0.40	0.80 and 0.80	1.73	0.54
Douban Book	0.07	0.55	1.38	1.00
Douban Movie	0.07	0.55	1.53	1.00
Douban Music	0.07	0.55	1.46	1.00
Sina Weibo	0.24	0.61	1.50	0.56

IV. RESULTS

A. Simulation results

To simulate the model, we consider a fixed group of $N_f = 1000$ users. At each time step t, $N_f \cdot P_A(t)$ randomly selected users become active. The interevent time and event determination activities of these users depend on two parameters: α and ρ. Specifically, for each activated user U_i, we generate a time series of forwarding actions based on (1) the time interval τ that a forwarding action occurs follows the distribution $p(\tau) \sim \tau^{-\alpha}$ and (2) the probability for this user to forward a new meme is ρ and that to forward an old meme is $(1-\rho)$. We then record the forwarding and exclusion probabilities, as well as the meme popularity at each time for different choices of the parameters. In the analysis of the empirical data, we truncate the F and W curves at the peak values to exclude artificial death events caused by the finite duration of the data sets, which typically occur at about 80% of the total duration.

Our hybrid model has four parameters: B and C which determine the users' sigmoid activation rate [Eq. (2)], α that regulates the interevent time intervals, and ρ that accounts for the event determination activities (probability of exploring new memes). We simulate the model for different choices of parameters to check the parameter sensitivity of the key quantities representing the model outcome: the forwarding and exclusion probabilities. We then estimate the four parameters directly from the five empirical data sets.

Figure 6 shows the simulation results of the forwarding and exclusion probabilities for different choices of the

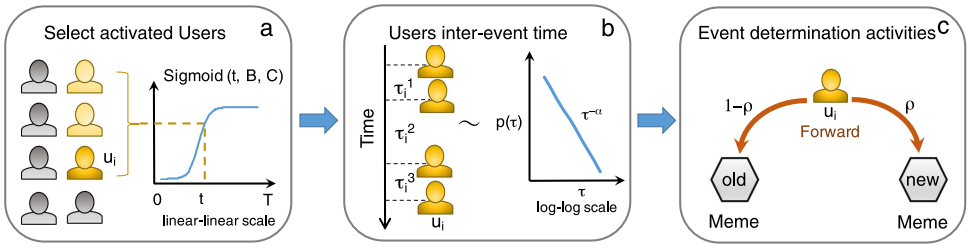

FIG. 5. A schematic illustration of the proposed biomimicry based hybrid meme diffusion model incorporating human interest dynamics and empirical observations. (a) At each time step t, a fraction of users are selected to be activated based on a *sigmoid* curve. (b) For a selected user U_i, the activities occur at different times according to the probability $p(\tau) \sim \tau^{-\alpha}$, where τ is the interval between two adjacent forwarding actions. (c) At time t, the selected user U_i has probability ρ to forward a new meme and probability $(1-\rho)$ to forward an old meme. The model is hybrid because (1) the dynamical evolution of the memes follows the rules of microbial cell diffusion (biomimicry) and (2) the probabilities $p(\tau)$ and ρ are from the human interest dynamics with key parameters extracted from actual data.

parameters α, ρ, and C. For panels in a row, we modify one parameter and compare the results, where the nominal parameter values are $B = 0.5$, $C = 0.5$, $\alpha = 1.5$, and $\rho = 0.5$, for the reasons that (1) a commonly used sigmoid function has $B = C = 0.5$, (2) users' interevent time distribution exponent α is typically in the range[14] from one to three, and (3) it is reasonable to assign equal probabilities for both old and new memes when the event determination activities are not known, leading to the choice $\rho = 0.5$. A key result of Fig. 6 is that the forwarding and exclusion probabilities are sensitive to user actions. For the sigmoid activation function, a small value of C means that the users get activated at an early time, resulting in increased forwarding and exclusion probabilities at the early time, as shown in Figs. 6(a) and 6(d). In terms of the interevent time distribution, a large value of α means that the distribution is more concentrated, indicating that many users have short interevent time. This leads to a slow growth in the forwarding probability at an early time, as shown in Fig. 6(b). While the increment is not dramatic as compared with that in the forwarding probability, there is a slight delay in the rising of the exclusion probability in early time when the value of α is large, as shown in Fig. 6(e). We note that the forwarding probability changes slightly with different values of the event determination parameter ρ. For example, a large value of ρ, which leads to a higher probability to forward new memes, results in a delayed increment in the forwarding probabilities, and the high probability of new memes indicates a small number of forwarding events per meme on the average, which causes the early increment of the exclusion probabilities. In all six panels, the amplitudes of increase in the forwarding probability with different parameters values are always larger than those of the exclusion probabilities. This is because in simulations, a meme

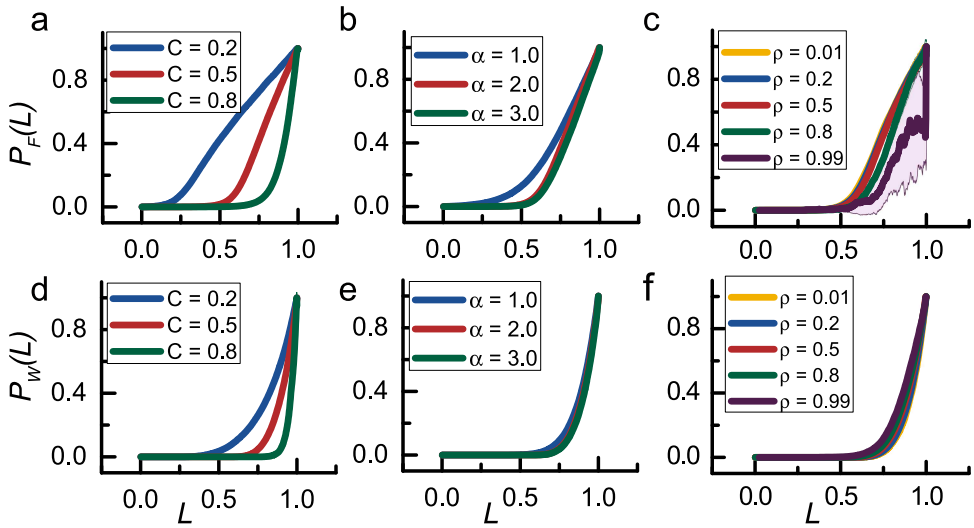

FIG. 6. Time evolution of the forwarding and exclusion probabilities generated from our hybrid model. Normalized probabilities for (a) and (d) different values of C in the users activation sigmoid function, (b) and (e) different values of α in the statistical distribution of user interevent time, and (c) and (f) different values of ρ in the user event determination probabilities. In all panels, the solid curves and the associated shadowed area indicate the average values and the error range, respectively.

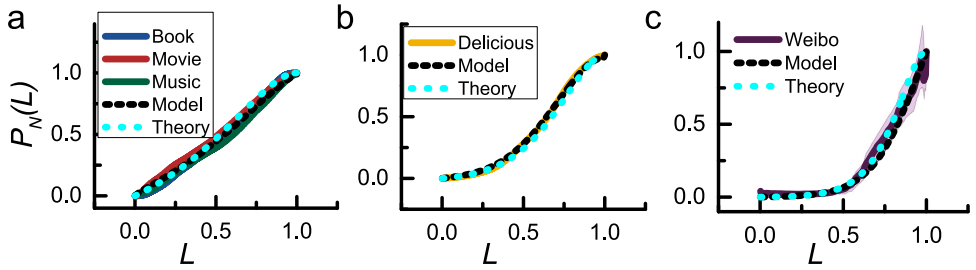

FIG. 7. Test of the predictive power of the proposed model: comparison between model predicted and real grow curves of meme popularity. Dashed and solid curves represent model prediction and the results from the real data, respectively: (a) linear growth of meme popularity for the *Douban Book, Douban Movie*, and *Douban Music* OSN systems, (b) an "S-shape" type of growth behavior from the data set *Delicious*, and (c) an approximately exponential growth curve for the data set *Weibo*. The four key model parameters B, C, α, and ρ are estimated from data. For the data set *Delicious*, the empirical fitting requires two sigmoid functions for the activation rate with different values of parameters B and C (see Table II). However, our hybrid model requires a single sigmoid function only (for $B = 0.15$ and $C = 0.5$) to generate the meme popularity growth curve that agrees well with the empirical results. The remarkable feature is that, regardless of the characteristically different growth behaviors in the meme popularity associated with diverse OSN systems, both model simulations and analytic solutions are capable of accurate prediction. Parameters used for solving Eq. (7) are $T = 1000$ and $dt = 0.01$. The model, analytic, and real results have passed the standard Kolmogorov-Smirnov (KS) test[50] with $D = 0.05$.

is regarded as being excluded only after the last recorded forwarding actions so that the influence of user actions on the overwriting action is postponed.

To validate our model, we set its four essential parameters to the values estimated from real data and compare the model predicted dynamical evolution of meme popularity with the real one. We do this for all five big OSN data sets, as shown in Fig. 7, where the growth curves of the normalized meme popularity with time $[P_N(t) \to P_N(L)]$ from model and data are displayed. The remarkable feature in Fig. 7 is that, regardless of the disparity in the nature of the data sets and regardless of the characteristic differences in the growth dynamics of meme popularity, our model predicts behaviors that agree with the actual behavior accurately at a detailed level. In particular, for the data sets *Douban Book*, *Movie*, and *Music* in Fig. 7(a), the meme popularity grows linearly with time and our model predicts this behavior precisely. For the data set *Delicious*, the meme popularity growth curve exhibits an "S-shape" feature, which is characteristically different from the linear behavior in Fig. 7(a), but the model prediction based on a single set of parameter values (B, C, α, ρ), which is in good agreement with the empirical result, captures this distinct feature unequivocally. For the growth behavior in Fig. 7(c) from the data set *Weibo*, the meme popularity exhibits an exponential behavior over time, which is predicted by our model.

We test the results from our mathematical analysis of the hybrid meme polarity growth model in Sec. IV B. Numerical solutions of Eq. (7) are shown by the cyan dashed lines in Fig. 7. For both analytic and numerical solutions, the parameters (B, C) associated with the users activation function and the event determination parameter ρ are key to determining the shape of the meme popularity growth curves. We find that fluctuations of the parameters about their "exact" values can be tolerated in our model without affecting its prediction. That is, our hybrid model not only predicts successfully the meme popularity growth dynamics but also is robust against parameter inaccuracy and uncertainties.

B. Mathematical analysis of meme popularity growth model

From Eq. (1) and Fig. 1, we see that the meme population at time t depends on the numbers of the survived and forwarded memes, which can be described by the following relation:

$$N(t) = S(t-1) + F(t-1) - W(t) + F_A(t),$$

where $F_A(t)$ is the number of newly added memes at time t. The recursive relation can be approximated by the following differential equation:

$$\frac{dN(t)}{dt} = F_A(t) - W(t). \tag{3}$$

In simulations, at each time step, we add some users into the system according to the sigmoid function. Based on natural human behavior as illustrated in Fig. 5, a subset of the newly added users contribute to new memes at the same time. The fraction of the added new memes is ρ, so the number of newly added memes is

$$F_A(t) = \rho N_i \cdot \frac{1}{1 + \exp\left[-\tilde{B}(t - \tilde{C})\right]}. \tag{4}$$

After the last forwarding event, a meme is regarded as being overwritten or excluded. Since the user interevent time intervals follow the distribution $p(\tau) \sim \tau^{-\alpha}$, it is reasonable to assume that all users' last interevent interval follows the same distribution, so does the memes' last interevent interval. The death term $W(t)$ in Eq. (3) depends on the total number of memes added into the system, which can be written as

$$W(t) = \int_0^t \frac{\rho N_i}{1 + \exp\left[-\tilde{B}(t' - \tilde{C})\right]} dt' \cdot (T - t)^{-\alpha}. \tag{5}$$

where T is the final observation time, as shown in Fig. 5(a). Substituting Eqs. (4) and (5) to Eq. (3), we obtain

$$\frac{dN(t)}{dt} = \rho N_i \cdot \frac{1}{1+\exp\left[-\tilde{B}(\hat{t}-\tilde{C})\right]} - \int_0^{\hat{t}} \frac{\rho N_i}{1+\exp\left[-\tilde{B}(t'-\tilde{C})\right]} dt' \cdot (\hat{T}-\hat{t})^{-\alpha}, \quad (6)$$

where $\hat{t}$ and $\hat{T}$ are normalized time defined as

$$\hat{t} = t \cdot \log(T)/T,$$
$$\hat{T} = \log(T).$$

The integral in Eq. (6) can be calculated analytically, leading to

$$\frac{dN(t)}{dt} = \rho N_i \cdot \frac{1}{1+\exp\left[-\tilde{B}(\hat{t}-\tilde{C})\right]} - \frac{\rho N_i}{\tilde{B}} \cdot \left(\tilde{B}\cdot\hat{t} - \log\left[1+\exp(\tilde{B}\cdot\tilde{C})\right] + \log\left\{1+\exp\left[\tilde{B}\cdot(\tilde{C}-t)\right]\right\}\right) \cdot (\hat{T}-\hat{t})^{-\alpha}. \quad (7)$$

The derivation of the analytic prediction of meme popularity growth, Eq. (7), relies on approximations such as taking the average of the growth and exclusion processes. For typical OSN systems where the meme and user numbers are large, the approximations are reasonable. Solutions of Eq. (7) for different choices of the parameters B, C, α, and ρ adopted from Table II are typically curves with both increasing and decreasing phases. To be consistent with the simulation settings, we truncate curves of the analytic solutions at the point where they start to decrease. As shown in Fig. 7, there is a good agreement among the three types results: analytic prediction, results from direct simulation of the hybrid model, and those from the empirical data.

A remarkable feature of Fig. 7 is that, associated with the three types of results there are characteristically different growth behaviors: linear, "S-shape," and exponential, depending on the choices of the parameters B, C, and ρ. To gain a qualitative understanding, we examine the effects of these parameters on the meme popularity dynamics. In particular, B and C are parameters in the activation sigmoid function, where B controls the shape of the function and C modulates the value of L when the activation rate reaches the value of 0.5 in Eq. (2). A large C value will then result in a late arrival of the half of the total meme popularity, and a large B value will make the slope of the sigmoid function steeper. A large increasing rate in the popularity can be expected in the region where the normalized time L is close to the C value. As a result, large values of B and C will delay the onset of the popularity growth to a later time, but when growth does start, it does so at a large rate. This explains the exponentially growing behavior in Fig. 7(c). For moderate values of B and C (e.g., B around 0.15 and C about 0.5), the onset of popularity growth occurs at an earlier time and reaches the maximum rate at a later time when the decrease due to exclusion and the increase due to the newly added memes are balanced, leading to an "S-shape" type of growth curves, as illustrated in Fig. 7(b). Finally, for small B and moderate C values, the slope of the sigmoid function is small, leading to an approximately linear growth behavior in the meme popularity.

Note that, the second term in Eq. (7) represents the number of excluded memes. For small $\hat{t}$, this exclusion term adds only a small modification to meme's popularity. For $\hat{t} \sim \hat{T}$, this term plays an important role in the growth behavior. The parameter ρ does not affect the shape of the growth curve and, in fact, its effect is quite straightforward: a large (small) value of ρ leads to a large (small) absolute popularity value. In a real system, for small values of ρ (e.g., 0.02), stochastic fluctuations will play a non-negligible role in the growth behavior, which can be seen through a comparison of Figs. 7(a)–7(c).

V. DISCUSSION

Recent years have witnessed an unprecedentedly rapid growth of OSN systems. These systems have become ubiquitous in the modern society with the tendency to eventually replace many of the traditional social networks. Understanding the dynamics and the dynamical processes on OSN systems is of importance to the well-being of the human society. There have been efforts in analyzing OSN systems in the past few years.[4,14–17,20,26,42,65–68] A common approach is to search for certain statistical or scaling relations through (big) data analysis and then to articulate a quantitative model to reproduce the specific scaling laws (often power-law scaling). While this approach has indeed yielded great insights into the dynamics of specific OSN systems, the inconvenient truth is that these systems are diverse and exhibit characteristically different behaviors even for a single quantity of interest.

A ubiquitous phenomenon with general interest is the growth dynamics of meme popularity in OSN systems. Our analysis of empirical data from diverse OSN systems reveals a lack of common growth behavior: depending on the system the growth dynamics can be linear, "S-shaped," or exponential. We ask the challenging question: can a single, universal model be articulated to capture and predict the characteristically different growth behaviors of meme popularity? Naturally, in order to produce distinct dynamics, the model should contain free parameters whose values depend on the specific system and should be estimated from data. In spite of this, previous phenomenological models were not generally applicable to diverse OSN systems. We are thus motivated to articulate our model based on the fundamental dynamical elements underlying the growth and spread of meme popularity. We find biomimicry to be highly inspirational, effective, and useful for achieving this goal. Specifically, we realize that the dynamics of cell evolution in microbial ecology bear similarities to the dynamics of meme popularity.

In microbial ecology, at any time a cell can experience one of the three events: division, death, or survival. Likewise, in OSN systems, at any time a meme can be forwarded, overwritten, or can survive. In both contexts, each event

is associated with a probability that is a function of time with free parameters. The close resemblance between cell and meme popularity growth dynamics provides the microscopic base for a population model of memes. Through data analysis, we find that the key probability functions have the shape of a sigmoid function, with two free parameters that can be determined from data. However, relying solely on the bio-inspired growth model is not sufficient to capture the detailed dynamical behaviors of meme popularity growth, as the OSN systems involve human behaviors and can thus be significantly more complex than just biological cell growth. For meme popularity, it is natural and reasonable to attribute the modeling element beyond cell growth to human interest dynamics, for which models have been developed recently. Incorporating the ingredient of human interest into basic cell growth, we end up with a hybrid model that contains four free parameters—all determinable from data. With the parameter values so estimated, we demonstrate that our model has the ability to predict the detailed, characteristically distinct growth dynamics of meme popularity in diverse OSN systems. To our knowledge, this is the first time that a universal model has been successfully developed to capture and predict the dynamical evolution of a broadly interested entity characteristic of modern social networking systems.

We remark that, in spite of the large meme varieties, it is still possible to construct a general model for the meme growth dynamics through the actions of forwarding (posting), staying survival, and being overwritten (exclusion). To capture the essential dynamics while making the model simple and analyzable, additional factors such as friendship reciprocity, visibility, intrinsic interaction among users, or other regular operations on past memes were not taken into account in our model. More specifically, in the real world, a user's activity is the overall result of its own intrinsic motivation and the influence of its peers.[69] Our current model does not have the ability to deal with this complicated issue mainly because not a single data set includes detailed information about these additional factors (although some data sets contain limited information about user behaviors). Moreover, one focused aspect of our model is the effects of user intrinsic activities on the meme popularity. To make this possible, we have neglected user-user interactions or the influences of peers, which are certainly important for further development of the model.

From the control point of view, our hybrid model leads to a new framework to understand the general mechanisms underlying the dynamics of OSN systems and how they may be manipulated or harnessed. Regardless of the specific features in the OSNs studied in this paper [from bookmarking shared networks (*Delicious*) to interest discovered media (*Douban*) and microblogging system (*Weibo*)], they all share the same "microscopic" dynamical elements (e.g., forwarding and exclusion) that results in certain rapid growth behavior. It is conceivable that the growth dynamics can be controlled through perturbations to these dynamical elements. This may lead to network design that, for instance, optimizes meme populations. It is also possible to optimize the robustness or resilience of information spreading in OSN systems to external attacks through modulating the forwarding and exclusion probabilities. Taken together, our work uncovers the fundamental principles underlying meme popularity in OSNs through the approach of biomimicry in combination with insights from human behavior dynamics, which may shed lights on further development of social networks and outstanding issues such as classification, robustness, optimization, and control.

ACKNOWLEDGMENTS

We thank Dr. J.-M. Huang for providing the Douban data set and Dr. X.-F. Wang for providing the Weibo data set. We would like to acknowledge the support from the Vannevar Bush Faculty Fellowship program sponsored by the Basic Research Office of the Assistant Secretary of Defense for Research and Engineering and funded by the Office of Naval Research through Grant No. N00014-16-1-2828. Z.-D.Z. was partially supported by the NSF of China under Grant No. 61603074.

APPENDIX A: RESULTS FROM EMPIRICAL DATA SETS

To gain quantitative insights into the number of occurrence for each event in the empirical data sets, we illustrate the results on the numbers of forwards and overwrites in Figs. 8 and 9, respectively.

APPENDIX B: EFFECTS OF PARAMETER VARIATIONS

As discussed in the main text, users' intra-event activities can be described by two probabilities: $\rho \cdot n^{-\beta}$ to forward an old meme and $(1 - \rho \cdot n^{-\beta})$ to forward a new meme, where n is the number of hopping events and β is an index (see user interest modeling in the main text). We find that, for different values of β, the forwarding, overwritten, and popularity rates hardly change, as shown in Fig. 10, justifying the use of a single parameter ρ to characterize users' intra-event activities.

We also examine the effects of small variations in the parameter B in the users activation Sigmoid function [Eq. (2) in main text] on the forwarding, overwritten, and popularity rates, as shown in Fig. 11. While varying the value of B can affect the rates of different events (especially the forwarding rate), the effects are negligible as compared with, e.g., those of varying the parameter C as shown in panels (a) and (d) in Fig. 5 of the main text. We thus fix $B = 0.5$ in our model analysis and simulations.

APPENDIX C: REGULARITIES IN THE EMPIRICAL DATA SETS

In population modeling of both cell growth in microbial ecology and meme popularity in social networking systems, a factor of consideration is cell or meme age.[16,20,27,41,45] In meme modeling, "age" is defined as the time duration between a meme's first appearance and death. Figure 12 shows the distribution of meme's age for all the empirical data sets studied. We see that, for data sets Douban Book, Movie, and *Music*,

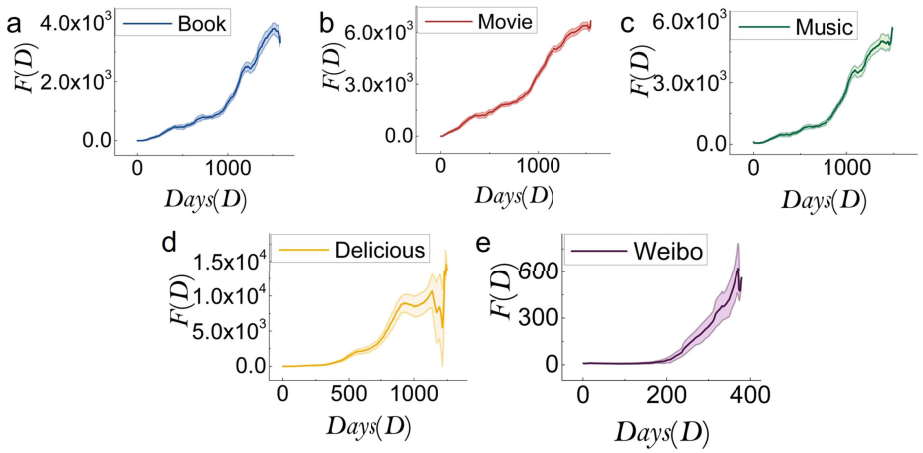

FIG. 8. Number of forwards from the five empirical data sets studied. In each panel, the line and the shadow area represent the average number of forwards and the error, respectively.

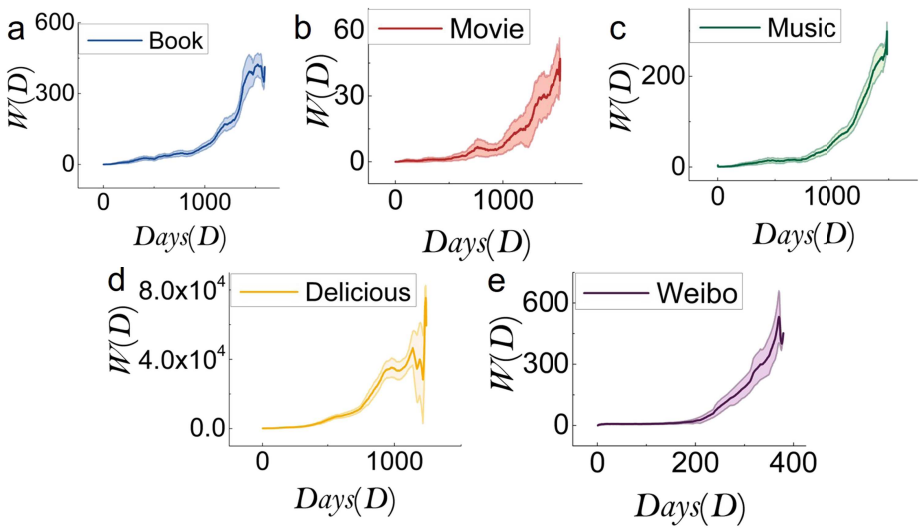

FIG. 9. Number of overwrites in the five empirical data sets studied. In each panel, the line and the shadow area represent the average number of forwards and the error, respectively.

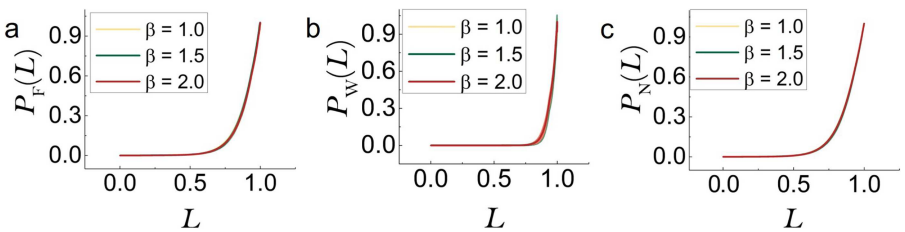

FIG. 10. Effects of the value of β on the dynamical rates. (a)–(c) The forwarding, overwritten, and popularity rates for different values of parameter β associated with modeling of users' intra-event process. Small variations of β have little effect on the rates.

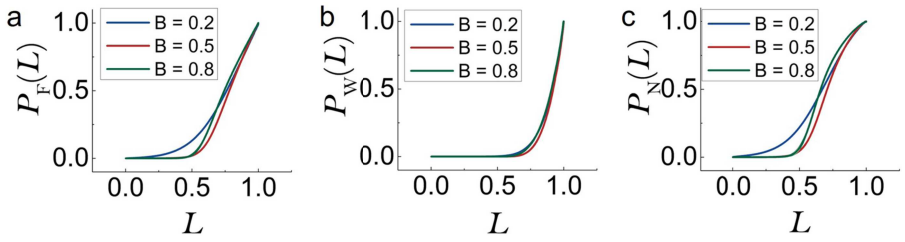

FIG. 11. Effects of the value of *B* on the dynamical rates. (a)–(c) The forwarding, overwritten, and popularity rates for different values of parameter *B* associated with modeling of users' intra-event process. Small variations of *B* have little effect on the rates.

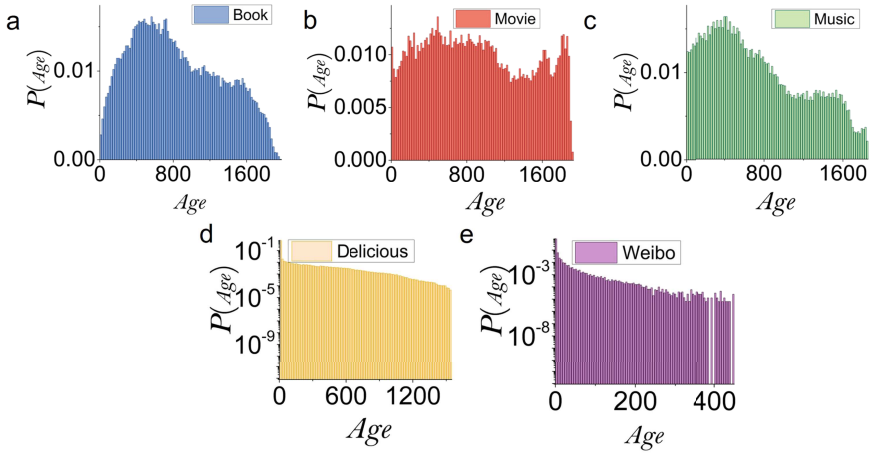

FIG. 12. Meme age distribution in the five empirical data sets studied. (a)–(c) Distributions of meme's age (in days) in *Douban Book*, *Movie*, and *Music* data sets, respectively, on a linear-linear scale. (d) and (e) Meme age distribution in *Delicious* and *Weibo* on a logarithmic-logarithmic scale.

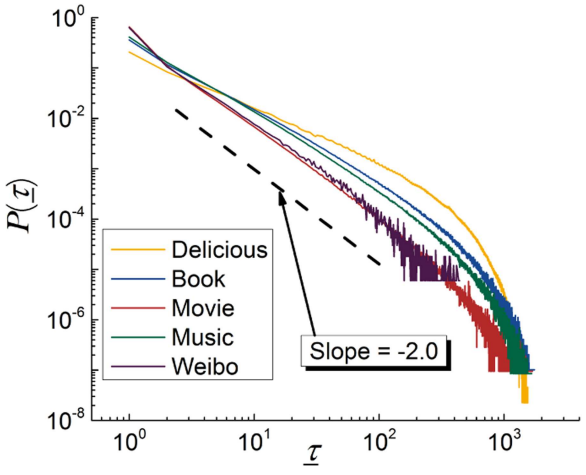

FIG. 13. Distributions of meme's forwarding intervals for the five empirical data sets studied. Yellow, blue, red, green, and purple solid curves represent the power-law distributions of meme forwarding intervals for *Delicious*, *Douban Book*, *Douban Movie*, *Douban Music*, and *Weibo*, respectively: $p(\tau) \sim \tau^{-\gamma}$.

the age distributions are relatively homogeneous, where most memes have the age of 400 to 1200 days. For *Delicious* and *Weibo*, memes have a heterogeneous age distribution, where many memes have short duration. The difference in the age distribution suggests that the corresponding OSN systems do possess different structures, but our framework is flexible and general enough to make unified modeling of popularity diffusion dynamics of different OSNs possible.

Heterogeneity in the OSN systems is also reflected in the distribution of the time interval τ for memes, as shown in Fig. 13 for the five data sets studied. We see that most memes have short time intervals (gaps) and only a few memes have large intervals.

REFERENCES

[1]P. Holme and M. E. J. Newman, "Nonequilibrium phase transition in the coevolution of networks and opinions," Phys. Rev. E **74**, 056108 (2006).
[2]J. Stehlé, A. Barrat, and G. Bianconi, "Dynamical and bursty interactions in social networks," Phys. Rev. E **81**, 035101 (2010).
[3]M.-Y. Cha, H. Haddadi, F. Benevenuto, K. Gummadi, and P. Krishna, "Measuring user influence in twitter: The million follower fallacy," in *4th International AAAI Conference on Weblogs and Social Media* (ICWSM) (2010), Vol. 10, p. 30.

[4]J.-P. Onnela and F. Reed-Tsochas, "Spontaneous emergence of social influence in online systems," Proc. Natl. Acad. Sci. U.S.A. **107**, 18375–18380 (2010).
[5]D. M. Romero, B. Meeder, and J. Kleinberg, "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), pp. 695–704.
[6]J. Yang and J. Leskovec, "Patterns of temporal variation in online media," in *Proceedings of the Fourth ACM International Conference on Web Search and Data Mining* (ACM, 2011), pp. 177–186.
[7]B. Gonçalves, N. Perra, and A. Vespignani, "Modeling users' activity on Twitter networks: Validation of Dunbar's number," PLoS One **6**, e22656 (2011).
[8]L. Weng, A. Flammini, A. Vespignani, and F. Menczer, "Competition among memes in a world with limited attention," Sci. Rep. **2**, 335 (2012).
[9]S. Goel, D. J. Watts, and D. G. Goldstein, "The structure of online diffusion networks," in *Proceedings of the 13th ACM Conference on Electronic Commerce* (ACM, 2012), pp. 623–638.
[10]J. Borondo, A. J. Morales, J. C. Losada, and R. M. Benito, "Characterizing and modeling an electoral campaign in the context of twitter: 2011 spanish presidential election as a case study," Chaos **22**, 023138 (2012).
[11]Q. Gao, F. Abel, G.-J. Houben, and Y. Yu, "A comparative study of users' microblogging behavior on Sina Weibo and Twitter," in *User Modeling, Adaptation, and Personalization* (Springer, 2012), pp. 88–101.
[12]L. Yu, S. Asur, and B. A. Huberman, "What trends in chinese social media," e-print arXiv:1107.3522 (2011).
[13]M. Oka, and T. Ikegami, "Exploring default mode and information flow on the web," PLoS One **8**, e60398 (2013).
[14]Z.-D. Zhao, Z. Yang, Z. Zhang, T. Zhou, Z.-G. Huang, and Y.-C. Lai, "Emergence of scaling in human-interest dynamics," Sci. Rep. **3**, 3472 (2013).
[15]L.-Z. Wang, Z.-G. Huang, Z.-H. Rong, X.-F. Wang, and Y.-C. Lai, "Emergence, evolution and scaling of online social networks," PLoS One **9**, e111013 (2014).
[16]J. P. Gleeson, J. A. Ward, K. P. O'Sullivan, and W. T. Lee, "Competition-induced criticality in a model of meme popularity," Phys. Rev. Lett. **112**, 048701 (2014).
[17]Y.-R. Lin, B. Keegan, D. Margolin, and D. Lazer, "Rising tides or rising stars?: Dynamics of shared attention on Twitter during media events," PLoS One **9**, e94093 (2014).
[18]J. P. Gleeson, D. Cellai, J.-P. Onnela, M. A. Porter, and F. Reed-Tsochas, "A simple generative model of collective online behavior," Proc. Natl. Acad. Sci. U.S.A. **111**, 10411–10415 (2014).
[19]K.-K. Kleineberg and M. Boguñá, "Digital ecology: Coexistence and domination among interacting networks," Sci. Rep. **5**, 10268 (2015).
[20]J. P. Gleeson, K. P. O'Sullivan, R. A. Baños, and Y. Moreno, "Effects of network structure, competition and memory time on social spreading phenomena," Phys. Rev. X **6**, 021019 (2016).
[21]Z.-K. Zhang, C. Liu, X.-X. Zhan, X. Lu, C.-X. Zhang, and Y.-C. Zhang, "Dynamics of information diffusion and its applications on complex networks," Phys. Rep. **651**, 1–34 (2016).
[22]C.-J. Wang, L. Wu, J. Zhang, and M. A. Janssen, "The collective direction of attention diffusion," Sci. Rep. **6**, 34059 (2016).
[23]X. Qiu, D. F. Oliveira, A. S. Shirazi, A. Flammini, and F. Menczer, "Limited individual attention and online virality of low-quality information," Nat. Hum. Behav. **1**, 0132 (2017).
[24]S. Luo, F. Morone, C. Sarraute, M. Travizano, and H. A. Makse, "Inferring personal economic status from social network location," Nat. Commun. **8**, 15227 (2017).
[25]J. Lehmann, B. Gonçalves, J. J. Ramasco, and C. Cattuto, "Dynamical classes of collective attention in twitter," in WWW'12—*Proceedings of the 21st International Conference on World Wide Web* (ACM, New York, 2012), pp. 251–260.
[26]K.-K. Kleineberg and M. Boguná, "Evolution of the digital society reveals balance between viral and mass media influence," Phys. Rev. X **4**, 031046 (2014).
[27]T. Zaman, E. B. Fox, E. T. Bradlow *et al.*, "A Bayesian approach for predicting the popularity of tweets," Ann. Appl. Stat. **8**, 1583–1611 (2014).
[28]J. Borge-Holthoefer, R. A. Banos, C. Gracia-Lazaro, and Y. Moreno, "Emergence of consensus as a modular-to-nested transition in communication dynamics," Sci. Rep. **7**, 41673 (2017).
[29]A. G. Hawkes, "Spectra of some self-exciting and mutually exciting point processes," Biometrika **58**, 83–90 (1971).
[30]Q. Zhao, M. A. Erdogdu, H. Y. He, A. Rajaraman, and J. Leskovec, "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* (ACM, 2015), pp. 1513–1522.
[31]R. Pastor-Satorras and A. Vespignani, "Epidemic spreading in scale-free networks," Phys. Rev. Lett. **86**, 3200 (2001).
[32]M. E. J. Newman, "Spread of epidemic disease on networks," Phys. Rev. E **66**, 016128 (2002).
[33]R. Pastor-Satorras, C. Castellano, P. Van Mieghem, and A. Vespignani, "Epidemic processes in complex networks," Rev. Mod. Phys. **87**, 925 (2015).
[34]H. von Foerster, "Some remarks on changing populations," in *The Kinetics of Cellular Proliferation*, edited by J. F. Stohlman (Grune and Stratton, 1959), pp. 382–407.
[35]E. Trucco, "Mathematical models for cellular systems the von Foerster equation. Part I," Bull. Math. Biol. **27**, 285–304 (1965).
[36]E. Trucco, "Mathematical models for cellular systems. the von Foerster equation. Part II," Bull. Math. Biol. **27**, 449–471 (1965).
[37]T. Lu, D. Volfson, L. Tsimring, and J. Hasty, "Cellular growth and division in the Gillespie algorithm," Syst. Biol. **1**, 121–128 (2004).
[38]Z. Kutalik, M. Razaz, and J. Baranyi, "Connection between stochastic and deterministic modelling of microbial growth," J. Theor. Biol. **232**, 285–299 (2005).
[39]P. Wang, L. Robert, J. Pelletier, W. L. Dang, F. Taddei, A. Wright, and S. Jun, "Robust growth of Escherichia coli," Curr. Biol. **20**, 1099–1103 (2010).
[40]J. Horowitz, M. D. Normand, M. G. Corradini, and M. Peleg, "Probabilistic model of microbial cell growth, division, and mortality," Appl. Environ. Microbiol. **76**, 230–242 (2010).
[41]E. B. Stukalin, I. Aifuwa, J. S. Kim, D. Wirtz, and S. X. Sun, "Age-dependent stochastic models for understanding population fluctuations in continuously cultured cells," J. R. Soc. Interface **10**, 20130325 (2013).
[42]B. Ribeiro, "Modeling and predicting the growth and death of membership-based websites," in *Proceedings of the 23rd International Conference on World Wide Web* (ACM, 2014), pp. 653–664.
[43]D. Antunes and A. Singh, "Quantifying gene expression variability arising from randomness in cell division times," J. Math. Biol. **71**, 437–463 (2015).
[44]C. D. Greenman and T. Chou, "Kinetic theory of age-structured stochastic birth-death processes," Phys. Rev. E **93**, 012112 (2016).
[45]T. Chou and C. D. Greenman, "A hierarchical kinetic theory of birth, death and fission in age-structured interacting populations," J. Stat. Phys. **164**, 49–76 (2016).
[46]S. Widder, R. J. Allen, T. Pfeiffer, T. P. Curtis, C. Wiuf, W. T. Sloan, O. X. Cordero, S. P. Brown, B. Momeni, W. Shou *et al.*, "Challenges in microbial ecology: building predictive understanding of community function and dynamics," ISME J. **10**, 2557 (2016).
[47]R. Wetzker, C. Zimmermann, and C. Bauckhage, "Analyzing social bookmarking systems: A delicious cookbook," in *Proceedings of the ECAI 2008 Mining Social Data Workshop* (2008), pp. 26–30.
[48]See https://en.wikipedia.org/wiki/Sina_Weibo for "Wikipedia: Sina Weibo" (2015).
[49]D. Boyd, S. Golder, and G. Lotan, "Tweet, tweet, retweet: Conversational aspects of retweeting on Twitter," in 2010 43*rd Hawaii International Conference on System Sciences (HICSS)* (IEEE, 2010), pp. 1–10.
[50]G. W. Corder and D. I. Foreman, *Nonparametric Statistics for Non-Statisticians: A Step-by-Step Approach* (John Wiley & Sons, 2009).
[51]A.-L. Barabási, "The origin of bursts and heavy tails in human dynamics," Nature **435**, 207–211 (2005).
[52]J. G. Oliveira and A.-L. Barabási, "Human dynamics: Darwin and Einstein correspondence patterns," Nature **437**, 1251–1251 (2005).

[53]Z. Dezsö, E. Almaas, A. Lukács, B. Rácz, I. Szakadát, and A.-L. Barabási, "Dynamics of information access on the web," Phys. Rev. E **73**, 066132 (2006).
[54]T. Zhou, H. A.-T. Kiet, B. J. Kim, B.-H. Wang, and P. Holme, "Role of activity in human dynamics," Europhys. Lett. **82**, 28002 (2008).
[55]B. Gonçalves and J. J. Ramasco, "Human dynamics revealed through Web analytics," Phys. Rev. E **78**, 026123 (2008).
[56]C. Song, T. Koren, P. Wang, and A.-L. Barabási, "Modeling the scaling properties of human mobility," Nat. Phys. **6**, 818–823 (2010).
[57]M. Szell, R. Sinatra, G. Petri, S. Thurner, and V. Latora, "Understanding mobility in a social petri dish," Sci. Rep. **2**, 457 (2012).
[58]F. Wu and B. A. Huberman, "Novelty and collective attention," Proc. Natl. Acad. Sci. U.S.A. **104**, 17599–17601 (2007).
[59]M. Ye, T. Sandholm, C. Wang, C. Aperjis, and B. A. Huberman, "Collective attention and the dynamics of group deals," in *Proceedings of the 21st International Conference on World Wide Web* (ACM, 2012), pp. 1205–1212.
[60]S. Brin and L. Page, "The anatomy of a large-scale hypertextual Web search engine," Comp. Netw. ISDN Syst. **30**, 107–117 (1998).
[61]N. Craswell and M. Szummer, "Random walks on the click graph," in *Proceedings of the 30th Annual International ACM SIGIR CRDIR* (ACM, 2007), pp. 239–246.
[62]R. Fagin, A. R. Karlin, J. Kleinberg, P. Raghavan, S. Rajagopalan, R. Rubinfeld, M. Sudan, and A. Tomkins, "Random walks with "back buttons"," Ann. Appl. Prob. **11**, 810–862 (2001).
[63]M. R. Meiss, B. Gonçalves, J. J. Ramasco, A. Flammini, and F. Menczer, "Agents, bookmarks and clicks: A topical model of web navigation," in *Proceedings of the 21st ACM CHH* (ACM, 2010), pp. 229–234.
[64]F. Chierichetti, R. Kumar, P. Raghavan, and T. Sarlós, "Are Web users really Markovian?" in *Proceedings of the 21st International Conference on World Wide Web* (ACM, 2012), pp. 609–618.
[65]L. Tang, X. Wang, and H. Liu, "Community detection via heterogeneous interaction analysis," Data Min. Knowl. Discov. **25**, 1–33 (2012).
[66]T. Kuhn, M. Perc, and D. Helbing, "Inheritance patterns in citation networks reveal scientific memes," Phys. Rev. X **4**, 041036 (2014).
[67]C. Sanlı and R. Lambiotte, "Local variation of hashtag spike trains and popularity in Twitter," PLoS One **10**, e0131704 (2015).
[68]H. Zhu, X. Yin, J. Ma, and W. Hu, "Identifying the main paths of information diffusion in online social networks," Physica A **452**, 320–328 (2016).
[69]P. Piedrahita, J. Borge-Holthoefer, Y. Moreno, and S. Gonzalez-Bailon, "The contagion effects of repeated activation in social networks," Soc. Netw. **54**, 326–335 (2018).