OPEN

Message spreading in networks with stickiness and persistence: Large clustering does not always facilitate large-scale diffusion

Pengbi Cui[1,2], Ming Tang[3] & Zhi-Xi Wu[1,2]

[1]Institute of Computational Physics and Complex Systems, Lanzhou University, Lanzhou Gansu 730000, China, [2]Key Laboratory for Magnetism and Magnetic Materials of the Ministry of Education, Lanzhou University, Lanzhou Gansu 730000, China, [3]Web Sciences Center, University of Electronic Science and Technology of China, Chengdu 610054, China.

Recent empirical studies have confirmed the key roles of complex contagion mechanisms such as memory, social reinforcement, and decay effects in information diffusion and behavior spreading. Inspired by this fact, we here propose a new agent–based model to capture the whole picture of the joint action of the three mechanisms in information spreading, by quantifying the complex contagion mechanisms as stickiness and persistence, and carry out extensive simulations of the model on various networks. By numerical simulations as well as theoretical analysis, we find that the stickiness of the message determines the critical dynamics of message diffusion on tree-like networks, whereas the persistence plays a decisive role on dense regular lattices. In either network, the greater persistence can effectively make the message more invasive. Of particular interest is that our research results renew our previous knowledge that messages can spread broader in networks with large clustering, which turns out to be only true when they can inform a non-zero fraction of the population in the limit of large system size.

O ver the last few years, many empirical works[1–7] or practical model[8,9] have identified the strong relevance of complex contagion mechanisms such as memory effect, social reinforcement and decay effects to information diffusion or behavior spreading. On account of memory effect, the previous contact activities can affect the current spreading process[4,10]. Specifically, individual's selection of message items can be naturally expedited by the increasing frequencies of the same choices of other people if they find the items interesting or crucial enough[3,8,11]. This is usually interpreted as the results of social reinforcement[9,12,13]. On the other hand, there are an increasing amount of new messages an individual is facing every day in modern real life, whereas the attention and processing abilities of people are finite and saturated[4,11,14]. The novelty of a message usually trend to fade with time and hence the attention people pay to it, which is normally described as decay effects[1,4,11,15]. It is shown that the social reinforcement effect could be weakened or even counterbalanced by decay effects[4,5,11].

Although the competition between social reinforcement and decay effects has been emphasized and used as a guideline to measure the natural time scale that attention fades away[11], to our best knowledge few works have been attempted to model explicitly the competition and memory effect, and study deeply how it shapes the spreading of information on complex networks. Here we want to point out that the three mentioned effects in information spreading are quite different from those have been considered in the studies on Naming Game (NG) and Category Game (CG), since either NG or CG is a two-step multi-state negotiation process[16–19], whereas information spreading is not. First, herein memory effect performs as the storing of the times of contact of people with recipients of information[3,8], rather than the possible words (or names) for the object (or a category) in NG (or CG)[16–19]. Second, decay effect in information spreading reflects the decay of people's interest or attention in a message owing to the competition with other news or stories[4,11], contrary to the NG (CG) in which it means the decrease of the number of different words used in the system (or average number of words per category)[16–19]. Third, unlike the phenomenon that an hearing would have more opportunities to add (or remain) one word only if more selected speakers try to transmit the same one to it[16–19], the reinforcement effect in information diffusion
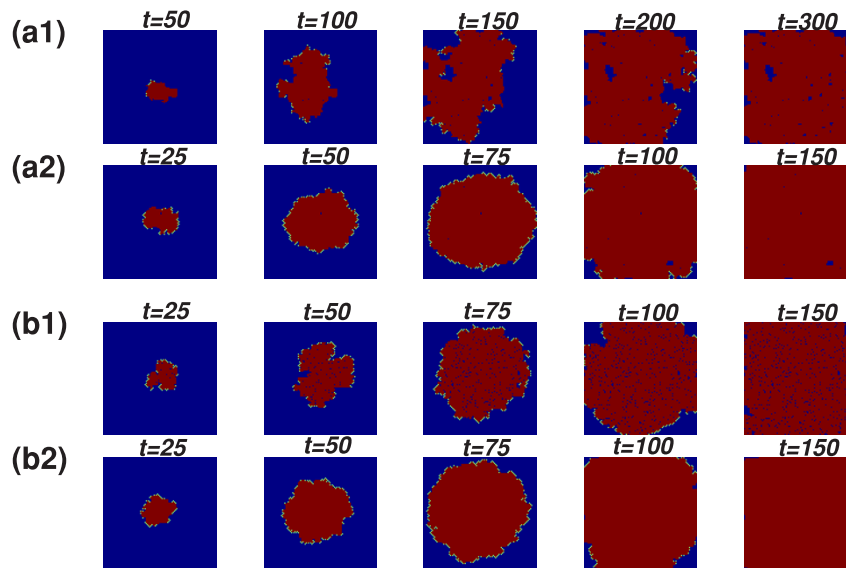
Figure 1 | The time evolution of spatial patterns, for two different values of stickiness (a1) $a = 0.40$, (a2) $a = 0.45$ (bottom panel) where $b = 0$; and for two values of persistence (b1) $b = -1.00$, (b2) $b = 1.00$ (bottom panel) where $a = 0.45$. Red sites represent recovered or alerted nodes, bright green sites represent infected ones, and blue sites denote susceptible nodes. Other parameters are chosen as $n_s = 2$ and $N = 101 \times 101$.

indicates the more simple situation that the more neighbors adopting the message, the higher likelihood an individual following them[3,5].

Next, the big challenge we are confronted with is the possibility of modeling and studying the message spreading along with both social reinforcement and decay effects based on the memory effect. Recent researches[6] have shown that the variation in the ways that different information spread is attributed to not only the stickiness – the probability of information adoption is mainly dominated by the first few exposures[6,7], but also the persistence – the relative extent to which more repeated exposures to the message continue to have durative effect. Similar results especially the exposure response behaviors were also confirmed by a lot of empirical studies[4,5,7]. The two mechanisms, stickiness and persistence, thus enable us to quantitatively study the joint action of the three effects together.

At the same time, the structures of complex social systems can be characterized by complex networks, on which many spreading activities may take place, ranging from the spreading of epidemics[20–23], the diffusion of behaviors and news[3,8], to the promotion of technique innovations[24], etc. Consequently, motivated by the empirical studies[3–8,11] mentioned above, we propose a new agent-based model offering an opportunity to explore the impact of social reinforcement and decaying effects quantified by stickiness and persistence on the message (information) diffusion on various networks. In the presence of strong decay effects, we find that a message is more likely to outbreak (i.e., it can reach a non-zero fraction of the population in the thermodynamic limit) on the tree-like networks such as scale-free (SF) networks and Erdős-Rényi (ER) random networks rather than on the regular lattices (RLs). Specifically, a message can spread broader in the RLs than that in the tree-like networks only if it can outbreak. The critical behaviors of the diffusion process can be reasonably estimated by the bond-percolation theory considering spatial correlations of the underlying networks through which message diffuses. In addition, we develop a verification approximation, whose solutions confirm well the non-negligible role of the dynamical correlations between transmission events in the RLs.

Results

Here, we first carry out extensive simulations for the agent-based model of message diffusion on square lattice. We then compare the simulation results with the predictions from the analytical bond percolation theory and verification method involving time correla-

tions of the spreading events. Finally we extend our model and analytical methods to other networks such as RLs, SF networks, and ER networks to validate the robustness of our findings.

Message diffusion on square lattice. We first consider the message diffusion on a square lattice of size $N = L \times L$ with periodic boundary conditions. The message starts spreading from the center node (selected as the seed), while all the others are in the susceptible state (i.e., they hear nothing about the message).

To intuitively grasp the roles of stickiness and persistence, we begin by presenting the time evolution of spatial patterns of message spreading in Fig. 1. The message with stickiness $a = 0.40$ (see the Methods for the precise definitions of a, b and other parameters) spreads in an irregular manner (see Fig. 1(a1)). By comparison, the message with a slightly stronger stickiness $a = 0.45$ diffuses outward to susceptible areas in a quasi-circular manner with a broader rim of informed (infected) individuals (see Fig. 1(a2)). This indicates that messages with different strength of stickiness could give rise quite different spreading patterns and behaviors. Figs. 1(b1) (b2) and Supplementary Fig. S1 show that the persistence b also affects considerably the whole spreading size, by governing the number of the isolated susceptible islands (blue domains surrounded by red areas, the emergence of these islands arises from the fact that continually increasing number of infected neighbors fail to infect those individuals owing to small persistence). The above arguments suggest that the stickiness and persistence have great but different influences on the spreading of message.

We next explore the critical behaviors of message diffusion for various stickiness and persistence, by providing a quantitative assessment of the message burstiness. By means of the theory of non-equilibrium phase transition in statistical physics which has been successfully generalized to study the epidemic dynamics, previous studies[25,26] on spreading dynamics have shown that the fluctuation of the order parameter is divergent at the critical point. We thus use the procedure proposed in[26] to numerically determine the critical areas. To be more specific, a series of variabilities $v(a, b)$ are firstly obtained as

$$v(a,b) = \frac{\sqrt{<\rho_R^2(a,b)> - <\rho_R(a,b)>^2}}{<\rho_R(a,b)>}, \qquad (1)$$
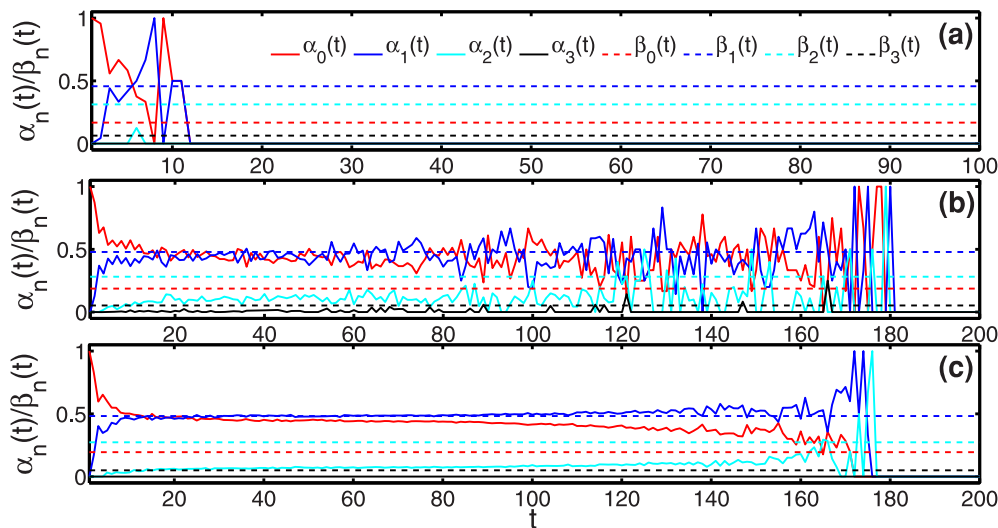
Figure 2 | The evolution of proportions of the transmission events. The parameters are chosen at (a) subcritical point $a = 0.20$, $b = 0.20$; (b) critical point $a = 0.35$, $b = 0.20$; and (c) supercritical point $a = 0.43$, $b = 0.20$.

where ρ_R and $v(a, b)$ denote, respectively, the density of recovered individuals in the population and the relative standard deviation (RSD) at parameter point (a, b). There exists a maximum variability $v_{max}(b)$ for each value of b when varying a from 0 to 0.5, and the values of $v_{max}(b)$ with $b \in [-1, 1]$ can be used as the numerical estimation of the threshold position.

Although a bond-percolation process can be mapped to the SIR model[27,28], its extension to the current model is not straightforward. First, in our model, the transmission probability that a susceptible individual approves the message varies with the times he has received it from his infected neighbors (i.e., the number of informed neighbor he has had). Second, the time correlations between different transmission events E_n[27] (the transmission event $E_n(t)$ represents that an individual, who has received the message at least once before, successfully approves the message when he has received it another n times until time t). To confirm the existence of the time correlation, in Fig. 2 we compare the dynamic proportions of the four transmission events ($\alpha_n(a,b,t) = \dfrac{\omega_n(a,b,t)}{\sum_{m=0}^{\langle k \rangle - 1} \omega_m(a,b,t)}$) obtained from numerical simulations with those predicted by the bond percolation considering spatial correlations ($\beta_n(a,b,t) = \dfrac{q_n T_n(a,b)}{\sum_{m=0}^{\langle k \rangle - 1} q_m T_m(a,b)}$) for three parameter points (a, b). Here, $\omega_n(a, b, t)$ represents the occurrence frequency of $E_n(t)$ obtained from numerical simulations, and $\langle k \rangle$ denotes the average degree of the networks. In Figs. 2 (b) and (c), we see that $\alpha_0(t)$ ($\alpha_2(t)$) is greater (less) than $\beta_0(t)$ ($\beta_2(t)$) during the spreading process, whereas $\alpha_1(t)$ ($\alpha_3(t)$) is equal or close to $\beta_1(t)$ ($\beta_3(t)$). The reason is that the existence of time correlations of the transmission events $E_n(t)$ can determine whether the subsequent events $E_m(t)$ ($m > n$) happen or not in the spreading process. If $E_n(t)$ does not happen, the events $E_m(t)$ ($m > n$) will probably happen; otherwise $E_m(t)$ ($m > n$) will never happen since an informed individual transmits the message only one time and then becomes recovered (i.e., completely ignores the message) forever. Consequently, E_0 (E_2) contributes more (less) than predicted by percolation theory to the spreading course (also see Supplementary Fig. S2). To overcome the challenge, we develop a verification approximation involving the time correlations of the transmission events, besides the spatial correlations originating from the spatial structure of the lattice[27]. It is necessary to mention that the discrete-time synchronous transmission of the message enables us to avoid

concerning about the additional synergistic effect[29]. Moreover, Figs. 2(b) and (c) show that $\alpha_n(a, b, t)$ is dynamically stable, which shows that this correlation always exists. It allows us to adopt average values of the four indices $\alpha_n(a, b, t)$ at the critical regions for the verification approximation.

Based on the proposed methods in the Methods section and Eq. (1), we yield both the analytical prediction and the verification threshold for $n_s = 2$, plus the numerical results for various n_s, as depicted in Fig. 3. We note that the numerical thresholds stay at $a \approx 0.32$, which is mainly determined by the stickiness of message (i.e., the parameter a), regardless of the values of b and n_s. This means that most informed individuals are actually infected by their first one or two infected neighbors (also see Supplementary Figs. S2–S6). Furthermore, the numerical estimations are fairly reproduced by the bond-percolation theory. Comparing the analytical boundary, the verification approximation involving both spatial and time correlations gives a higher accurate estimation than the bond-percolation method considering only spatial correlations (the dashed black
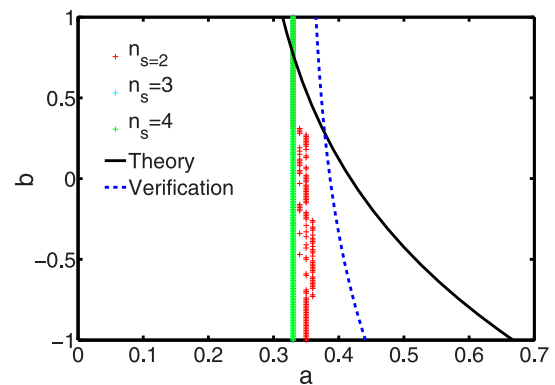
Figure 3 | The phase diagram of the message spreading. The real numerical critical boundaries (crosses) are obtained by Eq. (1) for various n_s on square lattice of size $N = 101 \times 101$. For comparison, analytical boundary (black line) and verification boundary (dash black line) for $n_s = 2$ are also shown (the cases for $n_s = 3$, 4, 5 do not allow us to analytically identify the critical lines). Herein, we select a narrow parameter ranges $b \in [-1, 1]$ and $a \in [0.32, 0.36]$ containing the numerical critical boundary for the calculation of verification threshold (see more detailed method in the Methods section).
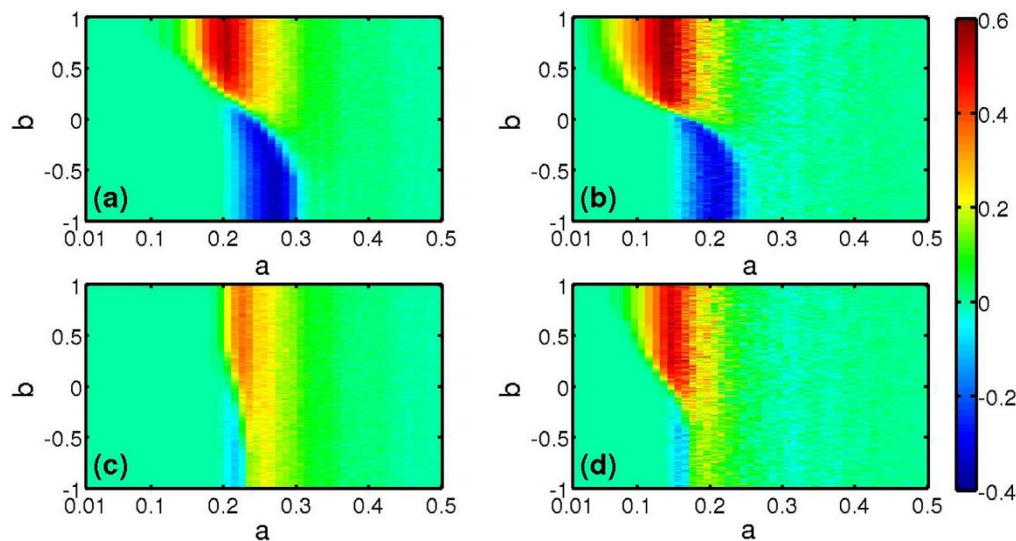
Figure 4 | The differences of spreading sizes between RLs and RRNs networks ($\rho_{RL} - \rho_{RRN}$). (a,b) $n_s = 2$, (c,d) $n_s = 3$. The average degrees of the networks are $\langle k \rangle = 6$ (a,c), and $\langle k \rangle = 8$ (b,d), respectively. ρ_{RL} (ρ_{RRN}) represents the recovered density of population (i.e., spreading size) on RLs (RRNs).

line is clearly closer to the numeric markers than the black solid line is).

Message diffusion on regular lattice networks and regular random networks. Centola's work[3] concludes that social behaviors can spread farther and faster across clustered-lattice networks than across corresponding regular random networks (RRNs), owing to the strong social reinforcement induced by clustered ties. RRNs are networks that all nodes have exactly the same degree while links are randomly distributed among nodes, avoiding self-connections and multiple connections. To check whether the findings by Centola are still fulfilled for information diffusion, we further investigate our model defined on the RLs (Hexagonal network and Moore network) and the RRNs.

To get a comparison, we present in Fig. 4 the differences of the final size of recovered population on the two networks with the same average degree. The blue areas characterize the parameter regions where the conclusion of Centola's experiment (that the information spreads farther across the RLs than across the corresponding RRNs) does not hold. The violation is attributed to the presence of strong decay effects ($b < 0$) in the vicinity of the critical regions. Specifically, the outbreaks of message can happen more easily in the RRNs than that in the RLs for negative persistence owing to that strong decay effects outcompetes the weak reinforcement effect (also see Supplementary Fig. S7, strong reinforcement effect (decay effects) is reflected by large stickiness and/or positive persistence (negative persistence) in our model). As a and b get larger, things turn out differently, the message is able to seize a larger population in the RLs, which is accordant with the anticipation of Centola's experiment. This means that high level of clustering created by redundant ties that linked each node's neighbors to one another in the RLs strengthens the reinforcement effect, and hence facilitates the diffusion of the message[3]. Moreover, larger n_s improves the performance of stickiness in facilitating the message diffusion, making social reinforcement effects be the most prominent for the RLs. Consequently, the blue areas shrink with increasing n_s. Thus, the above differences investigated indicate that network topology and n_s (which can be regarded as one of the intrinsic characteristics of the message) simultaneously determine the effects of stickiness and persistence on the spreading dynamics.

Using Eq. (1), we obtain the numerical thresholds for various values of n_s. According to the methods described in the Methods section, we can yield the analytical (theoretical) thresholds of mess-

age diffusion on both the RRNs and the RLs, in addition to the verification thresholds on the RLs for $n_s = 2$ (Supplementary Fig. S10 and Fig. 11 show that the time correlations between the transmission events in Hexagonal lattice and Moore lattice are noticeable). In the case of RRNs (Figs. 5(a) and (b)), we observe that the positions of the critical boundaries are mainly determined by both the stickiness and $\langle k \rangle$ instead of the persistence and n_s, on account of weaker social reinforcement[3,27] resulting from the low clustering coefficient. Unlike the case of RRNs, the theoretical analysis by means of bond-percolation theory gives rather accurate predications of the position of the threshold with strong persistence, but not for negative persistence (i.e, with the presence of strong decay effects) on the RLs. When the underlying networks for message diffusion are Hexagonal lattice and Moore lattice, more available edges for message spreading will further strengthen the role of persistence in message diffusion, especially with stronger social reinforcement (positive b). That enables all transmission events involve in the diffusion (see Supplementary Figs. S10–S13), so that $\beta_n(a, b, t)$ gets close to $\alpha_n(a, b, t)$ (see Supplementary Figs. S10 and S11). On the other hand, the trajectories obtained by verification approximation are in good agreement with the simulations, from which one can conclude that the effect of time correlations of transmission events is indeed general on the RLs. Additionally, the message steps forward to arrive in half of the neighbors of the same host on the RLs by flowing through almost $\dfrac{\langle k \rangle}{2}$ edges connecting it. This makes the message with positive persistence ($b > 0$) outbreak more easily in the presence of denser local connections, and the persistence thus imposes a greater influence on outbreaks of message on RLs with larger average degree for small n_s ($\dfrac{n_s}{\langle k \rangle} < \dfrac{1}{2}$). The message also has reached a saturation state when the subsequent events $E_i \left(i > \dfrac{\langle k \rangle}{2} \right)$ happen (see Supplementary Figs. S14–S21). Also in RLs, higher n_s limits the effect of persistence, and the phase transitions are determined by not only the topologies of the networks but also the stickiness and persistence of the message.

In addition, the reinforcement effect begins to work as the message is bursting and prevailing on both the RLs and RRNs. Therefore, the results for positive persistence near the thresholds (see Supplementary Figs. S12–S24) elucidate the actual phenomenon "Three men make a tiger" (or "A lie, if repeated often enough, will be accepted as truth"), where the majority do not believe the message

4
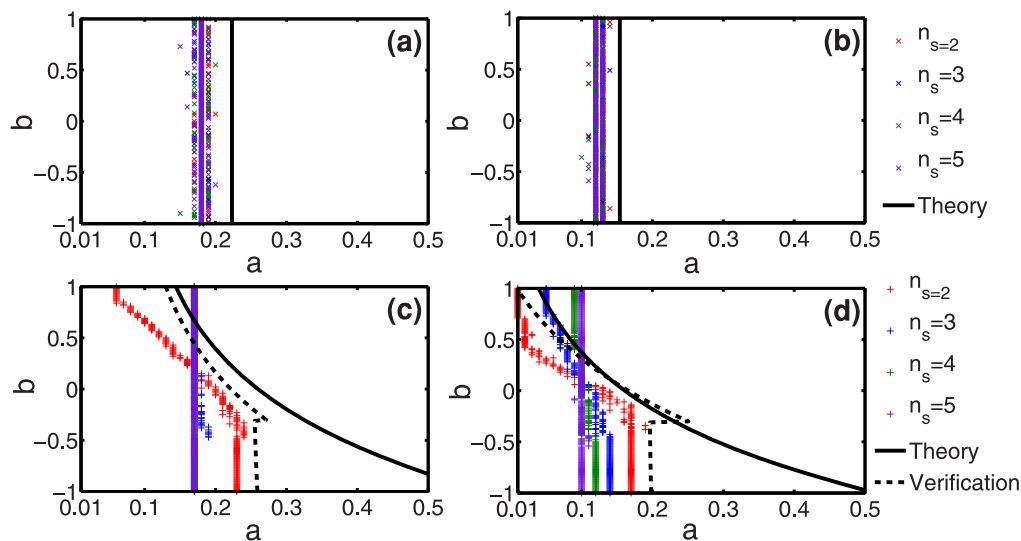
Figure 5 | The phase diagrams of the message spreading on RLs and RRNs. The underlying networks are (a) RRN with $\langle k \rangle = 6$, (b) RRN with $\langle k \rangle = 8$, (c) Hexagonal lattice, and (d) Moore lattice. The predictions from the bond-percolation theory (solid lines) and verification approximations (dashed lines) for $n_s = 2$ are illustrated to compare with the simulation data (markers) for various n_s. To get the verification trajectory on Hexagonal lattice (and Moore lattice), we select two parameter regions containing the numerical thresholds (see more detailed method in the Methods section). One region is $b \in [-0.3, 1.0]$ and $a \in [0.06, 0.24]$; the other one is $b \in [-0.3, 1.0]$ and $a \in [0.01, 0.19]$ in Hexagonal lattice. One region is $b \in [-1.0, -0.3]$ and $a \in [0.23, 0.24]$; the other one is $b \in [-1.0, 0.3]$ and $a \in [0.15, 0.18]$ in Moore lattice.

as a truth until at least three neighbors have tried to transmit this message to them.

Message diffusion on SF networks and ER networks. To further check the robustness of our above findings, we finally investigate our message diffusion model on SF and ER networks with $\langle k \rangle = 6, 8, 10, 12$. Since two or more transmission events fail to last over the long time at the critical points (see Supplementary Fig. S24), we do not take the time correlations into consideration in theoretical analysis. Compared with the simulation data, the analytical results for SF networks and ER networks with different average degrees show that the theoretical approaches are already sufficient to give fairly precise expressions of the outbreaks of message completely determined by the stickiness a (see Supplementary Fig. S25). Nevertheless, the persistence also partly boosts its impact on the size of message diffusion as $\langle k \rangle$ gets larger (see Supplementary Fig. S26).

Fig. 6 displays evident differences of the final spreading sizes on the RLs and SF (ER) networks with the same degree. The spreading sizes are larger in SF (ER) than those in the RLs at the parameter regions where the message has already outbroken on SF (ER) but not yet on the RLs, owing to the hubs and shorter shortest paths in SF (ER)[30] (see Supplementary Fig. S25 for the critical boundaries of message on SF and ER). Just for strong persistence, denser local connectivity of the RLs can make an invasion take place more easily. As a and b increase beyond the critical boundaries displayed in Fig. 5, the message can capture a larger population again on the RLs, despite of the existence of hubs the short characteristic path length in the SF (ER) networks. The reason is that very smaller clustering coefficient gives rise to weak social reinforcement effect[3,31], which again leads to weak performance of persistence in promoting the spread of message on the SF (ER) networks. Our results indicate that the role of reinforcement effect is more important than that of hubs or shortest paths in facilitating message spreading only when the message outbreaks on the RLs. Otherwise, especially for the presence of decay effects (negative persistence) the facilitation of hubs and shortest paths to the diffusion on SF (ER) cannot be neglected. In addition, the results for positive persistence on SF and ER networks can also be

treated as evidence of the mechanism "Three men make a tiger" (see Supplementary Fig. S27 and Fig. S28).

Discussion

In conclusion, taking into account social reinforcement and decay effects based on memory effect in reality, we have proposed a new agent-based message spreading model with stickiness and persistence, and carried out extensive computer simulations of our model on various types of networks. By means of the relative standard deviation method and the bond percolation theory involving the spatial correlations, we are able to determine numerically and analytically the positions of critical boundaries. Moreover, the remarkable accuracy of verification approximation involving the time correlations between different transmission events validates the wide existence of such correlations for the message diffusion on regular lattices.

Our preliminary results show that in RLs, the persistence depends greatly on the position of inflection point n_s and average degree $\langle k \rangle$ of the underlying networks, and begins to play a pivotal role in the spreading process with increasing $\langle k \rangle$ owing to the emergent large clustering coefficient[3]. Stronger social reinforcement arising from larger clustering coefficient leads readily to stronger infectivity of the message, which can invade a great number of susceptible individuals in the RLs, confirming the conclusion of Centola's work. By comparison, in tree-like networks such as RRNs, SF and ER, the critical thresholds of message diffusion are only dominated by the stickiness, and both the hubs and the short characteristic path length facilitate the outbreaks of message in the presence of decay effects. It worth emphasizing that the results presented in this paper has successfully substantiated the phenomenon "Three men make a tiger" (or "A lie, if repeated often enough, will be accepted as truth").

Placing our study in the context of social media, the hub nodes actually play a role of broadcasters, advertisements and so on, which are very important for the large scale spreading of information. However, the intrinsic contents of messages and their adaptation to hosts are extremely relevant in determining the message diffusion. In particular, the results for the message diffusion on the SF and ER networks constitute a proof that exposure to mass media can favour
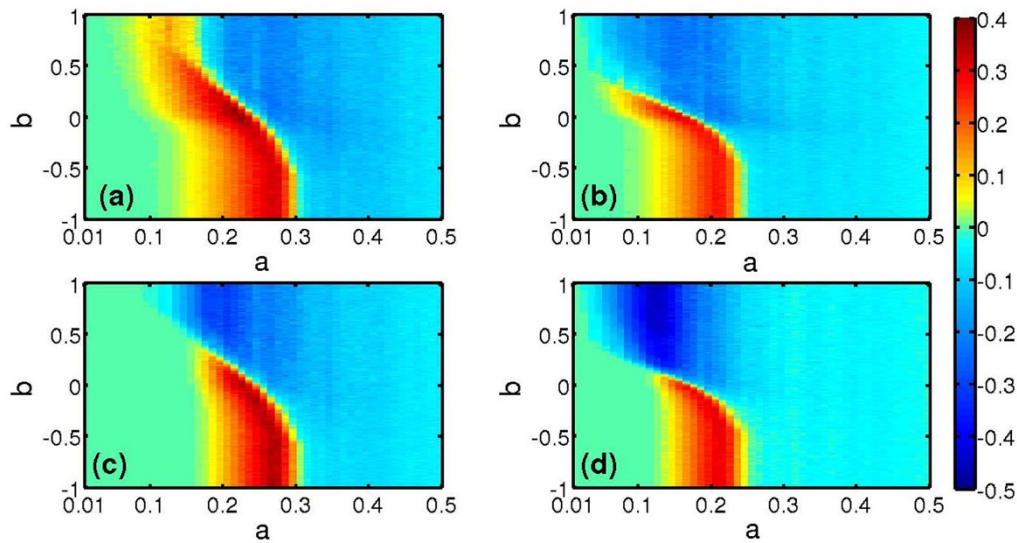
Figure 6 | The differences of spreading sizes between the RLs and SF (ER) networks ($\rho_{SF/ER} - \rho_{RL}$). (a,b) the difference $\rho_{SF} - \rho_{RL}$, (c,d) the difference $\rho_{ER} - \rho_{RL}$. (a,c) The average degrees of the networks are $\langle k \rangle = 6$ (a,c), and $\langle k \rangle = 8$ (b,d), respectively. Here, ρ_{SF} (ρ_{ER}, ρ_{RL}) represents the recovered density in SF networks (ER networks, RLs). Other parameter: $n_s = 2$.

the outbreaks of behaviors, news or messages, despite the decay effects, only if the stickiness of the message is large enough. On the other hand, the RLs are efficient in taking advantage of social reinforcement effects to promote the global spreading of the message, owning to more cluster ties (or local pressures) that function as form of 'initial groups' or 'small groups' through interpersonal communication or 'machine-interpersonal communication'[32]. It is also verified by our research that the local, personal communication is irreplaceable to lead to propagation of message despite of the developed media industry today, from the perspective of communication.

Recent studies have started considering the memory effect[3,8], social reinforcement and decay effects[4] in information spreading. The mechanisms were yet investigated in isolation. Our work is the first attempt to account explicitly for the three key mechanisms together, and to evaluate the joint action of them by quantifying their effects as stickiness and persistence. It provides a quantitative guideline for future social experiments for message spreading.

In reality, the ways in which information spread may be very complicated. In the present study, we do not capture the difference of individual-level preference[33] that might have also influenced their decisions to adopt one message. For example, in an online social network such as Twitter, individuals may prefer different hashtags, and significant variations in the ways that the hashtags on different topics spread were observed[7]. In addition, one need to concern about the diverse cultural and societal backgrounds[34] which would lead to the different styles by which individuals contact with the medium, or even hamper communications among different groups of members[32]. Moreover, the volatilities of complex contagion of controversial topics, psychological status of individuals[35] reveal that the status of the individuals in the communication systems are time dependent, which should be addressed in future research.

Methods

Message spreading model with stickiness and persistence. The message spreading model is implemented on a network consisting of N nodes and E edges, where the nodes represent the individuals in a population and the edges the social interactions among them, through which information propagates. Each individual is allowed to be in one of three states at each time step: (i) Susceptible (or uninformed) state—the individual has not yet heard the message or is aware of the news but not willing to transmit it. (ii) Infected state—the individual catches the message and forwards it to all his nearest neighbors. (iii) Recovered state—the individual will never transmit the message any more after having transmitted it once before.

Specifically, the information propagation is modeled in a probabilistic framework at individual level[8,9]. A susceptible individual i will adopt the message with a probability $\lambda_{n^\circ}(a, b)$, given that he has heard it from his informed neighbors n° times (i.e., n° informed neighbors he has had), plus the first time. In detail, $n^\circ = n + 1$ when individual i has owned at least one informed neighbors, otherwise $n^\circ = 0$. Here $\lambda_{n^\circ}(a, b)$ is a linear piecewise function of the times he has received message from his informed neighbors:

$$\lambda_{n^o}(a,b) = \begin{cases} \min\{an^o, 1\}, & 0 \le n^o \le n_s; \\ \min\{bn^o + n_s(a-b), 1\}, & n_s < n^o \le k_i; \end{cases}$$

$$\text{and} \quad \lambda_{n^o}(a,b) = 0, \quad \text{if} \quad \lambda_{n^o}(a,b) < 0; \tag{2}$$

where k_i is the degree of node i. n_s is the inflection point beyond which persistence is the dominant factor for the infection probability of message, and the parameters a (stickiness) and b (persistence) characterize how $\lambda_{n^\circ}(a, b)$ change with n°, as illustrated in Fig. 7. Since empirical data[3,6,7] have shown that social reinforcement sets in such that initial exposures generally increase infection probability, the parameter a should be non-negative when $n^\circ \le n_s$. For $n^\circ > n_s$, the competition between social reinforcement and decay effects, characterized by the parameter b, will be taken into account for the message adoption. If the reinforcement is strong enough the individual will be more likely to adopt the message with increasing n° ($b > 0$). Otherwise, even if many infected neighbors try to transmit the information to the focal individual i, the multiple exposures will lead to a decreased probability for the information adoption ($b < 0$). In the present study, we set $b \in [-1, 1]$ and $a \in (0, 0.5]$ so that the spreading dynamics of the message can be comprehensively investigated. For sim-
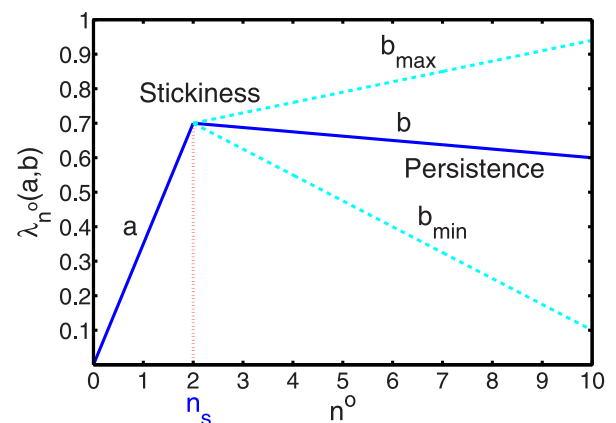
Figure 7 | The adopting probability of message as a function of n°. The degree of stickiness and persistence are quantified as an_s and b, respectively. n_s denotes the position of inflection point.

plicity, we do not consider the diversity of individuals' response to the message, and all individuals behave identically with the same values of parameters a and b.

We perform Monte Carlo (MC) simulations with synchronous updating of the states of all the individuals. Each MC step consists of the following three procedures: (i) All susceptible individuals decide whether or not to adopt the message with probability $\lambda_{n^\circ}(a, b)$; (ii) If an individual adopts the message, he will try to transmit what he has approved to all his nearest susceptible neighbors in the next step, and then becomes recovered immediately; (iii) Otherwise, the susceptible individuals will wait to repeat the procedure (i) in the following MC steps. The above elementary spreading processes are repeated $T'_S = 500$ steps until there are no infected individuals anymore in the population.

Theoretical analysis of the model. For the occurrence probability of transmission event E_n, $T_n(a,b) = 1 - e^{-\lambda_{n+1}(a,b)\tau} = 1 - e^{-\lambda_{n+1}(a,b)}$ with $\tau = 1$[36]. We consider the spatial correlations that affect the process of diffusion, but do not yet influence the critical behaviour of the message spreading[27]. The transition point from susceptible to infected phase is determined by[27,36]

$$T_C = \langle T \rangle, \tag{3}$$

where $\langle T \rangle$ and T_C are, respectively, the mean transmissibility and the critical topology-dependent bond-percolation threshold. On the other hand, the mean transmissibility can be gotten as

$$\langle T \rangle = \sum_{n=0}^{\langle k \rangle - 1} q_n T_n(a,b), \tag{4}$$

where $q_n = \binom{\langle k \rangle - 1}{n} p^n (1-p)^{\langle k \rangle - n - 1}$ is the probability that the recipient has other n ($n = 0, 1, 2, 3$) infected neighbors except for the one chosen beforehand when considering the spatial correlations (p is the probability that one nearest neighbor of the focal individual is in infected state).

Combine Eqs. (2), (3), (4), and $T_n(a,b) = 1 - e^{-\lambda_{n+1}(a,b)}$, the mean transmissibility for the discrete case reads as

$$
\begin{aligned}
\langle T \rangle &= \sum_{n=0}^{\langle k \rangle - 1} q_n \left(1 - e^{-\lambda_{n+1}(a,b)} \right) \\
&= \sum_{n=0}^{\langle k \rangle - 1} q_n - \sum_{n=0}^{n_s - 2} q_n e^{-\lambda_{n+1}(a,b)} - \sum_{n=n_s-1}^{\langle k \rangle - 1} q_n e^{-\lambda_{n+1}(a,b)} \\
&= 1 - \sum_{n=0}^{n_s - 2} q_n e^{-(n+1)a} - e^{-n_s a} \sum_{n=n_s-1}^{\langle k \rangle - 1} \binom{\langle k \rangle - 1}{n} e^{-(n-1)b}.
\end{aligned}
\tag{5}
$$

Then we have

$$\theta_2 e^{-2a} + \theta_1 e^{-a} + \theta_0 = 0, \tag{6}$$

where $\theta_2 = \sum_{n=1}^{\langle k \rangle - 1} \frac{(\langle k \rangle - 1)!}{n!(\langle k \rangle - n - 1)!} p^n (1-p)^{\langle k \rangle - n - 1} e^{(1-n)b}$, $\theta_1 = (1 - p)^{\langle k \rangle - 1}$ and $\theta_0 = T_C - 1 = -\frac{1}{2}$. The positive root is selected as the theoretical prediction

$$a(b) = \ln \left(\frac{2\theta_2}{\sqrt{\theta_1^2 - 4\theta_2 \theta_0} - \theta_1} \right). \tag{7}$$

In SF networks, the critical point T_C beyond which the message can reach a finite faction of the population can be obtained[36] as $T_c = \frac{\langle k \rangle}{\langle k^2 \rangle - \langle k \rangle}$. Since RRNs and ER networks are generated by connecting randomly selected pair of nodes, $T_c = \frac{1}{\langle k \rangle - 1}$. Due to the randomness of connections in these networks, each edge of a neighbor of one host can probably connected to any other $N - 2$ individuals. Therefore, the transmission of message from an informed neighbor to the susceptible host (i.e., message flows through the edge in the system) will happen only if there is at least one infected individual in the left $N - 2$ ones to transmit the message to the recipient. In other words, the probability that the message flow can reach the susceptible host through the edges connecting to the recipient is $p = \frac{1}{N - 2}$ for the three networks (i. e., the RRNs, SF networks, and ER networks). Also, two or more transmission events fail to last over the long time at the critical points (see Supplementary Fig. S8, Fig. S9, and Figs. S25), so the time correlations are ruled out in the theoretical analysis for the cases of RRNs, SF networks, and ER networks. In the regular lattices, $p = \frac{1}{\langle k \rangle - 1}$ since each one has exact $\langle k \rangle$ specific neighbors. Specifically, $p = \frac{1}{3} \approx 0.333$ for square lattice with von Neumann neighborhood, $p = \frac{1}{5}$ for Hexagonal lattice, and $p = \frac{1}{7} \approx 0.143$ for the lattice with Moore neighborhood. The bond percolation threshold $T_C = \frac{1}{2}$ for a

square lattice, $T_C = 0.347$ for Hexagonal lattice, and $T_C = 0.232$ for square lattice with Moore neighborhood[27,37].

Verification approximation of information threshold. In analogous to $q_n(a, b)$, we first set $Q'_n(a,b)(n=0, 1, 2, 3)$. According to Eq. (4), $Q'_n(a,b)(n=0, 1, 2, 3)$ can be calculated numerically by counting the relative number of successful attacks (i.e, transmission events) $\alpha_n(a, b)$ from infected neighbors to hosts for given values of a and b and equating this to $\alpha_n(a,b) = \frac{\sum_{t=0}^{T'_s} \omega_n(a,b,t)}{\sum_{m=0}^{\langle k \rangle - 1} \sum_{t=0}^{T'_s} \omega_m(a,b,t)} = \frac{Q'_n(a,b) T_i(a,b)}{\langle T(a,b) \rangle}$. In other words, $\alpha_n(a, b)$ represents numerically the proportion of E_n occurring at parameter point (a, b). Furthermore, $Q'_n(a,b)$ can be normalized as

$$
\begin{aligned}
Q_n(a,b) &= \frac{Q'_n(a,b)}{\sum_{m=0}^{\langle k \rangle - 1} Q'_m(a,b)} = \frac{\frac{\alpha_n(a,b)\langle T(a,b) \rangle}{T_n(a,b)}}{\sum_{m=0}^{\langle k \rangle - 1} \frac{\alpha_m(a,b)\langle T(a,b) \rangle}{T_m(a,b)}} \\
&= \frac{\frac{\alpha_n(a,b)}{T_n(a,b)}}{\sum_{m=0}^{\langle k \rangle - 1} \frac{\alpha_m(a,b)}{T_m(a,b)}}.
\end{aligned}
\tag{8}
$$

$Q_n(a, b)$ actually represents the probability that the recipient owns n ($n = 0, 1, 2, 3$) infected neighbors expect for the preselected informed one, involving both the spatial and time correlations of the message diffusion.

Herein, we select different parameter regions containing the numerical critical boundaries for corresponding lattices with $n_s = 2$ for the calculation of $Q_n(a, b)$. We average all non-zero $Q_n(a, b)$ in the selected parameter ranges for the expected indices Q_n, and substitute them into the equation

$$T_C = \langle T \rangle = \sum_{n=0}^{\langle k \rangle - 1} Q_n T_n(a,b) \tag{9}$$

to get the verification thresholds.

1. Centola, D. & Macy, M. Complex contagions and the weakness of long ties. *Am. J. Sociol.* **113**, 702–734 (2007).
2. Portrat, S., Barrouillet, P. & Camos, V. Time-related decay or interference-based forgetting in working memory. *J. Exp. Psychol. Learn Mem. Cogn.* **34**, 1561 (2008).
3. Centola, D. The spread of behavior in an online social network experiment. *Science* **329**, 1194 (2010).
4. Weng, L., Flammini, A., Vespignani, A. & Menczer, F. Competition among memes in a world with limited attention. *Sci. Rep.* **2**, 335 (2012).
5. Hodas, N. O. & Lerman, K. The Simple Rules of Social Contagion. *Sci. Rep.* **4**, 4343 (2014).
6. Cosley, D., Huttenlocher, D. P., Kleinberg, J. M., Lan, X.-Y. & Suri, S. Sequential influence models in social networks. In *Proc. 4th Int. AAAI Conf. on Weblogs and Social Media,* ICWSM' 10, (ACM, New York, NY, USA, 2010).
7. Romero, D. M., Meeder, B. & Kleinberg, J. Differences in the mechanics of information diffusion across topics: Idioms, political hashtags, and complex contagion on twitter. In *Proc. 20th Int. Conf. on World Wide Web,* WWW' 11, 695–704 (ACM, New York, NY, USA, 2011).
8. Lu, L., Chen, D. B. & Zhou, T. Small world yields the most effective information spreading. *New J. Phys.* **13**, 123005 (2011).
9. Krapivsky, P. L., Redner, S. & Volovik, D. Reinforcement-driven spread of innovations and fads. *J. Stat. Mech-Theory E.* **12**, 1104 (2011).
10. McQuail, D. & Windahl, S. *Communication Models for the Study of Mass Communications* (Longman, London, 1993).
11. Wu, F. & Huberman, B. A. Novelty and collective attention. *Proc. Nat. Acad. Sci.* **104**, 17599–17601 (2007).
12. Centola, D., Eguíluz, V. M. & Macy, M. W. Cascade dynamics of complex propagation. *Physica A* **374**, 449–456 (2007).
13. Watts, D. J. & Dodds, P. S. Influentials, networks, and public opinion formation. *J. Consum. Res.* **34**, 441–458 (2007).
14. Romero, D. M., Galuba, W., Asur, S. & Huberman, B. A. Influence and passivity in social media. In *Proc. 20th Int. Conf. on World Wide Web,* WWW' 11, 113–114 (ACM, New York, NY, USA, 2011).
15. Crane, R. & Sornette, D. Robust dynamic classes revealed by measuring the response function of a social system. *Proc. Nat. Acad. Sci.* **105**, 15649–15653 (2008).
16. Dall'Asta, L., Baronchelli, A., Barrat, A. & Loreto, V. Nonequilibrium dynamics of language games on complex networks. *Phys. Rev. E* **74**, 036105 (2006).
17. Barrat, A., Baronchelli, A., Dall'Asta, L. & Loreto, V. Agreement dynamics on interaction networks with diverse topologies. *Chaos* **17**, 026111 (2007).
18. Puglisi, A., Baronchelli, A. & Loreto, V. Cultural route to the emergence of linguistic categories. *Proc. Nat. Acad. Sci.* **105**, 7936–7940 (2008).
19. Baronchelli, A., Gong, T., Puglisi, A. & Loreto, V. Modeling the emergence of universality in color naming patterns. *Proc. Nat. Acad. Sci.* **107**, 2403–2407 (2010).
20. Tang, M., Liu, L. & Liu, Z.-H. Influence of dynamical condensation on epidemic spreading in scale-free networks. *Phys. Rev. E* **79**, 016108 (2009).
21. Tang, M., Liu, Z.-H. & Li, B.-W. Epidemic spreading by objective traveling. *Europhys. Lett.* **87**, 18005 (2009).

22. Zhao, Z.-D., Liu, Y. & Tang, M. Epidemic Variability in Hierarchical Geographical Networks with Human Activity Patterns. *Chaos* **22**, 023150 (2012).
23. Gong, K., Tang, M., Yang, H. & Shang, M.-S. Variability of contact process in complex networks. *Chaos* **21**, 043130 (2011).
24. Montanari, A. & Saberi, A. The spread of innovations in social networks. *Proc. Nat. Acad. Sci.* **107**, 20196–20201 (2010).
25. Ferreira, S. C., Castellano, C. & Pastor-Satorras, R. Epidemic thresholds of the susceptible-infected-susceptible model on networks: A comparison of numerical and theoretical results. *Phys. Rev. E* **86**, 041125 (2012).
26. Marro, J. & Dickman, R. *Nonequilibrium Phase Transitions Lattice Models* (Cambridge University Press, New York, 2005).
27. Pérez-Reche, F. J., Ludlam, J. J., Taraskin, S. N. & Gilligan, C. A. Synergy in spreading processes: from exploitative to explorative foraging strategies. *Phys. Rev. Lett.* **106**, 218701 (2011).
28. Newman, M. *Networks: An introduction.* (Oxford University Press, Oxford, 2010).
29. Taraskin, S. N. & Pérez-Reche, F. J. Effects of variable-state neighborhoods for spreading synergystic processes on lattices. *Phys. Rev. E* **88**, 062815 (2013).
30. Noah, E. Theoretical foundations for centrality measures. *Am. J. Sociol.* **96**, 1478–1504 (1991).
31. Gonzalez-Bailon, S., Borge-Holthoefer, J., Rivero, A. & Moreno, Y. The Dynamics of Protest Recruitment through an Online Network. *Sci. Rep.* **1**, 197 (2011).
32. Turow, J. *Media Today: An Introduction to Mass Communication* (Taylor & Francis, 2011).
33. Katz, E. & Lazarsfeld, P. F. *Personal influence: The part played by people in the flow of mass communications* (Transaction Publishers, 2006).
34. Samovar, L. A., Porter, R. E. & McDaniel, E. R. *Communication between cultures* (Cengage Learning, 2009).
35. Biagi, S. *Media impact: An introduction to mass media* (Wadsworth Publishing Company, 2011).
36. Newman, M. Spread of epidemic disease on networks. *Phys. Rev. E* **66**, 016128 (2002).
37. Sykes, M. F. & Essam, J. W. Exact critical percolation probabilities for site and bond problems in two dimensions. *J. Math. Phys.* **5**, 1117–1127 (2004).

Acknowledgments

Author contributions

P.B.C., M.T. and Z.X.W. designed the research, P.B.C. implemented and performed numerical experiments and simulations. P.B.C., M.T. and Z.X.W. analyzed the data and interpreted the results. P.B.C., M.T. and Z.X.W. wrote, reviewed and approved the manuscript.

Additional information

Supplementary information accompanies this paper at http://www.nature.com/scientificreports

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Cui, P., Tang, M. & Wu, Z. Message spreading in networks with stickiness and persistence: Large clustering does not always facilitate large-scale diffusion. *Sci. Rep.* **4**, 6303; DOI:10.1038/srep06303 (2014).