Phase Transitions in Information Spreading on Structured Populations

Jessica T. Davis[1], Nicola Perra[2,4], Qian Zhang[1], Yamir Moreno[3,4], and Alessandro Vespignani*[1,4]

[1]Laboratory for the Modeling of Biological and Socio-technical Systems, Northeastern University, Boston, MA, USA
[2]Networks and Urban Systems Centre, University of Greenwich, London, United Kingdom
[3]Institute for Biocomputation and Physics of Complex Systems (BIFI) and Department of Theoretical Physics, University of Zaragoza, 50018 Zaragoza, Spain
[4]ISI Foundation, Turin, Italy

Abstract

Mathematical models of social contagion that incorporate networks of human interactions have become increasingly popular, however, very few approaches have tackled the challenges of including complex and realistic properties of socio-technical systems. In this work we define a framework to characterize the dynamics of the Maki-Thompson rumor spreading model in structured populations, and analytically find a previously uncharacterized dynamical phase transition that separates the local and global contagion regimes. We validate our threshold prediction through extensive Monte Carlo simulations. Furthermore, we apply this framework in two real-world systems, the European commuting and transportation network and the Digital Bibliography and Library Project (DBLP) collaboration network. Our findings highlight the importance of the underlying population structure in understanding social contagion phenomena and have the potential to define new intervention strategies aimed at hindering or facilitating the diffusion of information in socio-technical systems.

The mathematical modeling of contagion processes is crucial in gaining insight into a broad range of phenomena from the spreading of infectious diseases to social collective behavior. While this avenue of research has a long tradition both in the biological and social sciences, in recent years there have been significant advancements triggered by increasing computational power and data availability characterizing socio-technical systems. These advances are particularly evident in the area of infectious disease forecasting where current models now incorporate realistic mobility and interaction data of human populations [1, 2, 3, 4, 5, 6]. Analogously, social contagion phenomena that were initially modeled using the same mathematical framework as epidemics [7, 8, 9, 10] are now described by *complex contagion* models [11, 12, 13] aimed at specifically characterizing processes such as the establishment of shared social norms and beliefs [14, 15, 16], the diffusion of knowledge and information [17, 18], and the emergence of political consensus [19]. These models consider complex factors such as reinforcement and threshold mechanisms [20, 21, 22, 23] and the loss of interest mediated by social interactions [24, 25, 26]. Furthermore, many of these theoretical approaches have put networks at the center of our understanding of social contagion phenomena and the information spreading process [27, 28, 8, 17, 29, 30, 31, 32, 12, 33]. However, most theoretical and numerical work on the dynamics of social contagion focuses on highly stylized models, trading off the realistic features of human interactions for analytical transparency and computational efficiency. As a result, social contagion models able to integrate the effects of human mobility, community structure, and time varying behavioral patterns are largely unexplored.

In this paper, we consider the classic rumor spreading model [24, 25] to study the effects of structured populations on the global diffusion of a rumor or piece of information. More specifically, we model the spatial structure of realistic populations and the behavior of individuals in virtual social networks through a reaction-diffusion model in a metapopulation network, and an activity-driven model with communities, respectively. We first identify analytically the necessary conditions for the social contagion to spread to a macroscopic fraction of the population. This analysis shows that although the rumor model is lacking any critical threshold, the population structure introduces a dynamical phase transition (global invasion threshold [34]) which is a function of the interactions between subpopulations. We validate the analytical results with large-scale numerical simulations on synthetic networks with different topological structures. Additionally, we recover the global threshold of the contagion process in data-driven models of the European transportation network and the Digital Bibliography and Library Project (DBLP) collaboration network.

Understanding how the social structure in both the physical and virtual worlds affects the emergence of contagion phenomena has the potential to indicate novel ways to utilize the network connectivity to develop efficient network-based interventions. The framework developed here opens a path to study the effects of communities and spatial structures in other complex contagion processes which can incorporate agent memory [35] and social reinforcement [22], or introduce other heterogeneous features such as age-dependent contact patterns and socio-economic conditions [36].

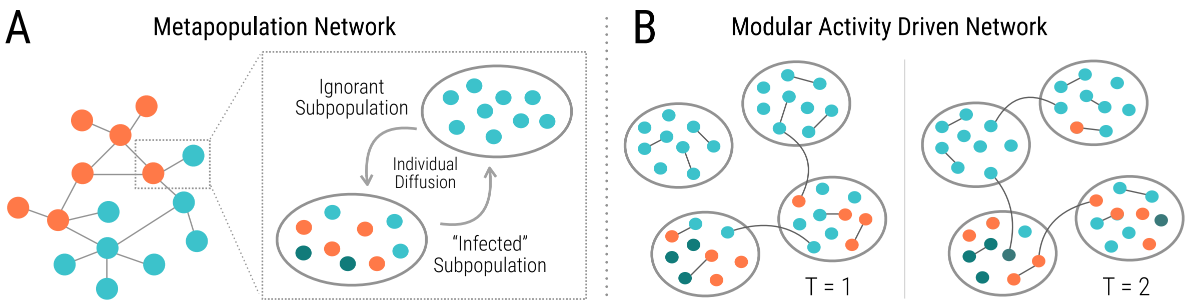

Figure 1: Types of structured populations considered in the modeling framework. **(A)** A schematic representation of a reaction-diffusion process on a metapopulation network, where individuals homogeneously interact within their current subpopulation, and then diffuse through the network constrained by the global structure. **(B)** A schematic representation of a modular activity-driven network at two points in time where individuals are confined to a single community, but when activated choose to form links to those outside of their current community based on a probability of inter-community interaction. Each instantaneous network is generated independently of prior networks.

1 Model Definition

Here we use a variant of the original rumor model [24], known as the Maki-Thompson (MT) model [25], to describe the spread of information through a population based on interactions between agents [37]. Similar to epidemic models, individuals can be classified into three compartments, *ignorants*: those who do not know the rumor, *spreaders*: those who know and are actively sharing the rumor, and *stiflers*: those who know the rumor but are no longer spreading it. The contagion process evolves through interactions between individuals in a population. If a spreader contacts an ignorant individual, with a probability λ, the ignorant will transition into a spreader. However, when a spreader contacts either a stifler or another spreader, with a probability α, the spreader will transition into a stifler. The stifling mechanism describes an individual's tendency to become uninterested in the rumor once the appeared novelty of the information is lost. In homogeneously mixed populations, this feature does not allow the presence of a *rumor threshold* [24, 38], meaning that for any $\lambda > 0$ the rumor will always spread to a macroscopic proportion of the population (see methods). We investigate the behavior of this model on two types of structured populations that incorporate the complexity observed in socio-technical systems (Fig. 1).

- **Rumor model in spatially structured populations**. We first consider a population where spatially defined groups of individuals (subpopulations) are coupled together by a mobility rate (Fig. 1a) [39, 33]. This structure, also called a metapopulation network, is used to model species persistence in ecosystems [40], the evolution of populations [41], and the global spreading of infectious diseases [42]. Specifically, we consider a metapopulation network with V subpopulations, each with an average population size of $\overline{N}$ individuals. Reaction-diffusion processes are used to characterize both the local interaction and global mobility dynamics. Individuals first *react* within their current subpopulation according to the rumor model dynamics and then *diffuse* between subpopulations based on a Markovian diffusion process. The probability that an individual will leave her current subpopulation and travel to a specific neighbor is p/k, where p is the mobility parameter and k is the number of neighboring subpopulations.

- **Rumor model in virtual structured populations**. In contrast to the reaction-diffusion scheme, rather than moving between communities, individuals may belong to a specific virtual community such as an interest or disciplinary group, online forum, or political affiliation etc., but interact occasionally with individuals in other virtual communities through collaborations, forum posts, or direct messages etc.(Fig. 1b). We model these interaction dynamics using a modular activity-driven network scheme [43, 44]. In particular, we consider a population with V communities whose sizes (s) follow a specific distribution, $P(s)$. Every individual is assigned an activity potential $a_i\Delta t$ that is sampled from a preset distribution $F(a)$. The activity of an individual corresponds to the probability with which the individual becomes active during a given time step [43]. Each active individual will form a single connection to another individual in the population creating an instantaneous network. To induce a community structure, an activated individual will choose to form a link to a randomly selected individual outside of her home community with a probability μ, otherwise, an intra-community link will be formed. The parameter μ allows us to tune the interaction between communities. After each single iteration of the rumor model, the network resets and a new instance is generated in the same manner.

2 Invasion threshold in structured populations

Although the rumor model in a single homogeneous population does not exhibit a spreading threshold, the presence of a subpopulation structure fundamentally alters the contagion dynamics. This can be clearly seen for the rumor model in virtual structured populations by examining two limits of the inter-community interaction term, μ. When $\mu = 0$, individuals will only interact with others in their community. Thus, the rumor will never escape the seed community. However, in the limit where $\mu = 1$, individuals effectively do not belong to any community and will always choose to form an external connection. Therefore, the rumor will certainly reach a macroscopic fraction of the population. The same reasoning can be applied to the limits of the mobility parameter, p, in the case of spatially structured populations. In both modeling frameworks, the population structure induces a transition point separating a dynamical regime where only local spreading is possible from a regime where the rumor spreads globally through the network.

To characterize this transition point quantitatively, we use a branching process framework to describe the rumor spreading dynamics across subpopulations [42, 45]. Let us consider a system that is structured into V subpopulations, each consisting of $\overline{N}$ individuals, on average, at any given time. Within the homogeneous population structure, we assume that all nodes are statistically equivalent and the connections formed between pairs of nodes are uncorrelated. Let D_n be the number of affected subpopulations where the rumor is known by at least one individual at generation n. We use a tree-like approximation to write an expression that captures the number of subpopulations that know the rumor at each generation of the spreading process, obtaining:

$$D_n = D_{n-1}\left(1 - \frac{\sum\limits_{m=0}^{n-1} D_m}{V}\right) C\Phi. \tag{1}$$

The above equation assumes that every affected subpopulation in the $(n-1)^{th}$ generation (D_{n-1}), may seed each one of its $(1 - \sum_{m=0}^{n-1} D_m/V)C$ unaffected neighbors with a probability Φ, where C indicates the average number of neighboring subpopulations and $1 - \sum_{m=0}^{n-1} D_m/V$ is the probability that the neighboring subpopulation is not already aware of the rumor during the $(n-1)^{th}$ generation. In a structured population model, a rumor epidemic occurs when each affected subpopulation, early in the contagion process, spreads the rumor on average to at least *one* fully ignorant subpopulation. Using the above expression, this global contagion condition reads as $D_n/D_{n-1} \geq 1$. Given that we are interested in the early time dynamics of the process, we assume that $\sum_{m=0}^{n-1} D_m/V << 1$, defining the *global contagion threshold* $D_n/D_{n-1} \simeq C\Phi \geq 1$. This effectively defines the *subpopulation reproductive number* $R^* = C\Phi$; i.e. the average number of communities becoming aware of the rumor from a single subpopulation. Analogously to the reproductive number in biological epidemics, in order for information to spread globally, R^* must be greater than or equal to one [42, 45]. The terms, C and Φ, depend explicitly on the type of structured population model as well as the contagion process. In the following sections we provide expressions for these parameters and the rumor invasion thresholds for both spatially structured and virtual populations.

2.1 Spatially structured populations

In a homogeneous metapopulation network the number of possible subpopulations that could be seeded by each affected subpopulation is $C = \langle k \rangle - 1$; i.e. the average number of neighboring subpopulations minus the one which originally seeded the contagion process. Due to the lack of a local rumor threshold within a single subpopulation, the probability Φ is simply given by the probability that at least one spreader will decide to leave an affected community and travel to a neighboring subpopulation where it will start spreading the rumor:

$$\Phi = 1 - \left(1 - \frac{p}{\langle k \rangle}\right)^{\beta}, \tag{2}$$

where β is the number of spreaders in an affected subpopulation that can travel out of their current subpopulation during the rumor epidemic. This value Φ, is calculated by considering one minus the probability that none of the spreaders will travel to a new community, $(1 - \frac{p}{\langle k \rangle})^{\beta}$. Here, $\beta = \frac{2(1+\lambda/\alpha)\overline{N}}{\alpha}$, is the product of the total number of individual spreaders generated by the contagion process within a single population and the average amount of time they are actively spreading the rumor (details in methods). Using the above expressions and considering small mobility probabilities, such that $p/\langle k \rangle << 1$, we can approximate the probability $\Phi \simeq \beta p/\langle k \rangle$. In this limit we obtain an explicit expression for the rumor invasion threshold as:

$$R^* = \frac{\langle k \rangle - 1}{\langle k \rangle} 2 \left(1 + \frac{\lambda}{\alpha}\right) \frac{p\overline{N}}{\alpha} \geq 1. \tag{3}$$

From the above expression it is possible to rewrite the necessary threshold condition to find the critical mobility p_c in the system required for a global spreading of the rumor as:

$$p_c = \frac{\langle k \rangle}{(\langle k \rangle - 1)} \; \frac{\alpha}{2(1 + \frac{\lambda}{\alpha})\overline{N}} \tag{4}$$

Below the critical value, p_c, the amount of individual mobility restricts the global propagation of the rumor. In this subcritical regime, spreaders in affected communities are generally unable to travel to a new subpopulation before they transition into stiflers, which consequently causes the rumor to go extinct in the early stages. This critical mobility is a function of both the network structure and the rumor model parameters, λ and α. However, for homogeneous networks with sufficiently large average degrees $\langle k \rangle$, the effect of the network structure is relatively insignificant. In the Supplementary Information we derive the critical mobility for metapopulation networks with heavy tailed degree distributions and find that the analytical expression depends not only the average degree of the network $\langle k \rangle$, but also the second moment $\langle k^2 \rangle$ of the degree distribution. Heterogeneous networks are characterized by having degree distributions with high variance (large $\langle k^2 \rangle$), thus considerably affecting the value of the mobility threshold. We also see that p_c is linearly dependent on α. When λ is small relative to α, the critical mobility is controlled predominantly by the stifling probability. Recall that the stifling probability characterizes the tendency for an individual to become disinterested in the rumor (i.e. transition into a stifler) when interacting with others that know the rumor. This finding is a feature worth remarking for the global spread of a rumor in a spatially structured environment, that places the emphasis not on how *appealing* a rumor is, but rather on the rate at which people decide that the rumor is not worth spreading.

2.2 Virtual structured populations

Now let us consider a modular activity-driven network where we assume discrete time, $\Delta t = 1$, a homogeneous activity rate (a) for all individuals in the network, and a homogenous distribution of community sizes. In this model, if an individual chooses to form an inter-community link, by construction it can choose any of the other $C = V - 1$ communities. The probability Φ, that at least one spreader from an affected population will choose to connect another individual outside of their current community and successfully transmit the rumor can be written as

$$\Phi = 1 - \Big(1 - \frac{\lambda\mu}{V-1}\Big)^{\beta}. \tag{5}$$

In this expression $\frac{\lambda\mu}{V-1}$ is the probability that an inter-community link successfully transmits the rumor to one of the $V - 1$ specific subpopulations and β translates to the number of potential chances that a single, affected community has to spread the rumor to another community. As described in the methods section, β is not explicitly dependent on the activity assigned to each individual as long as $a > 0$ and still remains a function of the total number of spreaders in the community and the average amount of time they were active. Therefore, the same equation used for spatially structured populations, $\beta = \frac{2(1+\lambda/\alpha)\overline{N}}{\alpha}$, holds. We can thus calculate the rumor invasion threshold by assuming that $\frac{\lambda\mu}{V-1} << 1$ in the limit of large V, obtaining:

$$R^* = 2(1 + \frac{\lambda}{\alpha})\frac{\lambda\overline{N}\mu}{\alpha} \geq 1. \tag{6}$$

The rumor invasion threshold in terms of the critical inter-community interaction rate, μ_c reads,

$$\mu_c = \frac{\alpha}{2\lambda\overline{N}(1 + \frac{\lambda}{\alpha})}. \tag{7}$$

This expression resembles the mobility threshold of Eq. 4 except for the addition of the λ parameter in the denominator, which comes from the node interaction process. In the spatially structured model, the mobility of an individual was the only factor that controlled whether the rumor spread to a new community. However, in the activity-driven model, active individuals do not move to another community, but rather may form a single connection through which the rumor has to be successfully transmitted to another individual in order to start the contagion process. This introduces a linear dependence on α/λ rather than α alone. When the spreading probability λ is high, the more likely an ignorant individual will transition into a spreader during a specific interaction. Thus, the rumor spreads more readily and does not require a high amount inter-community interaction to globally propagate. This result shows the inherent differences between Eq. 4 and Eq. 7 and brings attention to the importance of the type of structured population used when modeling a socio-technical system. In the SI we derive the critical interaction probability for populations with a heterogeneous size distribution.

2.3 Simulations in synthetic structured populations

In order to validate the analytical findings, we performed an extensive set of stochastic simulations of the rumor model on synthetic, structured populations. We generated homogeneous metapopulation networks as Erdös-Rényi random graphs with average degrees of $\langle k \rangle = 12$ and an average population size of $\overline{N} = 10^3$ individuals. To initiate the contagion process, one individual is made aware of the rumor in a single, randomly selected community. The microscopic *reaction* dynamics are mathematically defined by chain binomial and multinomial processes, which were used to update the stochastic transitions of individuals between compartments (details in SI). Following the reaction process, individuals *diffuse* along a specific link to a neighboring subpopulation with a probability p/k where k is the degree of the individual's current community. Our simulation results show the final fraction of affected communities as a function of the mobility probability p for various α parameters (see Fig. 2a), recovering the critical transition which separates the non-spreading and global spreading dynamical regimes. The vertical lines represent the values predicted from Eq. 4 and are in good agreement with our numerical findings. Furthermore, we see a clear dependence of the transition point on the stifling rate α. In the SI we show similar simulations for metapopulation networks with heavy tailed degree distributions, $p(k) \sim k^{-2.2}$. In this scenario, the network structure significantly reduces the threshold since, as mentioned above, it now depends on the second moment of the degree distribution which diverges for heterogeneous networks when $V \to \infty$.

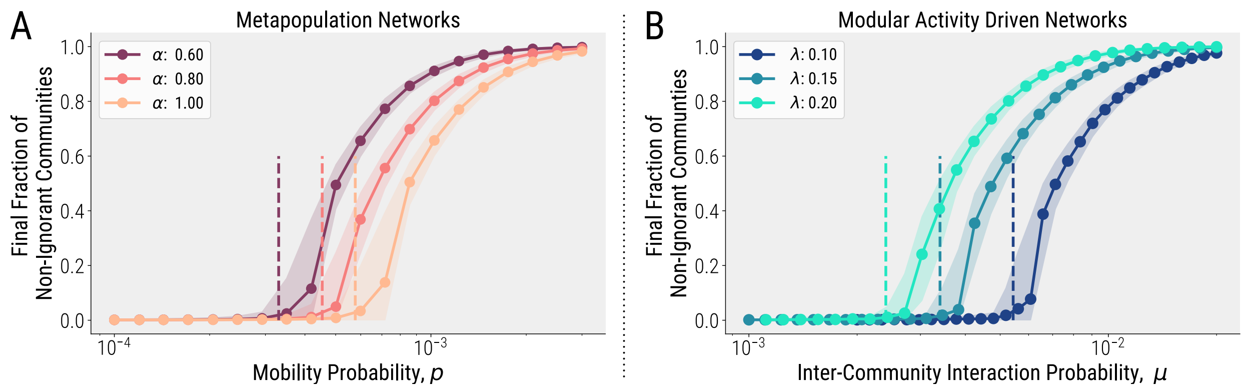

Figure 2: Results from numerical simulations of the rumor spreading process in homogeneous structured populations. **(A)** The final fraction of subpopulations where the rumor is known at the end of the rumor epidemic as a function of the mobility rate p, for varying α values averaged over 4,000 simulations. The networks have $V = 10^3$ subpopulations with an average of $\overline{N} = 10^3$ individuals and an average degree $\langle k \rangle = 12$. The vertical lines represent the predicted threshold values from Eq. 4 and the value $\lambda = 0.1$ was used for the spreading probability. **(B)** The final fraction of communities where the rumor is known as a function of the inter-community interaction probability for varying λ values averaged over 1,000 simulations. A total population had $V = 10^3$ communities, each containing 10^3 individuals. The vertical lines represent the predicted values from Eq. 7 and a value of $\alpha = 1$ was used as the stifling probability. The error regions represent the 90% reference range and the averages in the supercritical regime were calculated on the simulations where at least 5% of the subpopulations experienced a rumor epidemic.

For the second type of structured population, we generated modular activity-driven networks with $V = 10^3$ total communities, each with the same number of individuals, $\overline{N} = 10^3$, and every individual assigned the same activity probability, $a = 0.1$. To start the contagion process, a single individual from a randomly selected community is made aware of the rumor. An instantaneous network is generated by the modular activity-driven network model on which the rumor dynamics unfold for a single iteration. After the reaction process, a new network instance is generated and the process repeats until all individuals are either still ignorants or stiflers (more model details in SI). In Fig. 2b we show the final fraction of communities where a rumor epidemic occurred as a function of the inter-community interaction parameter μ, for multiple λ values. The phase diagram supports our theoretical findings, and confirms that for higher values of λ less inter-community interactions are required for the rumor to globally propagate. Additionally, we also model this system using a heterogeneous size distribution and report the results in the SI.

3 Data-driven Simulations

To further support the theoretical results obtained in the previous section, we analyze the rumor model on two real-world networks. Specifically, we simulate a rumor spreading across a metapopulation network modeling the transportation patterns in Europe and across a modular activity-driven network modeling scholarly collaborations from the DBLP collaboration network. The mobility of individuals throughout Europe is constructed by dividing the continent into spatial regions that are coupled together using data about commuting patterns and long-range transportation fluxes such as airline traffic (details in SI). This realistic, synthetic metapopulation network has been used in simulations of emerging infectious diseases as well as in the analysis and predictions of pandemic events [46, 47, 10]. In this framework, the mobility of individuals across subpopulations (analogous to the p parameter in spatially structured populations) is derived from actual transportation data. To study the effects of a reduction in mobility, we rescaled the proportion of individuals that travel at each time step by a factor ω. We show the results of the rescaled mobility on the spatial diffusion of a rumor simulated over the transportation network in Fig. (3a,c). Interestingly, we see a clear transition in the mobility required for the rumor to spread. The phase diagram reveals that the critical ω_c in this system is significantly small implying that the current human mobility pattern across Europe is orders of magnitude above the rumor threshold.

We also model the collaboration process of the DBLP co-authorship network using the modular activity-driven network scheme [48]. Nodes in the network are individual researchers that can form collaborations with others either within or outside of their own communities. In particular, a link represents co-authorship on at least one paper and each community is a specific publication venue. We measure the amount of interaction μ between communities by calculating the frequency of cross-community links relative to the total number of internal and external links. We simulated the rumor model to analyze how information would propagate in this system by rescaling the actual individual's tendency to link outside of their current community (analogous to the μ parameter in the virtual structured population framework) by a factor ω to study the effects of lowering inter-community interaction rates. In Fig. (3b,d), we show the results of this rescaling on the final fraction of affected disciplinary communities and observe a transition point characterizing the amount of inter-community collaboration needed in order for a rumor or idea to spread globally. Similar to the transportation network, the critical rescaling value ω_c is extremely small. Both data-driven network applications extend our modeling framework by incorporating heterogeneous and non-trivial subpopulation interactions. Consequently, the assumptions of statistical equivalence of nodes and an uncorrelated network structure made in our calculations are no longer valid. Therefore, in these realistic systems, the critical value ω_c, can not be easily computed analytically. However, we do see a similar phenomenology between the synthetic and data-driven structured populations in that there does exist a critical transition point in the amount of interaction between subpopulations or communities that is necessary for the a rumor to propagate. In both cases, the critical transition point (ω_c) is very small, highlighting the role of our interconnected world in facilitating the diffusion of information across geographical boundaries as well as through disciplinary communities. However, this result is not necessarily universal across all types of structured populations. Information spreading is fundamentally dependent on the strength of interactions among elements of the network, thus calling for specific case by case studies on the location of critical transition points in real world situations.

4 Discussion and Conclusion

In this work, by using a classic rumor spreading model lacking any critical threshold in a single homogeneous population, we show that the contagion process in structured populations exhibits a phase transition with a critical threshold dependent on the amount of interactions/coupling between subpopulations. The analytical and numerical results presented here emphasize the importance of accounting for the complex structure observed in socio-technical systems when studying social contagion processes. The features observed in real-world systems can potentially alter the theoretical picture and the understanding provided by only studying stylized models. Our results show that successful information or rumor spreading is the result of a complex interaction between the intrinsic properties of the contagion process and the dynamics of interactions between subpopulations/communities that comprise social systems.The flexibility of the framework allows for further study of different types of emergent behaviors that may be more complex than the rumor model used here. For example, in order for a contagion to spread, individuals must be contacted by multiple neighbors in their social network. Analogously, additional features can be incorporated into the interaction process and network structure such as age-dependent contact patterns, socio-economic conditions, and data-driven human mobility. These features have the potential to not only provide unexpected results of theoretical nature but also actionable insights crucial to understand and control social contagion phenomena.

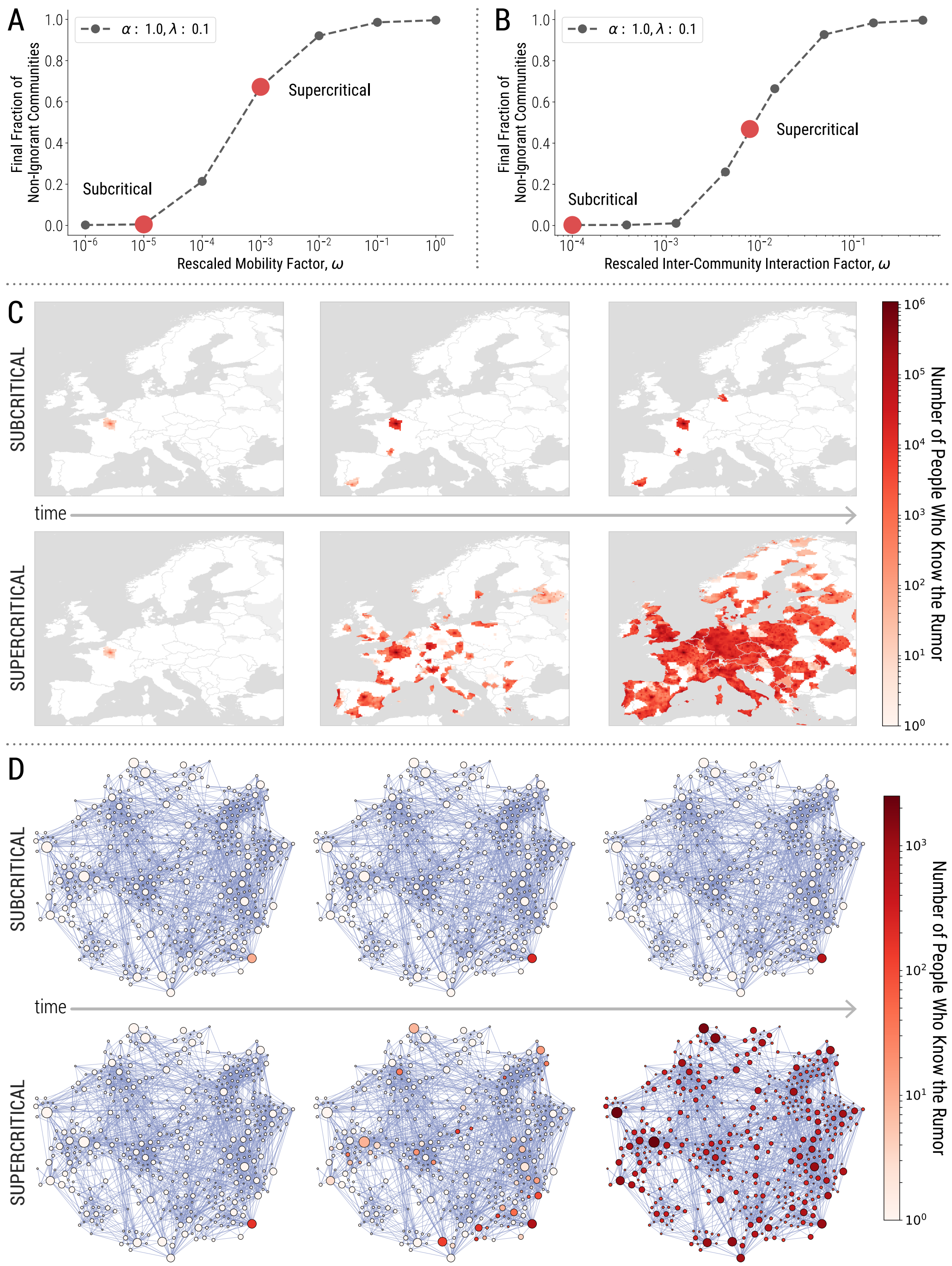

Figure 3: Results from numerical simulations of a rumor spreading in real-world networks. **(A-B)** The average final fraction of stiflers as a function of (A) the rescaling mobility factor in the European commuting and transportation network and (B) the rescaling factor of the inter-community interaction probability within the DBLP collaboration network. Simulations used a spreading probability of $\lambda = 0.1$ and stifling probability of $\alpha = 1$. Error bars represent the 90% reference range of the simulations where at least 5% of the subpopulations experienced a rumor epidemic. **(C-D)** The temporal evolution of the rumor spreading process taken from individual simulations corresponding to the ω values highlighted in (A-B) with red circles. (C) In the European commuting and transportation network, the rumor was initiated Paris, France by seeding one individual. (D) In the DBLP collaboration network, nodes represent publication venues where node size corresponds to the population size and the line thickness corresponds to the amount of inter-community interaction between each pair of communities. The activity probability per individual is $a = 0.1$, while the inter and intra-community probability is derived from the data.

Materials and Methods

Final Rumor Size in a Single Population: The mean field rate equations for the Maki-Thompson (MK) rumor model in a homogeneously mixed population are listed below. The densities of spreaders (S), ignorants (I) and stiflers (R) in a population are defined by $s = S/N, i = I/N, r = R/N$, respectively, where N is the total number of individuals in the population, yielding:

$$\begin{aligned} \frac{di}{dt} &= -\lambda s(t)i(t) \\ \frac{ds}{dt} &= \lambda s(t)i(t) - \alpha s(t)(s(t)+r(t)) \\ \frac{dr}{dt} &= \alpha s(t)(s(t)+r(t)). \end{aligned} \tag{8}$$

Using the initial conditions $i(0) \approx 1$ and $r(0) = 0$, a solution to these differential equations can be obtained analytically in the infinite time limit. The transcendental equation below has a trivial solution when $r_\infty = 0$, but a non-trivial solution, $r_\infty = 1 - e^{-(1+\frac{\lambda}{\alpha})r_\infty}$, when $\lambda/\alpha + 1 > 1$, confirming that a rumor will propagate through a population and reach a macroscopic fraction of individuals [38]. Assuming that $(1+\frac{\lambda}{\alpha})r_\infty << 1$, we can obtain the approximate solution:

$$r_\infty \simeq \frac{2\frac{\lambda}{\alpha}}{(1+\frac{\lambda}{\alpha})^2} \approx 2\frac{\lambda}{\alpha} \tag{9}$$

We can see that the final density of stiflers scales with λ/α. This relationship is verified through numerical simulations of the rumor model for a single homogeneously mixed population as detailed in the Supplementary Material.

Average Spreading Time: The average spreading time, $\langle\tau\rangle$, is the time elapsed since the individual was first told the rumor to the time the individual became a stifler. Fig. 1b in the SM shows the average spreading time as a function of $\frac{1}{\lambda} + \frac{1}{\alpha}$ from simulations done on a single population. A linear line is fit to the data, producing the equation:

$$\langle\tau\rangle = \frac{1}{\lambda} + \frac{1}{\alpha} + \frac{1}{2} \tag{10}$$

Number of potential spreaders: The number of potential spreaders β that could transmit the rumor to another population can be calculated as $\beta = \langle\tau\rangle\overline{N}r_\infty$ where $\langle\tau\rangle$ is the average amount of time an individual remains a spreader, and $\overline{N}r_\infty$ is the final average number of individuals that know the rumor at the end of the spreading process. Using the approximated equations for the final stifler density as well as the average spreading time, one obtains:

$$\beta = r_\infty\overline{N}\langle\tau\rangle = \frac{2(1+\frac{\lambda}{\alpha})\overline{N}}{\alpha}. \tag{11}$$

In the modular activity-driven network model, the expression for β is not altered by the fact that individuals are “activated" with probability a. The average spreading time $\langle\tau\rangle$ should be measured by considering the duration of the contagion process, which should be on the order of $1/a$ (average number of time steps between activations), and the activity of the individual at each time step, which is a. It follows that these terms cancel each other out, so the effective number of interactions of each spreader will not be dependent on a as shown also numerically in the SM.

Data Availability Statement

The data represented in Fig 3b are available through the Stanford Network Analysis Project (SNAP) [49]. All other data that supports the plots within this paper and other findings of this study are available from the corresponding author upon reasonable request.

Code Availability

Code is available upon request from the corresponding author.

Acknowledgments

N.P. was supported in part by the U. S. Army Research Laboratory and the U. S. Army Research Office under contract/grant number W911NF-18-1-0376. Y. M. acknowledges partial support from the Government of Aragón, Spain through grant E36-17R, by MINECO and FEDER funds (grant FIS2017-87519-P) and from Intesa Sanpaolo Innovation Center. The funder had no role in study design, data collection, and analysis, decision to publish, or preparation of the manuscript.

Author contributions

J.T.D., N.P., Y.M, and A.V. designed the research, J.T.D., A.V., and Q.Z., performed research and analyzed data, J.T.D., N.P., Q.Z., Y.M., and A.V. wrote the manuscript.

Competing Interests

The authors declare no competing financial interests.

References

[1] Hufnagel, L., Brockmann, D. & Geisel, T. Forecast and control of epidemics in a globalized world. *Proceedings of the National Academy of Sciences* **101**, 15124–15129 (2004).

[2] Colizza, V., Barrat, A., Barthelemy, M. & Vespignani, A. The role of the airline transportation network in the prediction and predictability of global epidemics. *Proceedings of the National Academy of Sciences* **103**, 2015–2020 (2006).

[3] Halloran, M. E. *et al.* Modeling targeted layered containment of an influenza pandemic in the united states. *Proceedings of the National Academy of Sciences* **105**, 4639–4644 (2008).

[4] Marathe, M. & Vullikanti, A. K. S. Computational epidemiology. *Commun. ACM* **56**, 88–96 (2013).

[5] Heesterbeek, H. *et al.* Modeling infectious disease dynamics in the complex landscape of global health. *Science* **347** (2015).

[6] Liu, Q.-H. *et al.* Measurability of the epidemic reproduction number in data-driven contact networks. *Proceedings of the National Academy of Sciences* **115**, 12680–12685 (2018).

[7] Goffman, W. Mathematical approach to the spread of scientific ideas - The history of mast cell research. *Nature* **212**, 449–452 (1966).

[8] Moreno, Y., Nekovee, M. & Pacheco, A. F. Dynamics of rumor spreading in complex networks. *Physical Review E - Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics* **69**, 7 (2004).

[9] Bettencourt, L. M., Cintrón-Arias, A., Kaiser, D. I. & Castillo-Chávez, C. The power of a good idea: Quantitative modeling of the spread of ideas from epidemiological models. *Physica A: Statistical Mechanics and its Applications* **364**, 513–536 (2006).

[10] GLEAMviz.org. URL `http://www.gleamviz.org/`.

[11] Centola, D. *How behavior spreads: The science of complex contagions*, vol. 3 (Princeton University Press, 2018).

[12] Guilbeault, D., Becker, J. & Centola, D. Complex contagions: A decade in review. In *Complex spreading phenomena in social systems*, 3–25 (Springer, 2018).

[13] Lehmann, S. & Ahn, Y.-Y. (eds.) *Complex spreading phenomena in social systems : influence and contagion in real-world social networks* (Springer International Publishing, 2018).

[14] Axelrod, R. The dissemination of culture: A model with local convergence and global polarization. *Journal of Conflict Resolution* (1997).

[15] Baronchelli, A., Felici, M., Loreto, V., Caglioti, E. & Steels, L. Sharp transition towards shared vocabularies in multi-agent systems. *Journal of Statistical Mechanics: Theory and Experiment* **2006**, P06014–P06014 (2006).

[16] Centola, D., Becker, J., Brackbill, D. & Baronchelli, A. Experimental evidence for tipping points in social convention. *Science* **360**, 1116–1119 (2018).

[17] Moreno, Y., Nekovee, M. & Vespignani, A. Efficiency and reliability of epidemic data dissemination in complex networks. *Physical Review E - Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics* **69**, 4 (2004).

[18] Gleeson, J. P., O'Sullivan, K. P., Baños, R. A. & Moreno, Y. Effects of network structure, competition and memory time on social spreading phenomena. *Phys. Rev. X* **6**, 021019 (2016).

[19] Volkening, A., Linder, D. F., Porter, M. A. & Rempala, G. A. Forecasting elections using compartmental models of infection. Tech. Rep. (2018).

[20] Centola, D. & Macy, M. Complex Contagions and the Weakness of Long Ties. *American Journal of Sociology* (2007).

[21] Granovetter, M. Threshold Models of Collective Behavior. *American Journal of Sociology* (1978).

[22] Watts, D. J. & Dodds, P. Threshold models of social influence. In *The Oxford Handbook of Analytical Sociology* (2017).

[23] Baronchelli, A. The emergence of consensus: a primer. *Royal Society open science* **5**, 172189 (2018).

[24] Daley, D. J. & Kendall, D. G. Epidemics and Rumours. *Nature* **204**, 1118 (1964).

[25] Maki, D. P. & Thompson, M. *Mathematical models and applications: with emphasis on the social life, and management sciences* (1973).

[26] Zhang, Q., Balcan, D. & Vespignani, A. Phase transition in rumor spread on complex contact networks. unpublished.

[27] Zanette, D. H. Dynamics of rumor propagation on small-world networks. *Physical Review E* **65**, 041908 (2002).

[28] Kempe, D., Kleinberg, J. & Tardos, É. Maximizing the spread of influence through a social network. In *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining - KDD '03*, 137 (ACM Press, New York, New York, USA, 2003).

[29] Kosfeld, M. Rumours and markets. *Journal of Mathematical Economics* **41**, 646–664 (2005).

[30] Liu, Z. & Bambi, H. U. Epidemic spreading in community networks. *Europhysics Letters* **72**, 315–321 (2005).

[31] Nekovee, M., Moreno, Y., Bianconi, G. & Marsili, M. Theory of rumour spreading in complex social networks. *Physica A* **374**, 457–470 (2007).

[32] Borge-Holthoefer, J., Meloni, S., Gonçalves, B. & Moreno, Y. Emergence of Influential Spreaders in Modified Rumor Models. *Journal of Statistical Physics* **151**, 383–393 (2013).

[33] Nematzadeh, A., Rodriguez, N., Flammini, A. & Ahn, Y.-Y. Optimal modularity in complex contagion (2018).

[34] Vespignani, A. Modeling dynamical processes in complex socio-technical systems. *Nature Physics* **8**, 32–30 (2012).

[35] Karsai, M., Perra, N. & Vespignani, A. Time varying networks and the weakness of strong ties. *Scientific Reports* **4**, 4001 (2014).

[36] Fumanelli, L., Ajelli, M., Manfredi, P., Vespignani, A. & Merler, S. Inferring the structure of social contacts from demographic data in the analysis of infectious diseases spread. *PLOS Computational Biology* **8**, 1–10 (2012).

[37] Daley, D. J. & Gani, J. *Epidemic Modelling: An Introduction*. Cambridge Studies in Mathematical Biology (Cambridge University Press, 1999).

[38] Barrat, A., Barthelemy, M. & Vespignani, A. *Dynamical Processes on Complex Networks* (Cambridge University Press, Cambridge, 2008).

[39] Levins, R. Some Demographic and Genetic Consequences of Environmental Heterogeneity for Biological Control. *Bulletin of the Entomological Society of America* **15**, 237–240 (1969).

[40] Keeling, M. J. Metapopulation moments: Coupling, stochasticity and persistence. *Journal of Animal Ecology* **69**, 725–736 (2000).

[41] Murrell, D. J. & Law, R. Beetles in fragmented woodlands: a formal framework for dynamics of movement in ecological landscapes. *Journal of Animal Ecology* **69**, 471–483 (2000).

[42] Colizza, V. & Vespignani, A. Epidemic modeling in metapopulation systems with heterogeneous coupling pattern: Theory and simulations. *Journal of Theoretical Biology* **251**, 450–467 (2008).

[43] Perra, N., Gonçalves, B., Pastor-Satorras, R. & Vespignani, A. Activity driven modeling of time varying networks. *Scientific reports* **2**, 469 (2012).

[44] Nadini, M. *et al.* Epidemic spreading in modular time-varying networks. *Scientific reports* **8**, 2352 (2018).

[45] Balcan, D. & Vespignani, A. Phase transitions in contagion processes mediated by recurrent mobility patterns. *Nature Physics* **7**, 581–586 (2011).

[46] Balcan, D. *et al.* Seasonal transmission potential and activity peaks of the new influenza a (h1n1): a monte carlo likelihood analysis based on human mobility. *BMC medicine* **7**, 45 (2009).

[47] Balcan, D. *et al.* Modeling the spatial spread of infectious diseases: The GLobal Epidemic and Mobility computational model. *Journal of Computational Science* **1**, 132–145 (2010).

[48] Yang, J. & Leskovec, J. Defining and evaluating network communities based on ground-truth. *Knowledge and Information Systems* **42**, 181–213 (2015).

[49] Leskovec, J. & Krevl, A. SNAP Datasets: Stanford large network dataset collection. `http://snap.stanford.edu/data` (2014).