REVIEW ARTICLES | INSIGHT
PUBLISHED ONLINE: 22 DECEMBER 2011 | DOI: 10.1038/NPHYS2160
nature physics

Modelling dynamical processes in complex socio-technical systems

Alessandro Vespignani[1,2]

In recent years the increasing availability of computer power and informatics tools has enabled the gathering of reliable data quantifying the complexity of socio-technical systems. Data-driven computational models have emerged as appropriate tools to tackle the study of dynamical phenomena as diverse as epidemic outbreaks, information spreading and Internet packet routing. These models aim at providing a rationale for understanding the emerging tipping points and nonlinear properties that often underpin the most interesting characteristics of socio-technical systems. Here, using diffusion and contagion phenomena as prototypical examples, we review some of the recent progress in modelling dynamical processes that integrates the complex features and heterogeneities of real-world systems.

Questions concerning how pathogens spread in population networks, how blackouts can spread on a nationwide scale, or how efficiently we can search and retrieve data on large information structures are generally related to the dynamics of spreading and diffusion processes. Social behaviour, the spread of cultural norms, or the emergence of consensus may often be modelled as the dynamical interaction of a set of connected agents. Phenomena as diverse as ecosystems or animal and insect behaviour can all be described as the dynamic behaviour of collections of coupled oscillators. Although all these phenomena refer to very different systems, their mathematical description relies on very similar models that depend on the definition and characterization of a large number of individuals and their interactions in spatially extended systems.

The modelling of dynamical processes is a research field that crosses different disciplines and has developed an impressive array of methods and approaches, ranging from simple explanatory models to realistic approaches capable of providing quantitative insight into real-world systems. Initially these models used simplistic assumptions for the micro-processes of interaction and were mostly concerned with the study of the emerging macro-level behaviour. This interest has favoured the use of techniques akin to statistical physics and the analysis of nonlinear, equilibrium and non-equilibrium physical systems in the study of collective behaviour in social and population systems. In recent years, however, the increase in interdisciplinary work and the availability of system-level high-quality data has opened the way to data-driven models aimed at a realistic description of complex socio-technical systems. Modelling approaches to dynamical processes in complex systems have been expanded into schemes that explicitly include spatial structures and have thus grown into a multiscale framework in which the various possible granularities of the system are considered through different approximations. These models offer a number of interesting and sometimes unexpected behaviours whose theoretical understanding represents a new challenge that has considerably transformed the mathematical and conceptual framework for the study of dynamical processes in complex systems.

Dynamical processes and phase transitions

The study of dynamical processes and the emergence of macro-level collective behaviour in complex systems follows a conceptual route essentially similar to the statistical physics approach to non-equilibrium phase transitions. A prototypical example is that of contagion processes. Epidemiologists, computer scientists and social scientists share a common interest in studying contagion phenomena and rely on very similar spreading models for the description of the diffusion of viruses, knowledge and innovations[1–5]. All these processes define a contagion dynamics that can be seen as an actual biological pathogen that spreads from host to host, or a piece of information or knowledge that is transmitted during social interactions. Let us consider the simple susceptible–infected–recovered (SIR) epidemic model. In this model, infected individuals (labelled with the state I) can propagate the contagion to susceptible neighbours (labelled with the state S) with rate λ, while infected individuals recover with rate μ and become removed from the population. This is the prototypical model for the spread of infectious diseases where individuals recover and are immune to disease after a typical time that, on average, can be expressed as the inverse of the recovery rate. A classic variation of this model is the susceptible–infected–susceptible (SIS) model, in which individuals revert to the susceptible state with rate μ, modelling the possibility of re-infection of individuals. The mapping between epidemic models and non-equilibrium phase transitions was pointed out in physics long ago, making those models of very broad relevance also outside the area of information and disease spreading. The static properties of the SIR model can indeed be mapped to an edge-percolation process[6]. Analogously, the SIS model can be regarded as a generalization of the contact-process model[7], widely studied as the paradigmatic example of an absorbing-state phase transition with a unique absorbing state[8].

A cornerstone feature of epidemic processes is the presence of the so-called epidemic threshold[1]. In a fully homogeneous population, the behaviour of the SIR model is controlled by the reproductive number $R_0 = \beta/\mu$, where $\beta = \lambda\langle k\rangle$ is the per-capita spreading rate, which takes into account the average number of contacts $\langle k\rangle$ of each individual. The reproductive number simply identifies the average number of secondary cases generated by a primary case in an entirely susceptible population and defines an epidemic threshold such that only if $R_0 \geq 1$ ($\beta \geq \mu$) can epidemics reach an endemic state and spread into a closed population. The SIS and SIR models are indeed characterized by a threshold defining the transition between two very different regimes. These regimes are determined by the values of the disease parameters, and characterized by

[1]Department of Physics, College of Computer and Information Sciences, Bouvé College of Health Sciences, Northeastern University, Boston, Massachusetts 02115, USA, [2]Institute for Scientific Interchange (ISI), Torino, 10133, Italy. e-mail: a.vespignani@neu.edu.

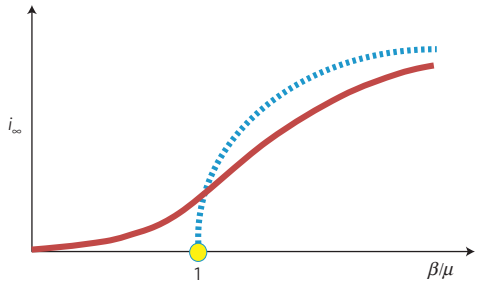

Figure 1 | Phase diagram of epidemic models. Illustration of the behaviour of the prevalence i_∞ for the SIS and SIR model in a heterogeneous network (solid line) as a function of the spreading rate β/μ, compared with the theoretical prediction for a homogeneous network (dashed line). The figure clearly shows the difference between homogeneous and heterogeneous networks, where the epidemic threshold is shifted to very small values. For scale-free networks with degree distribution exponent $\gamma \leq 3$, however, the associated prevalence i_∞ is extremely small over a large range of values of β/μ. In other words, as noted since the first work on epidemic spreading in complex networks, the bad news about the suppression (or very small value) of the epidemic threshold is balanced by the very low prevalence attained by the epidemic[46].

the global parameter i_∞, which identifies the density of infected individuals (or nodes in a network) in the infinite-time limit. In the limit of an infinitely large population, this density is zero below the threshold and assumes a finite value above the threshold. From this perspective we can consider the epidemic threshold as the critical point of the system and i_∞ as representing the order parameter characterizing the transition. Below the critical point the system relaxes in a frozen state with null dynamics—the healthy phase. Above this point, a dynamical state characterized by a macroscopic number of infected individuals sets in, defining an infected phase (Fig. 1).

Many other pioneering works in the area of social sciences use simple dynamical models to explore the emergence of macro-level collective behaviour as a function of the micro-level processes acting among the agents of a large population[9–11], and the incursions by statistical physicists in the area of social sciences have become very frequent (see, for example, the recent review by Castellano *et al.*[12]). A first class of models is represented by behavioural models where the attributes of agents are binary variables similar to Ising spins, as in the case of the voter model[13], the majority-rule model[14,15] and the Sznajd model[16]. In other instances, further realism has been introduced by the use of continuous opinion variables[17–19]. Along the path opened by Axelrod[11], models in which opinions or cultures are represented by vectors of cultural traits have introduced the notion of bounded confidence: an agent will not interact with any other agent, independently of their opinions, unless the opinions are close enough.

Finally, there is a vast class of models that focus on the analysis of diffusion processes as a tool to study phase transitions and emergent phenomena in simple models mimicking the routing of information packets in technological systems and networks. In this case the focus is on what lies behind the appearance of congestion and traffic self-similarity[20–26]. In traffic problems, one of the main issues is that the diffusion process is not random but determined by recurrent patterns, reinforcing mechanisms and routing strategies that represent formidable challenges to the modelling of systems[27]. Interestingly, it is the study of traffic dynamics in the Internet and the World Wide Web that has made clear the central role of networks and their structural properties in the understanding and characterization of dynamical processes in real-world systems.

Box 1 | The heterogeneous mean-field approach.

The heterogeneous mean-field approach generalizes, for the case of networks with arbitrary degree distribution, the equations describing the dynamical process, by considering degree-block variables grouping nodes within the same degree class k. If we consider the SIS model, the variables describing the system are i_k and s_k, which respectively represent the fraction of nodes with degree k in the infected and susceptible class. The evolution equation for the infected individual is

$$\frac{\mathrm{d}i_k(t)}{\mathrm{d}t} = -\mu i_k + \lambda[1-i_k(t)]k\Theta_k(t)$$

The first term just expresses the fact that any node in the infected state may recover with rate μ. The second term, which generates new infected individuals, is proportional to the probability of transmission λ, the degree k, the probability $1-i_k$ that a vertex with degree k is not infected, and the density Θ_k of infected neighbours of vertices of degree k, which is the probability of contacting an infected individuals. As we are still assuming a mean-field description of the system, the latter term is the average probability that any given neighbour of a vertex of degree k is infected. This quantity can be expressed as $\Theta_k(t) = \sum_{k'} P(k'|k) i_{k'}(t)$, which is the average over all possible degrees k' of the probability $P(k'|k)$ that any edge of a node of degree k is pointing to a node of degree k' times the probability $i_{k'}$ that the node is infected. This expression can be further simplified by considering a random network in which the conditional probability does not depend on the originating node. In this case we have that $P(k'|k) = k'P(k')/\langle k\rangle$, following simply from the fact that any edge has a probability proportional to the degree itself of pointing to a node with degree k' (see ref. 38). On substituting the expression for Θ in the main equation and adopting the early-epidemic assumption (that is, assuming that all second-order terms of i_k and r_k can be neglected), we readily recover the topology-dependent epidemic threshold result, $\lambda/\mu = \langle k\rangle/\langle k^2\rangle$.

Following the results obtained with the HMF assumption, a number of rigorous results that link the network topology to the epidemic threshold have been derived[53,57,58]. These results relate the epidemic threshold to the largest eigenvalue of the adjacency matrix of the network, showing that the HMF does not recover the correct behaviour for the SIS model when the degree distribution of the graph $P(k) \sim k^{-\gamma}$ has $\gamma > 3$. The rigorous results refer to quenched networks where the adjacency matrix is fixed in time. The HMF assumption instead, in its mean-field perspective, is equivalent to a system in which edges are continuously reshuffled so that the elements of the adjacency matrix are defined by the effective probabilities $k_i k_j / \sum_i k_i$ that two nodes i and j with degree k_i and k_j, respectively, are connected. This consideration clearly shows the shortcomings of the HMF assumption in the case of systems where the timescale of the transmission or infection is very short with respect to the duration of the contact and the adjacency matrix can be considered as quenched. The HMF can be considered, however, as a description of the system closer to reality in situations where the transmission occurs on rapidly varying networks; this is for instance the case for many influenza-like illnesses, where the infectious period is much longer than the duration of contacts responsible for the transmission[57].

Complex networks and dynamical processes

We live in an increasingly interconnected world, where infrastructures composed of different technological layers inter-operate

NATURE PHYSICS DOI:10.1038/NPHYS2160

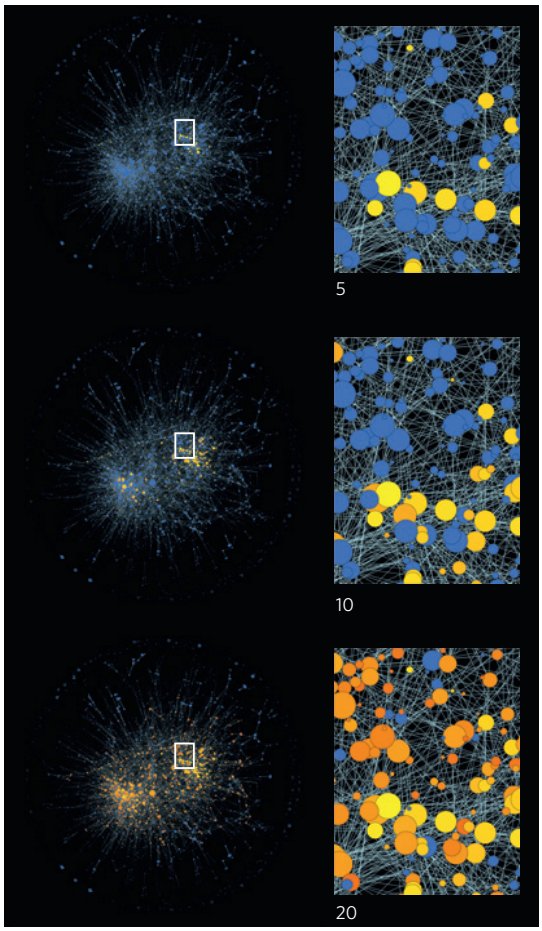

Figure 2 | Progression of an epidemic process. The progression of a susceptible–infected (SI) epidemic in a heavy-tailed network at three snapshots of the process, corresponding to time $t = 5$, 10 and 20, measured in unitary time integration steps of the model. The SI model assumes that infected nodes will spread the infection indefinitely to neighbours with rate α. In this case we know that the system is eventually completely infected, whatever the spreading rate of the infection. However, we can highlight the effect of topological fluctuations on the spreading hierarchy. Susceptible nodes are coloured blue and infected nodes are coloured from yellow to red according to the time of infection (red corresponding to later times). The size of a node is proportional to the node degree. In general, the first nodes to be infected are the large hubs with high degree, then the epidemic progresses in time by a dynamical cascade through degree classes, finally affecting low-degree nodes.

within the social component that drives their use and development. Examples are the Internet, the World Wide Web, mobile technologies, and transportation and mobility infrastructures[28–34]. The multiscale nature and complexity of these networks are crucial features in understanding and managing socio-technical systems and the dynamical processes occurring on top of them. For this reason, in the past decade, the study of models unfolding on complex networks has generated a body of work that includes results of conceptual and practical relevance[35–40]. The resilience of networks, their vulnerability to attacks, and their synchronization properties are all drastically affected by topological heterogeneities. Consensus formation, disease spreading and the accessibility of information can benefit or be impaired by the connectivity pattern of the population or infrastructure we are looking at. Network science has thus become pervasive in the study of complex systems and presented us with a number of surprising discoveries that have steered our way of thinking on dynamical processes in socio-technical systems.

One of the most important features affecting dynamical processes in real-world networks is the presence of dynamic self-organization and the lack of characteristic scales—typical hallmarks of complex systems[40–44]. Although those characteristics have long been acknowledged as a relevant factor in determining the properties of dynamical processes, many real-world networks exhibit levels of heterogeneity that were not anticipated until a few years ago. In particular, the various statistical distributions characterizing these networks are generally heavy-tailed, skewed, and varying over several orders of magnitude. This is a very peculiar feature, typical of many natural and artificial complex networks, characterized by virtually infinite degree fluctuations, where the degree k of a given node represents its number of connections to other nodes. In contrast to regular lattices and homogeneous graphs, characterized by nodes having a typical degree k close to the average $\langle k \rangle$, such networks are structured in a hierarchy where a few nodes (the hubs) have very high degree whereas the vast majority of nodes have lower degrees. This feature is usually manifest in a heavy-tailed degree distribution, often approximated by a power-law behaviour of the form $P(k) \sim k^{-\gamma}$, which implies a non-negligible probability of finding vertices with very high degree[40,42–44]. Furthermore, the presence of large-scale fluctuations associated with heavy-tail distributions is also observed for the intensity carried by the connecting links, transport flows, and other basic quantities that go beyond the connectivity description of the network[45].

The presence of large-scale fluctuations virtually acting at all scales of the network connectivity pattern calls for a mathematical analysis where the variables characterizing each node of the network explicitly enter the description of the system. Unfortunately, the general solution, handling the master equation of the system, is hardly, if ever, achievable—even for very simple dynamical processes. For this reason, a viable theoretical approach has to be based on techniques such as mean-field and deterministic continuum approximations, which usually provide the understanding of the basic phenomenology and phase diagram of the process under study. In both cases, the heterogeneous nature of the network-connectivity pattern is introduced by aggregating variables according to a degree-block formalism that assumes that all nodes with the same degree k are statistically equivalent[38,46,47]. This assumption allows the grouping of nodes in degree classes, yielding a convenient representation of the system. For instance, if for each node i we associate a corresponding state σ_i characterizing its dynamical state, a convenient representation of the system is provided by the quantity S_k, which indicates the number of nodes of degree k in the dynamical state $\sigma = s$, and the corresponding degree-block density of nodes of degree k in the state s

$$s_k = \frac{S_k}{V_k}$$

where V_k is the number of nodes of degree k. Finally, the global averages on the network are given by the expression

$$\rho_s = \sum_k P(k) s_k$$

where ρ_s is the probability that any given node is in the state s. This formalism defines a mean-field approximation within each degree class, relaxing, however, the overall homogeneity assumption on the degree distribution[38]. This framework, first introduced for the description of epidemic processes, is at the basis of the heterogeneous mean-field (HMF) approach that allows the analytical study of dynamical processes in complex networks by writing mean-field dynamical equations for each degree class variable. An example

Box 2 | The particle–network framework.

The particle–network framework extends the HMF approach to the case of a reaction–diffusion system in which particles (or individuals) diffuse on a network with arbitrary topology. A convenient representation of the system is therefore provided by quantities defined in terms of the degree k

$$N_k = \frac{1}{V_k}\sum_{i|k_i=k} N_i$$

where V_k is the number of nodes with degree k and the sums run over all nodes i having degree k_i equal to k. The degree-block variable N_k represents the average number of particles in nodes with degree k. The use of the HMF approach amounts to the assumption that nodes with degree k, and thus the particles in those nodes, are statistically equivalent. In this approximation the dynamics of particles randomly diffusing on the network is given by a mean-field dynamical equation expressing the variation in time of the particle subpopulations $N_k(t)$ in each degree block k. This can simply be written as:

$$\frac{\partial N_k}{\partial t} = -d_k N_k(t) + k\sum_{k'} P(k'|k) d_{k'k} N_{k'}(t)$$

The first term of the equation just considers that only a fraction of particles d_k moves out of the node per unit time. The second term accounts for particles diffusing from its neighbours into the node of degree k. This term is proportional to the number of links k times the average number of particles coming from each neighbour. The number of particles arriving from each neighbour is thus equal to that of particles $d_{k'k}N_{k'}(t)$ diffusing on any edge connecting a node of degree k' with a node of degree k, averaged over the conditional probability $P(k'|k)$ that an edge belonging to a node of degree k is pointing to a node of degree k'. Here the term $d_{k'k}$ is the diffusion rate along the edges connecting nodes of degree k and k'. The rate at which individuals leave a subpopulation with degree k is then given by $d_k = k\sum_{k'} P(k'|k) d_{kk'}$. The function $P(k'|k)$ encodes the topological connectivity properties of the network and allows the study of different topologies and mixing patterns. The above equation explicitly introduces the diffusion of particles into the description of the system. The equation can easily be generalized to particles with different states, and reacting among themselves, by adding a reaction term to the above equations. For instance, the generalization of the SIR model described in the main text would consider three types of particle, denoting infected, susceptible and recovered individuals. The reaction taking place among individuals in the same node would be the usual contagion process among susceptibles and infected individuals, and the spontaneous recovery of infected individuals.

The analysis of a simple diffusion process immediately indicates the importance of network topology. In a random network with arbitrary degree distribution, the stationary state reached by a swarm of particles diffusing with the same diffusive rate yields $N_k \sim k$ and the probability to find a single diffusing walker in a node of degree k is

$$p_k = \frac{k}{\langle k\rangle}\frac{1}{V}$$

where V is the total number of nodes in the network. This expression implies that the higher the degree of the nodes, the greater the probability to be visited by the walker. This observation has profound consequences for the way we can discover, retrieve and rank information in complex networks. The PageRank algorithm[117] is in this respect a major breakthrough, based on the idea that a viable ranking depends on the topological structure of the network, and is defined by essentially simulating the random surfing process on the web graph. The most important pages are simply those with the highest probability of being discovered if the web-surfer had infinite time to explore the web. Analogously, search processes can take advantage of this property using degree-biased searching algorithms that bias the routing of messages towards nodes with high degree[115,116].

of the HMF approach is given in Box 1 for the case of the SIS model. The HMF technique is often the first line of attack towards understanding the effects of complex connectivity patterns on dynamical processes and it has been used widely in a broad range of phenomena, although with different names and specific assumptions, depending on the problem at hand. Although it contains several approximations, the HMF approach readily shows that the heterogeneity found in the connectivity pattern of many networks may drastically affect the unfolding of the dynamical process.

The classic example for the effect of degree heterogeneity on dynamical processes in complex networks is epidemic spreading. The previously discussed result of the presence of an epidemic threshold in the SIR and SIS models is obtained under the assumption that each individual in the system has, to a first approximation, the same number of connections $k \simeq \langle k\rangle$. However, social heterogeneity and the existence of 'super-spreaders' have long been known in the epidemics literature[48]. Generally, it is possible to show that the reproductive rate R_0 is renormalized by fluctuations in the transmissibility or contact pattern as $R_0 \to R_0(1+f(\nu))$, where $f(\nu)$ is a positive and increasing function of the standard deviation ν of the individual transmissibility or connectivity pattern[49]. In particular, by generalizing the dynamical equations of the SIS model, the HMF approach yields that the disease will affect a finite fraction of the population only if $\beta/\mu \geq \langle k\rangle^2/\langle k^2\rangle$, that is the ratio between the first and second moments of the degree distribution[38,46,47]. This readily suggests that the topology of the network enters the very definition of the epidemic threshold. Furthermore, this implies that in heavy-tailed networks such that $\langle k^2\rangle \to \infty$, in the limit of infinite network size, we have a null epidemic threshold. Although this is not the case in any finite-size real-world network[50,51], larger heterogeneity levels lead to smaller epidemic thresholds (Fig. 1). This is an important result, which indicates that heterogeneous networks behave very differently from homogeneous networks with respect to physical and dynamical processes. Indeed, the heterogeneous connectivity pattern of networks affects also the dynamical progression of the epidemic process, which results in a striking hierarchical dynamics in which the infection propagates from higher-degree to lower-degree classes. The infection first takes control of the high-degree vertices in the network, then rapidly invades the network via a cascade through progressively lower-degree classes (Fig. 2). It also turns out that the time behaviour of epidemic outbreaks and the growth of the number of infected individuals are governed by a timescale τ proportional to the ratio between the first and second moment of the network's degree distribution, thus suggesting a velocity of progression that increases with the heterogeneity of the network[52].

The change of framework suggested by the network heterogeneity in the case of epidemic processes has triggered many studies

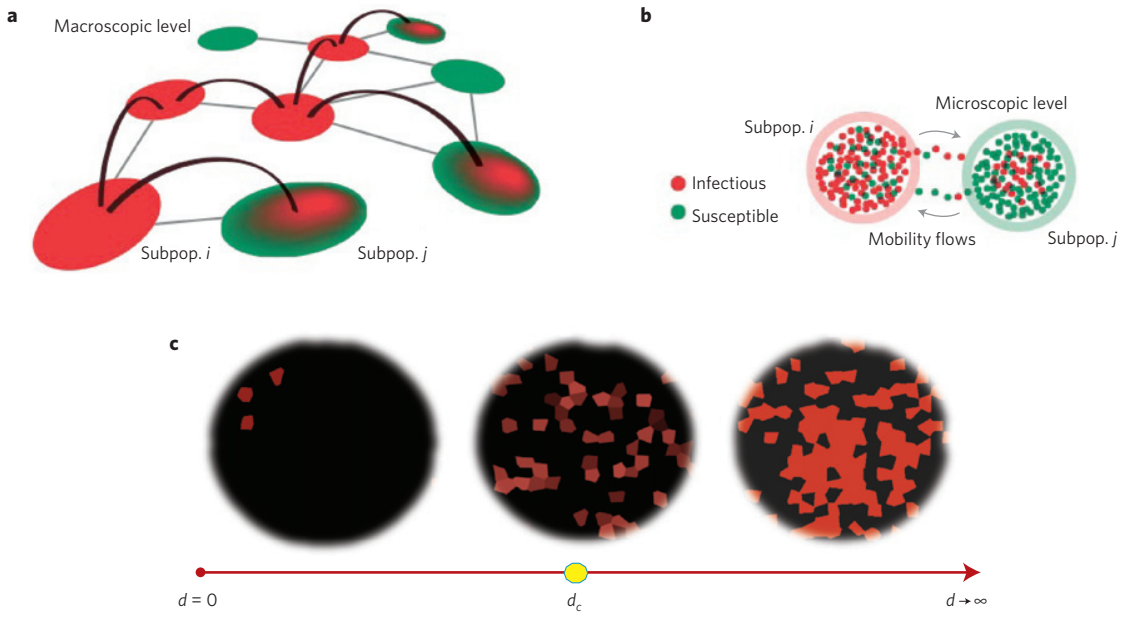

Figure 3 | Illustration of the global threshold in reaction–diffusion processes. a, Schematic of the simplified modelling framework based on the particle–network scheme. At the macroscopic level the system is composed of a heterogeneous network of subpopulations. The contagion process in one subpopulation (marked in red) can spread to other subpopulations as particles diffuse across subpopulations. **b**, At the microscopic level, each subpopulation contains a population of individuals. The dynamical process, for instance a contagion phenomena, is described by a simple compartmentalization (compartments are indicated by different coloured dots). Within each subpopulation, individuals can mix homogeneously, or according to a subnetwork, and can diffuse with rate d from one subpopulation to another, following the edges of the network. **c**, A critical value d_c of the diffusion strength for individuals or particles identifies a phase transition between a regime in which the contagion affects a large fraction of the system and one in which only a small fraction is affected (see the discussion in the text). Panels **a** and **b** reproduced from ref. 118.

aimed at providing a more rigorous analytical basis for the results obtained with the HMF and other approximate methods exploring different spreading models[53–58]. Equally important is the research activity concerned with developing dynamical ad hoc strategies for network protection; targeted immunization strategies and targeted prophylaxis that evolve with time might be particularly effective in the control of epidemics on heterogeneous patterns, compared with massive uniform vaccinations or stationary interventions[59–62]. Following the results on epidemic processes, an avalanche of studies addressed the study of the effect of the network's structure on the behaviour of the most widely used classes of dynamical processes. For instance, in the area of synchronization it has been shown that networks with heavy-tailed degree distributions, and therefore a large number of hubs, are more difficult to synchronize than homogeneous networks, a counterintuitive insight dubbed the paradox of heterogeneity[63–66]. In the case of packet-traffic routing, homogeneous networks have typically much larger congestion thresholds than heterogeneous graphs[67–69]. Finally, a wealth of surprising results, often overturning the common wisdom obtained by studies on regular networks, have been harvested on the voter and the Axelrod models[70–73], and many other models for the emergence of cooperation[38,74].

Reaction–diffusion processes and computational thinking

Although most approaches assume systems in which each node of the network corresponds to a single individual, it is of crucial importance for the study of many phenomena to provide a general understanding of processes where the multiple occupancy of nodes is a key feature. Examples of multiple occupancy are provided by chemical reactions, in which different molecules or atoms diffuse in space and may react whenever in close contact. Mechanistic metapopulation epidemic models, where particles represent people moving between different locations, and the routing of information packets in technological networks provides relevant examples in the case of socio-technical systems[75–79]. All those phenomena fall into the category of reaction–diffusion processes, where each node i is allowed to have any non-negative integer number of particles N_i so that the total particle population of the system is $N = \sum N_i$. The particle–network framework extends the heterogeneous mean-field approach to reaction–diffusion systems in networks with arbitrary degree distribution (Box 2). Particles diffuse along the edges connecting nodes, with a diffusion coefficient that depends on the node degree and/or other nodes' attributes. Within each node, particles may react according to different schemes characterizing the interaction dynamic of the system.

The consideration of complex networks in reaction–diffusion systems has broadened our knowledge of non-equilibrium reaction–diffusion systems in heterogeneous systems. For instance, the Turing mechanism represents a classical model for the formation of self-organized spatial structures in non-equilibrium activator–inhibitor systems. By studying the Turing mechanism[80] in systems with heterogeneous connectivity patterns it has been found that the relevant instabilities of the systems are localized in a set of vertices with degree inversely proportional to the characteristic scale of diffusion[81]. Interestingly, and contrary to other models and systems where the hubs are the playmakers, the segregation process takes place mainly in vertices of low degree.

Another interesting example is that of simple epidemic processes, such as the SIR model in a metapopulation context[79,82–90]. In this case, each node of the network is a subpopulation (ideally an urban area) connected by a transportation system (the edges of the network) that allows individuals to move from one subpopulation to another (Fig. 3). If we assume a diffusion rate d for each individual and consider that the single-population reproductive number of the SIR model is $R_0 > 1$, we can easily identify two different limits. If $d = 0$, any epidemic occurring in a given subpopulation

will remain confined; no individual could travel to a different subpopulation and spread the infection across the system. In the limit $d \to \infty$ we have that individuals are constantly wandering from one subpopulation to the other and the system is in practice equivalent to a well-mixed unique population. In this case, as $R_0 > 1$, the epidemic will spread across the entire system. A transition point between these two regimes occurs at a threshold value d_c of the diffusion rate, identifying a global invasion threshold. This threshold cannot be uncovered by continuous models as it is related to the stochastic diffusion rate of single individuals. Furthermore, the global invasion threshold is affected by the connectivity fluctuations of the metapopulation network. In particular, the greater the network heterogeneity, the smaller the value of the diffusion rate above which the epidemic may globally invade the metapopulation system. This result assumes a particular relevance, as it explains why travel restrictions seem to be highly ineffective in containing epidemics: the complexity and heterogeneity of present-day transport networks favour considerably the global spread of infectious diseases. Only infeasibly tight mobility restrictions, reducing global travel fluxes by 90% or more, would be effective[84,91,92].

Reaction–diffusion models lend themselves to the implementation of large-scale computer simulations (Monte-Carlo and individual-based simulations) that allow one to track microscopically the state of each node and the evolution of the dynamical process. At the most detailed level, the introduction of agent-based models has enabled the usual modelling perspective to be extended further by simulating the population and embedding environment on an individual-by-individual basis. An example is epidemic modelling, where spatially structured and agent-based models at various granularities (country, inter-city, intra-city) have been pushed to the computational limits with the integration of huge amount of data describing the flows of people and/or animals[93–97]. These models can generate results at an unprecedented level of detail and have been used successfully in the analysis and anticipation of real epidemics, such as the 2009 H1N1 pandemic[98,99]. Computer simulations thus become valuable in allowing both *in silico* experiments that would be infeasible in real systems and the capability to analyse and forecast scenarios. This computational approach is also helping to guide researchers in identifying typical nonlinear behaviour and tipping points[100] not accessible by analytical means, using the numerical simulations as a novel experimental workbench[101,102].

Co-evolution, timescale and control

Although in recent years our understanding of dynamical processes in complex networks has progressed at an exponential pace, there are still a number of major challenges that keep the research community actively engaged. The first challenge stems from the fact that the analysis of dynamical processes is generally performed in the presence of a timescale separation between the network evolution and the dynamical process unfolding on its structure. In one limit we can consider the network as quenched in its connectivity pattern, thus evolving on a timescale that is much longer that the dynamical process itself. In the other limiting case, the network evolves on a timescale much shorter than the dynamical process, which thus effectively disappears from the definition of the interaction among individuals such that this interaction can be conveniently replaced by effective random coupling. Although the timescale separation is extremely convenient with a view to the numerical and analytical tractability of the models, networks generally evolve on a timescale that might be comparable to that of the dynamical process. Furthermore, the network properties used in defining models generally represent a time-integrated static snapshot of the system. However, in many systems the timing and duration of interactions define processes on a timescale very different from, and often conflicting with, those of the time-integrated view. This highlights the importance of considering the concurrency of network evolution and dynamical processes in realistic models to avoid misleading conclusions[103–106].

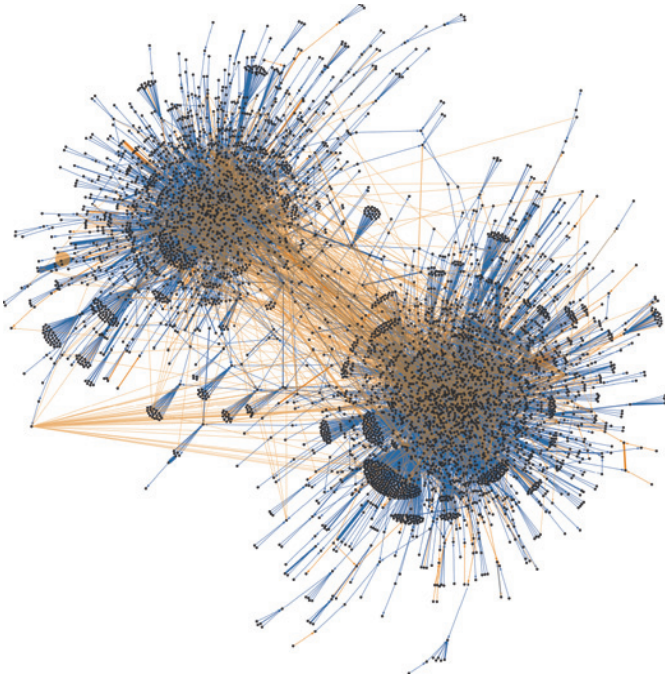

Figure 4 | Visualization of the dynamical network generated by Twitter interactions. Twitter is a microblogging tool that allows users to post and relay ('re-tweet') short messages. The topic of the message is signalled by short identifiers (@mentions, #hash-tags and urls). This feature allows one to trace the spreading of specific discussion topics (also called memes). The figure shows the diffusion network for the tag #gop. Each node corresponds to an individual user. Blue edges represent re-tweets and orange edges represent mentions. Two communities are clearly visible, corresponding to politically left- and right-leaning users[113]. Communications between the two communities take place primarily through the use of mentions, while within a group communication occurs through re-tweets. The figure, obtained using the Truthy infrastructure[114], clearly exemplifies the co-evolution of the communication network with the spreading process.

A second challenge is the co-evolution of networks with the dynamical process. Access to the mathematical and statistical laws that characterize the interplay and feedback mechanisms between the network evolution and the dynamical processes is extremely important, especially in social systems, where the adaptive nature of agents is of paramount importance[106–108]. The spreading of an opinion is affected by the interaction among individuals, but the presence and/or establishment of interaction among individuals is affected by their opinion. This issue is increasingly relevant in the area of the modern social networks populating the information-technology ecosystem, such as those defined by the Facebook and Twitter applications. In this case the network and the spread of information cannot be defined in isolation, because of rapidly changing interactions and modes of communication that depend on the type of information exchanged and the adaptive behaviour of individuals (Fig. 4).

The adaptive behaviour of individuals to the dynamical processes they are involved in represents another modelling challenge, as it calls for the understanding of the feedback among different and competing dynamical processes. For instance, relatively little systematic work has been done to provide coupled behaviour–disease models able to close the feedback loop between

behavioural changes triggered in the population by an individual's perception of the disease spread and the actual disease spread[109,110]. Similar issues arise in many areas where we find competing processes of adaptation and awareness to information or knowledge spreading in a population[111].

Finally, the overall goal is not only to understand complex systems, mathematically describe their structure and dynamics, and predict their behaviour, but also to control their dynamics. Also in this case, although control theory offers a large set of mathematical tools for steering engineered and natural systems, we are just taking the first steps towards a full understanding of how the network heterogeneities influence our ability to control the network dynamics and how the network evolution impacts controllability[112].

Conclusions

There are no doubts that a complete understanding of complex socio-technical systems requires diving into the specifics of each system by adopting a domain-specific perspective. Data-driven models, however, are generating new questions, the answers to which should preferably be analytical and applicable to a wide range of systems. What are the fundamental limits to predictability with computational modelling? How does our understanding depend on the level of accuracy of our description and knowledge of the state of the system? The research community needs, now more than ever, the kind of basic theoretical understanding that would help discriminate between what is relevant and what is superfluous in the description of socio-technical systems. This is a crucial endeavour if we want to complement data-driven approaches with a conceptual understanding that would help guide the management, prediction and control of dynamical processes in complex systems—a conceptual understanding that necessarily descends from the study of the dynamical models and processes presented here.

References

1. Keeling, M. J. & Rohani, P. *Modeling Infectious Diseases in Humans and Animals* (Princeton Univ. Press, 2008).
2. Goffman, W. & Newill, V. A. Generalization of epidemic theory: An application to the transmission of ideas. *Nature* **204,** 225–228 (1964).
3. Rapoport, A. Spread of information through a population with socio-structural bias: I. Assumption of transitivity. *Bull. Math. Biol.* **15,** 523–533 (1953).
4. Tabah, A. N. Literature dynamics: Studies on growth, diffusion, and epidemics. *Annu. Rev. Inform. Sci. Technol.* **34,** 249–286 (1999).
5. Lloyd, A. L. & May, R. M. How viruses spread among computers and people. *Science* **292,** 1316–1317 (2001).
6. Grassberger, P. On the critical behavior of the general epidemic process and dynamical percolation. *Math. Biosci.* **63,** 157–172 (1983).
7. Harris, T. E. Contact interactions on a lattice. *Ann. Prob.* **2,** 969–988 (1974).
8. Marro, J. & Dickman, R. *Nonequilibrium Phase Transitions in Lattice Models* (Cambridge Univ. Press, 1999).
9. Granovetter, M. Threshold models of collective behavior. *Am. J. Sociol.* **83,** 1420–1443 (1978).
10. Nowak, A., Szamrej, J. & Latané, B. From private attitude to public opinion: A dynamic theory of social impact. *Psychol. Rev.* **97,** 362–376 (1990).
11. Axelrod, R. *The Complexity of Cooperation* (Princeton Univ. Press, 1997).
12. Castellano, C., Fortunato, S. & Loreto, V. Statistical physics of social dynamics. *Rev. Mod. Phys.* **81,** 591–646 (2009).
13. Krapivsky, P. L. Kinetics of monomer–monomer surface catalytic reactions. *Phys. Rev. A* **45,** 1067–1072 (1992).
14. Galam, S. Minority opinion spreading in random geometry. *Eur. Phys. J. B* **25,** 403–406 (2002).
15. Krapivsky, P. L. & Redner, S. Dynamics of majority rule in two-state interacting spin systems. *Phys. Rev. Lett.* **90,** 238701 (2003).
16. Sznajd-Weron, K. & Sznajd, J. Opinion evolution in closed community. *Int. J. Mod. Phys. C* **11,** 1157–1165 (2000).
17. Deffuant, G., Neau, D., Amblard, F. & Weisbuch, G. Mixing beliefs among interacting agents. *Adv. Complex Syst.* **3,** 87–98 (2000).
18. Hegselmann, R. & Krause, U. Opinion dynamics and bounded confidence models, analysis and simulation. *J. Art. Soc. Soc. Sim.* **5,** 2 (2002).
19. Ben-Naim, E., Krapivsky, P. L. & Redner, S. Bifurcations and patterns in compromise processes. *Physica D* **183,** 190–204 (2003).
20. Leland, W. E., Taqqu, M. S., Willinger, W. & Wilson, D. V. On the self-similar nature of Ethernet traffic. *IEEE/ACM Trans. Netw.* **2,** 1–15 (1994).
21. Csabai, I. 1/f noise in computer network traffic. *J. Phys. A* **27,** L417–L42 (1994).
22. Solé, R. V. & Valverde, S. Information transfer and phase transitions in a model of internet traffic. *Physica A* **289,** 595–605 (2001).
23. Willinger, W., Govindan, R, Jamin, S., Paxson, V. & Shenker, S. Scaling phenomena in the Internet: Critically examining criticality. *Proc. Natl Acad. Sci. USA* **99,** 2573–2580 (2002).
24. Valverde, S. & Solé, R. V. Internet's critical path horizon. *Eur. Phys. J. B* **38,** 245–252 (2004).
25. Tadić, B., Thurner, S. & Rodgers, G. J. Traffic on complex networks: Towards understanding global statistical properties from microscopic density fluctuations. *Phys. Rev. E* **69,** 036102 (2004).
26. Crovella, M. E. & Krishnamurthy, B. *Internet Measurements: Infrastructure, Traffic and Applications* (John Wiley, 2006).
27. Helbing, D. Traffic and related self-driven many particle systems. *Rev. Mod. Phys.* **73,** 1067–1141 (2001).
28. Albert, R., Jeong, H. & Barabási, A-L. Internet: Diameter of the World-Wide Web. *Nature* **401,** 130–131 (1999).
29. Pastor-Satorras, R. & Vespignani, A. *Evolution and Structure of the Internet: A Statistical Physics Approach* (Cambridge Univ. Press, 2004).
30. Brockmann, D., Hufnagel, L. & Geisel, T. The scaling laws of human travel. *Nature* **439,** 462–465 (2006).
31. Onnela, J-P. *et al.* Structure and tie strengths in mobile communication networks. *Proc. Natl Acad. Sci. USA* **104,** 7332–7337 (2007).
32. González, M. C., Hidalgo, C. A. & Barabási, A-L. Understanding individual human mobility patterns. *Nature* **453,** 779–782 (2008).
33. Lazer, D *et al.* Life in the network: The coming age of computational social science. *Science* **323,** 721–723 (2009).
34. Vespignani, A. Predicting the behavior of tecno-social systems. *Science* **325,** 425–428 (2009).
35. Albert, R. & Barabási, A-L. Statistical mechanics of complex networks. *Rev. Mod. Phys.* **74,** 47–97 (2002).
36. Boccaletti, S. *et al.* Complex networks: Structure and dynamics. *Phys. Rep.* **424,** 175–308 (2006).
37. Dorogovtsev, S. N., Goltsev, A. V. & Mendes, J. F. F. Critical phenomena in complex networks. *Rev. Mod. Phys.* **80,** 1275–1335 (2008).
38. Barrat, A., Barthelemy, M. & Vespignani, A. *Dynamical Processes on Complex Networks* (Cambridge Univ. Press, 2008).
39. Cohen, R. & Havlin, S. *Complex Networks: Structure, Robustness and Function* (Cambridge Univ. Press, 2010).
40. Newman, M. E. J. *Networks: An Introduction* (Oxford Univ. Press, 2010).
41. Watts, D. J. & Strogatz, S. H. Collective dynamics of 'small-world' networks. *Nature* **393,** 440–442 (1998).
42. Barabási, A-L. & Albert, R. Emergence of scaling in random networks. *Science* **286,** 509–512 (1999).
43. Dorogovtsev, S. N. & Mendes, J. F. F. *Evolution of Networks: From Biological Nets to the Internet and WWW* (Oxford Univ. Press, 2003).
44. Amaral, L. A. N., Scala, A., Barthlemy, M. & Stanley, H. E. Classes of small-world networks. *Proc. Natl Acad. Sci. USA* **97,** 11149–11154 (2005).
45. Barrat, A., Barthlemy, M., Pastor-Satorras, R. & Vespignani, A. The architecture of complex weighted networks. *Proc. Natl Acad. Sci. USA* **101,** 3747–3752 (2004).
46. Pastor-Satorras, R. & Vespignani, A. Epidemic spreading in scale-free networks. *Phys. Rev. Lett.* **86,** 3200–3203 (2001).
47. Moreno, Y., Pastor-Satorras, R. & Vespignani, A. Epidemic outbreaks in complex heterogeneous networks. *Eur. Phys. J. B* **26,** 521–529 (2002).
48. Hethcote, H. W. & Yorke, J. A. Gonorrhea: Transmission and control. *Lect. Notes Biomath.* **56,** 1–105 (1984).
49. Anderson, R. M. & May, R. M. *Infectious Diseases in Humans* (Oxford Univ. Press, 1992).
50. May, R. M. & Lloyd, A. L. Infection dynamics on scale-free networks. *Phys. Rev. E* **64,** 066112 (2001).
51. Pastor-Satorras, R. & Vespignani, R. Epidemic dynamics in finite size scale-free networks. *Phys. Rev. E* **65,** 035108(R) (2002).
52. Barthelemy, M., Barrat, A., Pastor-Satorras, R. & Vespignani, A. Velocity and hierarchical spread of epidemic outbreaks in scale-free networks. *Phys. Rev. Lett.* **92,** 178701 (2004).
53. Wang, Y., Chakrabarti, D., Wang, G. & Faloutsos, C, in *Proc. 22nd International Symposium on Reliable Distributed Systems (SRDS'03)* 25–34 (IEEE, 2003).
54. Boguna, M., Pastor-Satorras, R. & Vespignani, A. Absence of epidemic threshold in scale-free networks with degree correlations. *Phys. Rev. Lett.* **90,** 028701 (2003).
55. Castellano, C. & Pastor-Satorras, R. Routes to thermodynamic limit on scale-free networks. *Phys. Rev. Lett.* **100,** 148701 (2008).
56. Chatterjee, S. & Durrett, R. Contact processes on random graphs with power law degree distributions have critical value 0. *Ann. Probab.* **37,** 2332–2356 (2009).

57. Castellano, C. & Pastor-Satorras, R. Thresholds for epidemic spreading in networks. *Phys. Rev. Lett.* **105,** 218701 (2010).
58. Durrett, R. Some features of the spread of epidemics and information on a random graph. *Proc. Natl Acad. Sci. USA* **107,** 4491–4498 (2010).
59. Pastor-Satorras, R. & Vespignani, A. Immunization of complex networks. *Phys. Rev. E* **65,** 036104 (2001).
60. Cohen, R., Havlin, S. & Ben-Avraham, D. Efficient immunization strategies for computer networks and populations. *Phys. Rev. Lett.* **91,** 247901 (2003).
61. Holme, P. Efficient local strategies for vaccination and network attack. *Europhys. Lett.* **68,** 908–914 (2004).
62. Goldenberg, J., Shavitt, Y., Shir, E. & Solomon, S. Distributive immunization of networks against viruses using the 'honey-pot' architecture. *Nature Phys.* **1,** 184–188 (2005).
63. Motter, A. E., Zhou, C. S. & Kurths, J. Enhancing complex-network synchronization. *Europhys. Lett.* **69,** 334–340 (2005).
64. Motter, A. E., Zhou, C. S. & Kurths, J. Network synchronization, diffusion, and the paradox of heterogeneity. *Phys. Rev. E* **71,** 016116 (2005).
65. Gómez-Gardeñes, J., Campillo, M., Floria, L. M. & Moreno, Y. Dynamical organization of cooperation in complex topologies. *Phys. Rev. Lett.* **98,** 108103 (2007).
66. Korniss, G. Synchronization in weighted uncorrelated complex networks in a noisy environment: Optimization and connections with transport efficiency. *Phys. Rev. E* **75,** 051121 (2007).
67. Arenas, A., Díaz-Guilera, A. & Guimerà, R. Communication in networks with hierarchical branching. *Phys. Rev. Lett.* **86,** 3196–3199 (2001).
68. Guimerà, R., Arenas, A., Díaz-Guilera, A. & Giralt, F. Dynamical properties of model communication networks. *Phys. Rev. E* **66,** 026704 (2002).
69. Sreenivasan, S., Cohen, R., López, E., Toroczkai, Z. & Stanley, H. E. Structural bottlenecks for communication in networks. *Phys. Rev. E* **75,** 036105 (2007).
70. Castellano, C., Loreto, V., Barrat, A., Cecconi, F. & Parisi, D. Comparison of voter and Glauber ordering dynamics on networks. *Phys. Rev. E* **71,** 066107 (2005).
71. Sood, V. & Redner, S. Voter model on heterogeneous graphs. *Phys. Rev. Lett.* **94,** 178701 (2005).
72. Suchecki, K., Eguíluz, V. M. & San Miguel, M. Conservation laws for the voter model in complex networks. *Europhys. Lett.* **69,** 228–234 (2005).
73. Klemm, K., Eguíluz, V. M., Toral, R. & San Miguel, M. Nonequilibrium transitions in complex networks: A model of social interaction. *Phys. Rev. E.* **67,** 026120 (2003).
74. Santos, F. C., Pacheco, J. M. & Lenaerts, T. Evolutionary dynamics of social dilemmas in structured heterogeneous populations. *Proc. Natl Acad. Sci. USA* **103,** 3490–3494 (2006).
75. van Kampen, N. G. *Stochastic Processes in Physics and Chemistry* (North-Holland, 1981).
76. Bolker, B. M. & Grenfell, T. Chaos and biological complexity in measles dynamics. *Proc. Trans. R. Soc. Lond. B* **251,** 75–81 (1993).
77. Keeling, M. J. & Rohani, P. Estimating spatial coupling in epidemiological systems: A mechanistic approach. *Ecol. Lett.* **5,** 20–29 (2002).
78. Sattenspiel, L. & Dietz, K. A structured epidemic model incorporating geographic mobility among regions. *Math. Biosci.* **128,** 71–91 (1995).
79. Watts, D., Muhamad, R., Medina, D. C. & Dodds, P. S. Multiscale resurgent epidemics in a hierarchical metapopulation model. *Proc. Natl Acad. Sci. USA* **102,** 11157–11162 (2005).
80. Turing, A. M. The chemical basis of morphogenesis. *Phil. Trans. R. Soc. Lond.* **B237,** 37–72 (1952).
81. Nakao, H. & Mikhailov, A. S. Turing patterns in network-organized activator-inhibitor systems. *Nature Phys.* **6,** 544–550 (2010).
82. Colizza, V., Pastor-Satorras, R. & Vespignani, A. Reaction–diffusion processes and metapopulation models in heterogeneous networks. *Nature Phys.* **3,** 276–282 (2007).
83. Colizza, V. & Vespignani, A. Invasion threshold in heterogeneous metapopulation networks. *Phys. Rev. Lett.* **99,** 148701 (2007).
84. Colizza, V. & Vespignani, A. Epidemic modeling in metapopulation systems with heterogeneous coupling pattern: Theory and simulations. *J. Theor. Biol.* **251,** 450–467 (2008).
85. Barthélemy, M., Godrèche, C. & Luck, J-M. Fluctuation effects in metapopulation models: Percolation and pandemic threshold. *J. Theor. Biol.* **267,** 554–564 (2010).
86. Saldana, J. Continuous-time formulation of reaction–diffusion processes on heterogeneous metapopulations. *Phys. Rev. E* **78,** 012902 (2008).
87. Ni, S. & Weng, W. Impact of travel patterns on epidemic dynamics in heterogeneous spatial metapopulation networks. *Phys. Rev. E* **79,** 016111 (2009).
88. Ben-Zion, Y., Cohena, Y. & Shnerba, N. M. Modeling epidemics dynamics on heterogenous networks. *J. Theor. Biol.* **264,** 197–204 (2010).
89. Balcan, D. & Vespignani, A. Phase transitions in contagion processes mediated by recurrent mobility patterns. *Nature Phys.* **7,** 581–586 (2011).
90. Belik, V., Geisel, T. & Brockmann, D. Natural human mobility patterns and spatial spread of infectious diseases. *Phys. Rev. X* **1,** 011001 (2011).
91. Cooper, B. S., Pitman, R. J., Edmunds, W. J. & Gay, N. J. Delaying the international spread of pandemic influenza. *PLoS Med.* **3,** e12 (2006).
92. Hollingsworth, T. D., Ferguson, N. M. & Anderson, R. M. Will travel restrictions control the international spread of pandemic influenza? *Nature Med.* **12,** 497–499 (2006).
93. Hufnagel, L., Brockmann, D. & Geisel, T. Forecast and control of epidemics in a globalized world. *Proc. Natl Acad. Sci. USA* **101,** 15124–15129 (2004).
94. Eubank, S. *et al.* Modelling disease outbreaks in realistic urban social networks. *Nature* **429,** 180–184 (2004).
95. Longini, I. M. *et al.* Containing pandemic infleunza at the source. *Science* **309,** 1083–1087 (2005).
96. Ferguson, N. M. *et al.* Strategies for containing an emerging influenza pandemic in Southeast Asia. *Nature* **437,** 209–211 (2005).
97. Colizza, V., Barrat, A., Barthlemy, M., Valleron, M. A. J. & Vespignani, A. Modeling the worldwide spread of pandemic influenza: Baseline case and containment interventions. *PLoS Med.* **4,** e13 (2007).
98. 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 Med.* **7,** 45 (2009).
99. Merler, S., Ajelli, M., Pugliese, A. & Ferguson, N. M. Determinants of the spatiotemporal dynamics of the 2009 H1N1 pandemic in Europe: Implications for real-time modelling. *PLoS Comput. Biol.* **7,** e1002205 (2011).
100. Gladwell, M. *The Tipping Point: How Little Things Can Make a Big Difference* (Little, Brown and Company, 2002).
101. Helbing, D. & Yu, W. The outbreak of cooperation among success-driven individuals under noisy condition. *Proc. Natl Acad. Sci. USA* **106,** 3680–3685 (2009).
102. Xie, J. *et al.* Social consensus through the influence of commited minorities. *Phys. Rev. E* **84,** 011130 (2011).
103. Morris, M. & Kretzschmar, M. Concurrent partnerships and the spread of HIV. *AIDS* **11,** 641–648 (1997).
104. Moody, J. The importance of relationship timing for diffusion: Indirect connectivity and STD infection risk. *Soc. Forces* **81,** 25–56 (2002).
105. Isella, L. *et al.* What's in a crowd? Analysis of face-to-face behavioral networks. *J. Theor. Biol.* **271,** 166–180 (2011).
106. Volz, E. & Meyers, L. A. Epidemic thresholds in dynamic contact networks. *J. R. Soc. Interface* **6,** 233–241 (2009).
107. Holme, P. & Newman, M. E. J. Nonequilibrium phase transition in the coevolution of networks and opinions. *Phys. Rev. E* **74,** 056108 (2006).
108. Centola, D., Gonzalez-Avella, J. C., Eguiluz, V. M. & San Miguel, M. Homophily, cultural drift, and the co-evolution of cultural groups. *J. Conflict Resolution* **51,** 905–929 (2007).
109. Funk, S., Salathé, M. & Jansen, V. A. A. Modelling the inuence of human behaviour on the spread of infectious diseases: A review. *J. R. Soc. Interface* **7,** 1247–1256 (2010).
110. Perra, N., Balcan, D., Goncalves, B. & Vespignani, A. Towards a characterization of behavior–disease models. *PLoS ONE* **6,** e23084 (2011).
111. Bauch, C. T. & Earn, D. J. Vaccination and the theory of games. *Proc. Natl Acad. Sci. USA* **101,** 13391–13394 (2004).
112. Liu, Y-Y., Slotine, J-J. & Barabasi, A-L. Controllability of complex networks. *Nature* **473,** 167–173 (2011).
113. Conover, M. *et al. Proc. 5th International Conference on Weblogs and Social Media (ICWSM)* 89–96 (2011).
114. Ratkiewicz, J. *et al. Proc. 20th International Conference Companion on World Wide Web (WWW '11)* 249–252 (ACM, 2001).
115. Kim, B. J., Yoon, C. N., Han, S. K. & Jeong, H. Path finding strategies in scale-free networks. *Phys. Rev. E* **65,** 027103 (2002).
116. Adamic, L. A., Lukose, R. M., Puniyani, A. R. & Huberman, B. A. Search in power-law networks. *Phys. Rev. E* **64,** 046135 (2001).
117. Brin, S. & Page, L. The anatomy of a large-scale hypertextual Web search engine. *Comput. Netw. ISDN Syst.* **30,** 107–117 (1998).
118. Bajardi, P. *et al.* Human mobility networks, travel restrictions, and the global spread of 2009 H1N1 pandemic. *PLoS ONE* **6,** e16591 (2011).

Acknowledgements

I thank B. Goncalves and N. Perra for their help with the figures and a critical reading of the manuscript. This work has been partially funded by the NIH R21-DA024259, DTRA-1-0910039 and NSF CCF-1101743 and NSF CMMI-1125095 awards. The work has been also partly sponsored by the Army Research Laboratory and was accomplished under Cooperative Agreement Number W911NF-09-2-0053. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Army Research Laboratory or the US Government.

Additional information

The authors declare no competing financial interests. Reprints and permissions information is available online at http://www.nature.com/reprints.