Eur. Phys. J. Special Topics **222**, 1389–1401 (2013)

DOI: 10.1140/epjst/e2013-01933-9

THE EUROPEAN PHYSICAL JOURNAL SPECIAL TOPICS

Regular Article

Towards understanding the behavior of physical systems using information theory

Rick Quax[1,a], Andrea Apolloni[2], and Peter M.A. Sloot[1,3,4]

[1] Computational Science, University of Amsterdam, Science Park 904, 1098 XH Amsterdam, The Netherlands
[2] I.S.I. Foundation, 11/C Via Alassio, 10126 Turin, Italy
[3] National Research University of Information Technologies, Mechanics and Optics (ITMO), Kronverkskiy 49, 197101 Saint Petersburg, Russia
[4] Nanyang Technological University, 50 Nanyang Avenue, 639798 Singapore

Received 3 June 2013 / Received in final form 10 July 2013
Published online 13 September 2013

Abstract. One of the goals of complex network analysis is to identify the most influential nodes, i.e., the nodes that dictate the dynamics of other nodes. In the case of autonomous systems or transportation networks, highly connected hubs play a preeminent role in diffusing the flow of information and viruses; in contrast, in language evolution most linguistic norms come from the peripheral nodes who have only few contacts. Clearly a topological analysis of the interactions alone is not sufficient to identify the nodes that drive the state of the network. Here we show how information theory can be used to quantify how the dynamics of individual nodes propagate through a system. We interpret the state of a node as a storage of information about the state of other nodes, which is quantified in terms of Shannon information. This information is transferred through interactions and lost due to noise, and we calculate how far it can travel through a network. We apply this concept to a model of opinion formation in a complex social network to calculate the impact of each node by measuring how long its opinion is remembered by the network. Counter-intuitively we find that the dynamics of opinions are not determined by the hubs or peripheral nodes, but rather by nodes with an intermediate connectivity.

1 Introduction

In a complex network of interactions among dynamical nodes, some nodes may be more important than others. These "driver nodes" drive the dynamics of the network [29]. The two main factors that determine the driver nodes are the dynamics of the nodes and the topology of the network. The dynamics dictate how nodes influence each other, and the topology propagates this influence to more distant nodes. The topology and node dynamics are therefore complementary and must be considered

[a] e-mail: r.quax@uva.nl

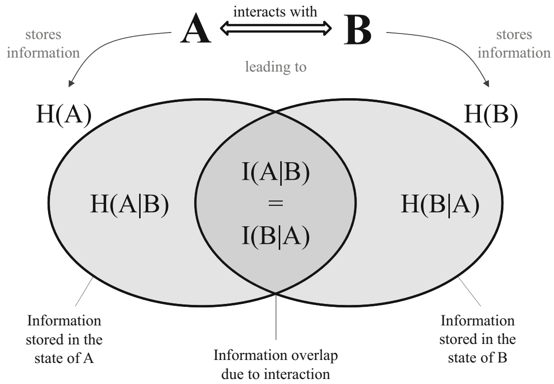

Fig. 1. Venn-diagram of two elements which store information about each other. H denotes Shannon's entropy and I denotes mutual information.

together: understanding the dynamics of the network requires an understanding of the interplay between the rules of dynamics and the topology of the interactions.

A general theory to quantify the influence of a single element on the dynamics of the system is lacking. A common approach is to remove an element from the system and measure the change of behavior of the remaining elements. For instance, in an N-body gravity simulation of planets one could define influence as a measure of how much one planet incurs a deviating path of another planet [32,33]. In social networks, a person's social influence could be taken as the number of people that would not change opinion of this person were not present [4,16]. In protein-protein interactions, the influence of one type of proteins could be the difference in concentrations of all other proteins in its absence ("knock-out" experiments) [22,23].

The problem with such definitions of "influence" is that they typically have arbitrary units and lack well-defined procedures for summing up and subtracting. One cannot sum up these influences in some way to decide what fraction of the state of one node is dictated by interactions with other nodes, and what fraction is noise or randomness. It is also difficult in general to disentangle the individual contributions of nodes from their combined effect.

Moreover, a list of interventions and system responses is generally not sufficient to understand the natural behavior of a system. For instance, a knock-out experiment removes a specific protein from an organism and observes the average difference in concentration of all other proteins. However, a knock-out is typically not part of the natural behavior of a network of interacting proteins, and the problem of understanding natural behavior transfers recursively to the set of remaining proteins.

Here we show that the illusive term "influence" can be concretized into an amount of *bits* for any system. The state of each element is capable of storing a certain number of bits. For instance, if a person can have one out of two opinions then each person's opinion is capable of storing 1 classical bit [12]. This means that each opinion can at most store 1 bit of information about the opinion of other persons in the social network. This provides an absolute measure of influence: person A influences person B as much as the number of bits that person B stores about person A. See Figure 1.

How do these bits from person A end up in the state of person B? The answer lies in the network of interactions. Two persons interact by pursuading each other of their own opinion. In other words, the opinion of person A can exert influence on the opinion of person B, and vice versa. This means that the opinion of B is partially a reflection of the opinion of A, which can be quantified as the number of bits stored in the state of B which it copied from A. This fraction would be zero if there would be no interaction between A and B, so we can say that bits of information are transported through interactions. This phenomenon is transitive: bits can travel from A to B, then from B to C, etc., each time diminished by randomness and coinciding with influence (information) from others.

In Sect. 2 we review the literature about information processing in dynamical systems, as well as the modeling of opinion forming and innovation diffusion in social networks. In Sect. 3 we show by example how to quantify the influence of an element in amounts of bits of information, and the difference between static and dynamical information. Then in Sect. 4 we show how to apply this concept to a model of opinion forming in social networks, of which the results are presented in Sect. 5. We conclude in Sect. 6.

2 Related work

In this article we study the diffusion of information in a model of opinion formation in a social network. Therefore we first review the literature on information processing in dynamical systems, and then we review the modeling of social networks and the process of forming opinions.

2.1 Information processing in dynamical systems

Information theory is traditionally used for statistical inference, where an external observer attempts to describe the state of a system and its behavior [12,13,15,41]. In information processing of dynamical systems, however, there is no external observer. Each component (particle, agent, or element) can instead be considered as an observer which stores information about the state and behavior of other components in the system.

How this information is processed in a system is still unknown, but energy and information are related concepts. Jaynes [24] studied this connection between statistical mechaniscs and information theory. He suggests a re-intepretation that relates thermodynamic entropy and information-theory entropy through the second law of thermodynamics. Landauer [25,26] goes on to show that erasing information leads to increasing entropy. Specifically, losing 1 bit of information corresponds to the exchange of $kTln2$ heat [7,8], and information about a system that is erased and truly lost cannot be retrieved [17]. This supports our view of information as a physical quantity that is processed in dynamical systems.

Lloyd [30,31] considers each physical system as a quantum computer: "Nature intrinsically computes. (...) Every physical system registers information, and just by evolving in time, by doing its thing, it changes that information, transforms that information, or, if you like, processes that information." The starting point of our research is to find an answer to the question: can the behavior of a physical system be understood from how it processes information? Here, information refers to the extent that (local) topology and/or the states of nodes in a network are remembered as they evolve by some dynamical process. The dynamical process can involve either the state of the constituents, either can involve the network structure itself, or both. Once the storage and transfer of this information is quantified it may provide a useful tool for studying complex systems.

Applications of the concept of information processing is scarce. Various researchers such as Ebeling [18] uses higher order Shannon entropies to measure the uncertainty of predicting the next system's state by observing n foregoing states. This is a measure of memory of a dynamical system calculated in the same way as the entropy of text

[18]. Crutchfield et al. employ excess entropy and entropy rate to quantify the amount of information that is stored in a dynamical system on a lattice [14,15] and how it is transferred in time and space [13]. Benett [7] reviewed the connection between statistical mechanics and information theory.

2.2 Modeling collective phenomena

2.2.1 Scale-free networks

The first example of a social network dates back during the '60 with the Erdos-Renyi model [19]. The basis of this model is that all the nodes are equivalent, and that the probability of establishing a link between two nodes is p for every possible couple. As a consequence, the degree distribution can be drawn from a binomial distribution, and it is peaked around the average value so that most of the nodes have average degree $\langle k \rangle = Np$. These random networks have been the paradigm for over 50 years.

With the advent of new technologies for collecting data, researchers collected evidence that many real-world networks have a different topology. Examples are technological networks such as the internet and the World Wide Web; scientific and social networks such as co-authorhsips, citations, actors, and prostitution; and communication networks such as telephone calls and mail networks. These networks exhibit a heavy-tailed distribution [5,35]. They are often modeled with a scale-free degree distribution $P(k) \propto k^{-\gamma}$, which specifies the probability that a node has degree k. As a consequence, the network consists of many peripheral nodes and only few hubs. Because of this heterogeneity, the properties of the network are different from the random networks.

Nodes with different degree play different roles in the communication flow and in enhancing the resilience of the network [1]: hubs collect information from peripheral nodes and distribute it to other nodes, which could otherwise be reached only through a long path in the network. Another characteristic is that scale free networks are compact: the diameter is of order $\ln \ln N$ compared to N^a typical of random and regular networks. Also, additional structure such as clustering can be present in the network [35].

Despite the similarity in degree distribution and compactness, technological and social networks differ regarding assortativity, i.e. the tendency of nodes with similar topological properties of being connected. While in technological networks, hubs are connected to nodes with few degrees, in social networks, gregarious nodes tend to establish links with other gregarious nodes. Thus, while in the first case hubs act principally as information bridges, acquiring tokens of information at the periphery of the network and then transmitting to the rest of the network, in the second case hubs impose their opinion more easily.

2.2.2 Modeling opinion formation in social networks

Statistical mechanics and the social sciences share a common point of view: understanding how macroscopic structures emerge from the microscopic interactions among atomic elements. The term 'consensus' is used in sociology to express that an opinion is shared by the majority of the population; in the physical sciences this translates to the term "ordered state" or symmetry-breaking. Although the physics of spin particles or diffusion cannot completely describe the evolution of opinion dynamics, it provides a qualitative and quantitative description of the phenomena [9]. Everyday life indicates that people are sometimes confronted with a limited number of positions on a specific issue, which often are as few as two: right/left, Windows/Linux, buying/selling, etc. From this point of view, an agent can be described as a spin particle where the rules of spin alignment are specific to the specific problem.

Different models of opinion formation have been developed. The main assumption is that an agent can change his own opinion depending on the current state of the system, and the stochasticity of the system reflects the fuzziness of individuals. The classical Ising spin model has been used extensively to model the dynamics of opinions [21]. In this case the spin state represents a particular personal opinion, such as voting for either the Republicans or the Democrats. The local magnetic field of an individual is felt by its nearest neighbors, representing the tendency of an agent to pursuade his friends to adopt the same opinion. Optionally, an external magnetic field represents the cultural bias or propaganda which equally affects all agents in the system. Depending on the strength of propaganda and pairwise interactions, the system may evolve towards a consensus state or remain in a disordered state.

One of the first models of opinion dynamics is the voter model [39]. In this model agents are endowed with a binary variable $s_i(= \pm 1)$. At each time step an agent i is chosen together with one of his neighbor and the agent copies the state of his neighbor. Agent fell the pressure of the peers in average sense and fluctuations are expected. The question is whether full consensus can be reached in a system of infinite dimensions. The analysis of these models has shown that in the case of scale free networks, hubs reach immediately the final state and then, depending on the size and the exponent γ the system reach consensus in few steps. Several modifications have been included to the models, to take account of different social effects, like for example, modulating the flipping probability according to the majority state of the neighbors, or introducing a random flipping of the spin state as a proxy for introduction of novelty in the society [20].

Topology plays an important role in reaching consensus. Considering the case of scale free networks, the threshold temperature value, change according to exponent γ and diverges with the size of the network [2,27]. In the case of voter model, the time to reach consensus can scale with the size N if the exponent $\gamma > 3$ and is independent of N otherwise [39].

2.2.3 Modeling diffusion of innovation in social networks

The problem of innovation diffusion is a central argument in sociology: what are the acharacteristics of the social network, its innovators and the novelty of ideas that determine whether a new practice will be adopted by the entire population? Classical studies by Ryan and Gross [38] on hybrid corn seeds and Coleman, Katz and Menzel [11] on the adoption of new medicine among doctors have underlined some of the common characteristics of the innovation diffusion and adoption: the process of innovation adoption is slow; innovations are costly and then difficult to adopt immediately; decisions about adoption were made in a context where individuals can observe what the other were doing.

Rogers [37] went one step further: while analyzing different processes of novelty adoption, from using preservatives in South East Asia to adoption of boiling water among Peruvian villagers, he revealed the presence of different actors in the process and their position in the social network. Rogers' mechanisms have been mathematically formulated in the the Bass model, in its deterministic![6] and stochastic [36] variants.

Since the different roles in the process correspond to different positions in the network, the topology can influence the spread of the new habits. A convenient way to model the innovation diffusion on a network is using cooperative games. Suppose that at the beginning only few nodes have adopted the novelty (nodes A) while the rest of the population hasn't (nodes B). In this situation some nodes could be at the interface between nodes A and B, having a certain fraction of links p_A towards innovators and $p_B = 1 - p_A$ towards resistant ones. The node under observation can decide to switch to the novelty if it could perceive the positive effect. Quantifying the possible earning by q_A for each link established between with node of type A, and q_B with nodes B, the node will adopt the novelty if:

$$p_A q_A \geq p_B q_B \tag{1}$$

thus fixing a condition on the fraction of adopters necessary for triggering the innovation diffusion. Novelty is most of the cases introduced in a population by an *innovator*, somebody at the periphery of community with some connections toward the exterior of the community, then capable of retrieving information that other individuals can't. The novelty can initially spread among innovators' peers (*early adopters*). Although highly advantageous, the novelty is also costly, so only when the novelty is adopted by *opinion leaders* (people trusted by the community) the innovation can spread through all the society. Using the terminology of social network, early adopters correspond to nodes that can retrieve a lot of information from other communities (weak links). However this informational advantage is not an immediate help for the adoption of one practice: in order to be adopted, a novelty should be tested by other individuals. On the other hand, opinion leaders represent nodes with strong ties with other members of the communities, mostly they are in the core of the network, end they are highly clustered among them: they aren't the most connected ones, but they represent the most preeminent. Similar results has been found by Centola [10] and Valente [40], in studying the diffusion of habits among american adolescents.

3 Information storage and transmission: An example

We illustrate the concepts of information storage, transfer and dissipation using a one-dimensional system of coin flips. Suppose that an array of fair coins are numbered S_1, S_2, S_3, and S_4. First we flip S_1, then S_2, then S_3 and then S_4, independently of the previous outcome. We are interested in calculating how much the state of S_1, i.e., heads or tails, influences the state of the subsequent states. We do this by calculating how much information is stored in the state of each coin about S_1, using the definitions of Shannon entropy H in the form of mutual information I [12]:

$$H(X) = -\sum_{x \in X} p(x) \log_2 p(x) \tag{2}$$

$$H(X|Y) = -\sum_{y \in Y} p(y) \sum_{x \in X} p(x|y) \log_2 p(x|y) \tag{3}$$

$$I(X|Y) = H(X) - H(X|Y). \tag{4}$$

Firstly, the state of S_1 trivially stores complete information about itself, namely 1 bit (heads or tails). The interpretation is that the state of S_1 fully determines its own state. The second coin, on the other hand, was not influenced by the outcome of the first coin and therefore stores 0 bits of information about the state of S_1. The same is true for S_3 and S_4. This is depicted in Figure 2(a). In other words, the information stored at S_1 is not transmitted.

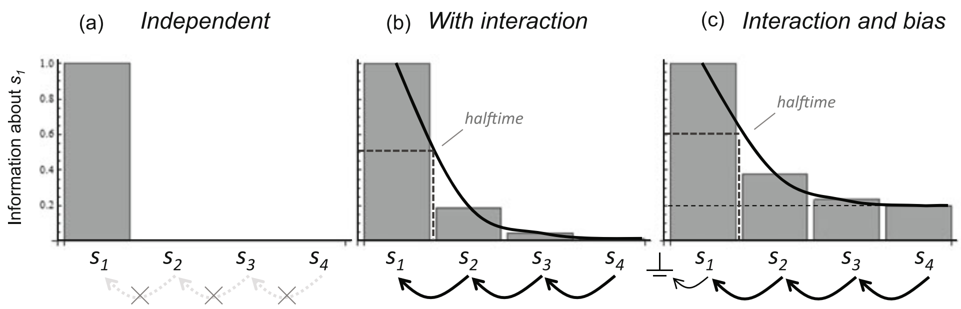

Fig. 2. Three forms of information transfer in the coins example. (a) If there is no interaction between the coin flips, then no information about S_1 is transferred. (b) If a stochastic interaction is introduced – here a 75%-25% chance of reproducing the outcome of the previous coin flip – information is transferred. The state of S_2 depends on the state of S_1, so the state of S_2 partially reflects the state of S_1. This information about S_1 is stored in S_2 and is quantified here using Eq. (4). (c) If S_1 has a structural bias towards heads, then each coin state already has 'static' information about the state of S_1 to begin with. Information due to interaction, or 'dynamic' information, is additional to the static information. The halftime of dynamic information is unaffected by the presence of static information.

3.1 Information transfer

Suppose now that each coin flip tends to reproduce the outcome of the previous coin flip. Let us assume that each coin flip succeeds in reproducing a heads or tails with a 75% probability, while the first coin still flips randomly. This introduces a direct influence between the states of subsequent coins which we can reveal in the same manner as before.

The first coin still stores the maximum 1 bit of information about itself. S_2 can now use its own state, either 'heads' (↑) or 'tails' (↓), to partially reconstruct the state of S_1 using Bayes' rule:

$$P(S_1 = \uparrow \mid S_2 = \uparrow) = \frac{P(S_2 = \uparrow \mid S_1 = \uparrow) \cdot P(S_1 = \uparrow)}{P(S_2 = \uparrow)},$$

and

$$P(S_1 = \downarrow | S_2 = \uparrow) = 1 - P(S_1 = \uparrow \mid S_2 = \uparrow).$$

In this example, S_2 infers that the probability that S_1 has the same state is simply 75%. So the ability to reconstruct the state of S_1 by S_2 has improved from $50\% - 50\%$ to $75\% - 25\%$. We can quantify this reconstruction in a number of bits using the formula of mutual information:

$$\begin{aligned} I(s_1|s_2) = &- \sum_{X \in \{\uparrow,\downarrow\}} P(s_1 = X) \log_2 P(s_1 = X) \\ &- \sum_{X \in \{\uparrow,\downarrow\}} P(s_1 = X|s_2 = \uparrow) \log_2 P(s_1 = X|s_2 = \uparrow) \\ &\approx 0.19. \end{aligned}$$

This means that 19% of the state of S_2 is actually a reflection of the state of S_1; the remaining 81% is still noise, or randomness, as before.

S_3 did not interact directly with S_1, but he did interact with S_2. He can partially reconstruct the state of S_2, which in turn can partially reconstruct the S_1. S_3, given his own state, can use Bayes' rule to find that the probability that S_1 has the same state as S_3 is $0.75^2 + 0.25^2 = 0.68$. Using the formula of mutual information we find that approximately 4.6% of the information stored in the state of S_3 originated from S_1. Similarly, S_4 stores approximately 1.1% of the information about S_1. This is shown in Figure 2(b).

3.2 Static and dynamic information

The final change that we discuss for the coin example is the use of an unfair coin. Suppose that the first coin result in "heads" with a probability of 75% and "tails" with 25% probability. This introduces additional information about the first outcome. See Figure 2(c).

Unlike the previous information, this additional information is not transferred through interactions. It would be present even if all coin flips were completely independent, and it does not change with dynamical behavior. Therefore we refer to this information as static. The previous information, in contrast, was transferred by dynamical interactions and is therefore dynamic.

This means that explaining the observed dynamical behavior of a physical system requires distinguishing the dynamical information from the static information. An example characteristic of the processing of dynamical information is the information dissipation length, which we discuss next.

3.3 Information dissipation length

It is clear that some information is lost during each transfer. In other words, information dissipates. This is caused by imperfect interactions where the state of an element only partially depends on another element. In the coin example, roughly 81% of the outcome of a coin flip consists of noise due to the randomness in flipping a coin. This noise combines until eventually all information about the first outcome is lost. The *information dissipation length* can be computed and is a measure of the size of the subsystem that is affected by a particular element.

Typically the amount of information about an element never becomes truly zero and approximates an exponential decay. Therefore we compute the time it takes to lose 50% of the information, analogous to the halftime of the decay of radioactive particles. The rate of losing information in the example is

$$\frac{I(S_1|S_n)}{I(S_1|S_{n+1})} = \frac{H(S_1) - H(S_1|S_n)}{H(S_1) - H(S_1|S_n + 1)} \tag{5}$$

$$= \frac{\left(\frac{1}{2} - \frac{1}{2^n}\right)\log\left(\frac{1}{2} - \frac{1}{2^n}\right) + \left(\frac{1}{2^{n+1}}\right)\log\left(\frac{1}{2^{n+1}}\right) - \log 2}{\left(\frac{1}{2} - \frac{1}{2^{n+1}}\right)\log\left(\frac{1}{2} - \frac{1}{2^{n+1}}\right) + \left(\frac{1}{2} + \frac{1}{2^{n+1}}\right)\log\left(\frac{1}{2} + \frac{1}{2^{n+1}}\right) - \log 2}, \tag{6}$$

which is independent of the bias of S_1. This rate is constant except for a small deviation for the lowest n. We find the exact rate by taking the limit $n \to \infty$:

$$\lim_{n\to\infty} \frac{\left(\frac{1}{2} - \frac{1}{2^n}\right)\log\left(\frac{1}{2} - \frac{1}{2^n}\right) + \left(\frac{1}{2^{n+1}}\right)\log\left(\frac{1}{2^{n+1}}\right) - \log 2}{\left(\frac{1}{2} - \frac{1}{2^{n+1}}\right)\log\left(\frac{1}{2} - \frac{1}{2^{n+1}}\right) + \left(\frac{1}{2} + \frac{1}{2^{n+1}}\right)\log\left(\frac{1}{2} + \frac{1}{2^{n+1}}\right) - \log 2} = 4. \tag{7}$$

In other words, every subsequent coin flip reduces the amount of information about the state of S_1 by a factor 4 in the example. The characteristic halftime of the information is then

$$\frac{\log \frac{1}{2}}{\log \frac{1}{4}} = 0.5, \tag{8}$$

which is depicted in Figure 2.

A related quantity is information dissipation *time*, which is the time it takes for information about a coin's state to be forgotten by the network. In this example it equals the dissipation length if we interpret S_i as the outcome of flipping a single coin repeatedly at fixed time intervals.

4 Information processing in social networks

We use the concepts of information storage and transfer to explain a counterintuitive observation in social networks: persons with many friends (connections) do not dictate the observed dynamics of opinion forming [3,28]. In a social network, the opinions of individuals change over time continuously. A change of opinion of one person may cause a change of opinion of friends, who may in turn cause changes of opinion among the friends of friends, and so on.

We use the prevailing model of Ising-spin particles connected by a network with power-law degree distribution. Each particle can be "up" or "down" and represents the opinion of one person; each connection in the network represents a social interaction. The size of our system is 6000 spins and the degree distribution of the network is set to $p(k) \propto k^{-1.6}$. Our results do not change qualitatively with the size and power-law exponent. The temperature T of the system is a free parameter which determines the sensitivity of a person's opinion based on the opinions of his neighbors. We simulate the continuous change of opinions using the Metropolis algorithm [34].

4.1 Measuring dynamical influence: Information dissipation time

Our goal is to decide which nodes in the network are responsible for the observed dynamics of individual opinions. For most systems, the information dissipation length as described above would be an intuitive measure of dynamical influence of a node. But the diameter of scale-free networks is very small, on the order of $\ln \ln N$, so a measure of spatial distance is not applicable in this case.

A related concept that we will use is information dissipation time. Instead of measuring the total information storage about a node as function of distance, we measure it as function of time. More precisely, we measure the amount of information that the social network $\mathbb{N}$ stores as a whole about the state s_i of an individual node as function of time t, i.e., $I(s_i|\mathbb{N}(t))$. The information dissipation time can be interpreted as how long the network remembers the instantaneous states of a particular node, which is a measure of its dynamical influence. See Figure 3.

We measure how much information is stored in the network about each individual node as follows. At random intervals during simulation we take a snapshot of the state of the network, and for each snapshot we sample many state trajectories that lead to it. Observing only the network state at a particular time, each state trajectory represents one possible history of individual states over time. The probability of having opinion 1 for each individual i at each previous time step $t - \delta$ can be calculated simply by counting the fraction of state trajectories that have $s_i(t - \delta) = 1$.

We sample 90,000 state trajectories according to their likelihood using a Markov-Chain Monte Carlo algorithm. Starting at a snapshot $\mathbb{N}(t_0)$ we select a candidate

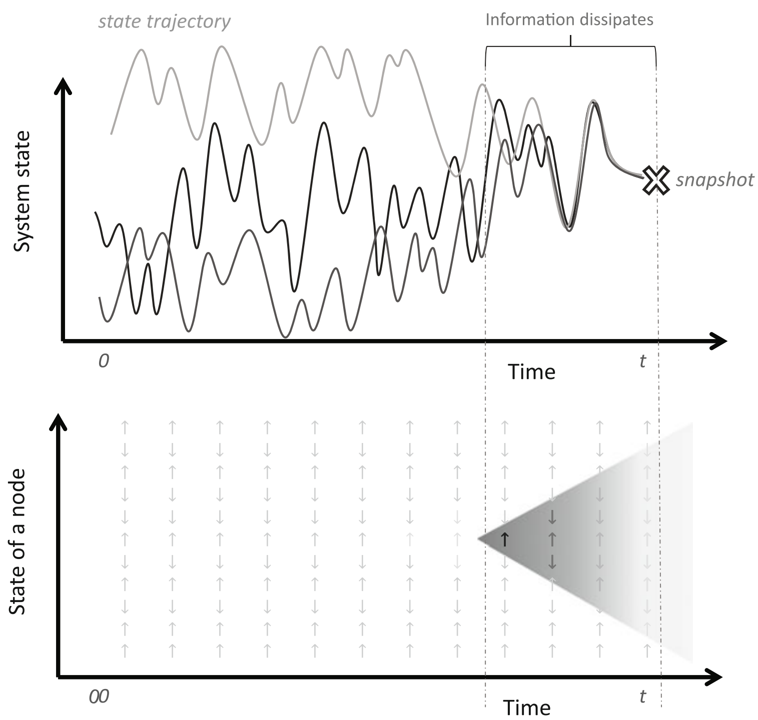

Fig. 3. From a global point of view, the state of a system stores information and loses it over time. This implies that a system state contains information about recent states only (upper panel). From a local point of view, the information stored in each single element dissipates, i.e., diffuses through neighbors and neighbors-of-neighbors while it diminishes at the same time (lower panel).

preceding network state using Bayes' rule:

$$P(\mathbb{N}(t_0-1) \mid \mathbb{N}(t_0)) = P(\mathbb{N}(t_0) \mid \mathbb{N}(t_0-1))\frac{\mathbb{N}(t_0-1)}{\mathbb{N}(t_0)}. \tag{9}$$

In this way we recursively select preceding network states until the candidate state trajectory is long enough for estimating the information dissipation time. This algorithm samples each possible state trajectory with a probability proportional to how likely it is, given the rules of dynamics.

For each spin we calculate its information dissipation time using the set of state trajectories. At each time point t we count how many times the spin was observed in the $\uparrow$-state and in the $\downarrow$-state. This yields an estimated probability distribution $p(s_i(t) =\uparrow)$ of the i'th spin state as function of time $t_0 - t$ leading up to the snapshot. Of the corresponding amount of information $I(s_i(t)| \ \mathbb{N}(t))$ we estimate the rate of dissipation by fitting an exponentially decaying curve $1 - b + b \cdot e^{-a \cdot t}$. Here, $1 - b$ is the static information and e^{-a} is the rate of decay of the dynamic information. The information dissipation length is then $-\log \frac{1}{2}/a$.

5 Results

In Figure 4 we first show the static information of each spin in the network as function of its degree. This is the equilibrium information $I(s_i| \ \mathbb{N}, T)$ about each spin, which

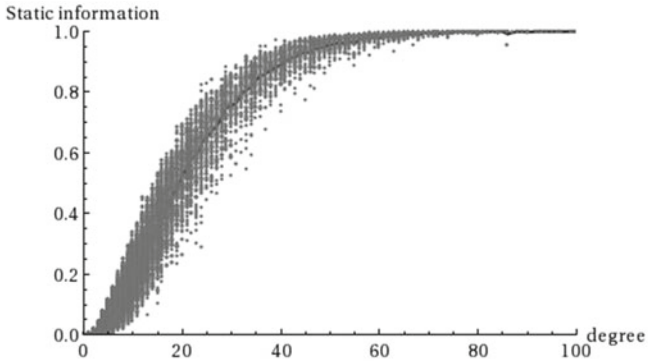

Fig. 4. The static information about the state of each spin as function of its degree. This shows how much is known about a spin only by observing the static network topology and temperature. It is therefore the minimum information about each spin that could be stored in the dynamic network state. The temperature was set to 20.

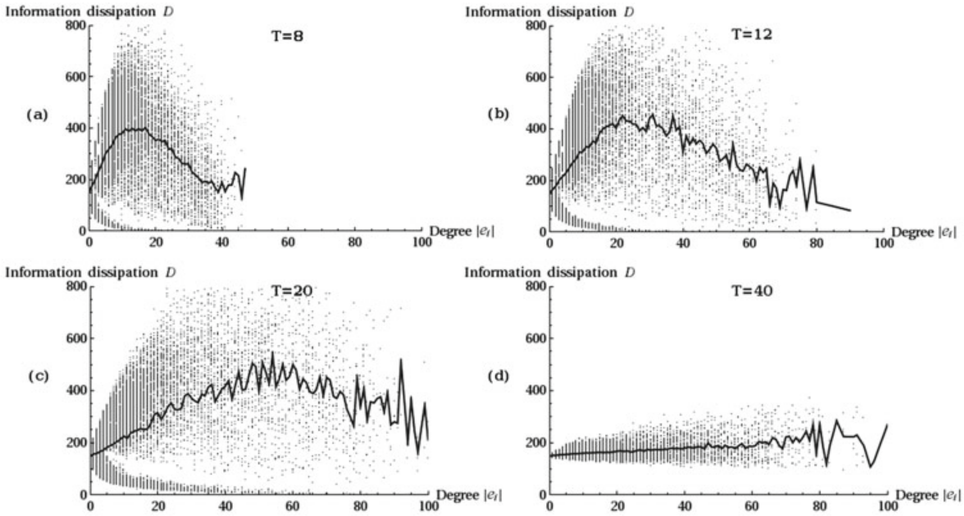

Fig. 5. The dissipation time of the dynamic information about the state of each spin as function of its degree. This shows how long it takes for dynamic information about a spin's state to be lost. The solid line is the average dissipation time per degree. Each plot consists of more than 35000 data points, generated by six simulations; the solid curve connects the average for each degree. Nodes with a flip probability of less than 0.01% (1.5% of the nodes) were ignored because their state probability over time could not be reliably estimated. The distinct lower cluster of dissipation times pertain to nodes which were incidentally in a high energy state given their local topology and the states of the neighbors. This high energy state is transient and due to fluctuations so the information about this state is forgotten quickly.

is determined solely by its position in the network and the temperature. It is clear that the state of high-degree spins is 'known in advance', while the state of low-degree nodes is transient and uncertain.

Additional information about the state of each spin comes from the states of other spins and their changes. This dynamic information is the total information about a spin $I(s_i(t)|\ \mathbb{N}(t_0), T)$ given the instantaneous states of all spins, minus the static information $I(s_i|\,\mathbb{N}, T)$. It is the amount of information that is stored in the instantaneous spin states in the network beyond what is already known due to the static network topology and temperature. We show the measured dissipation time of this dynamic information for each spin as function of its degree in Figure 5.

Surprisingly, hubs do not dictate the observed dynamics of the spin states in the network. To explain the behavior of the spin states over time in the network, intermediately connected nodes provide more information than hubs or peripheral nodes. Although a hub restricts the equilibrium state space more than any other node (Fig. 4), a specific instance of its internal state is apparently irrelevant and soon forgotten by the system.

One possible interpretation is the following. The state of high-degree nodes is insensitive to change, i.e., is dominated by determinism, so most information that it transmits is already present in the network. On the other hand, states of lowest-degree nodes are sensitive to change, i.e., are dominated by thermal noise, so their information storage is minimal: they mostly transmit noise. It is apparently a trade-off between randomness and determinism that optimizes the percolation of information through a system.

Another observation is that the optimum number of connections k_{max} increases as the temperature increases. This supports our interpretation of a trade-off between randomness and determinism. Sparsely connected spins become more random, losing their ability to transmit information, whereas hubs become more dynamic and start to transmit information.

6 Conclusions

We have shown that information theory can be used to discover how the dynamical influence of each element percolates through a system of coupled units. We find that the dynamics of opinions in a model social network with a power-law degree distribution are determined mostly by intermediately connected persons. In order to explain how opinions change in a social network it would be most informative to observe the opinions of these persons. This finding is counterintuitive from the perspective of the network topology, which suggests that a person with many connections is capable of influencing many friends simultaneously and would therefore be most informative. From the perspective of processing information, however, we find that hubs transmit redundant information and peripheral nodes transmit noise. We believe that the concept of information processing in dynamical systems can improve our understanding of the behavior of complex systems in general.

We acknowledge the financial support of DynaNets within the Seventh Framework Program for Research of the European Commission, under FET-Open grant number: 233847. PMAS also acknowledges the support from a grant from the "Leading Scientist Program" of the Government of the Russian Federation, under contract 11.G34.31.0019 and the support from the FET-Proactive grant TOPDRIM, number FP7-ICT-318121 as well as the Complexity program of the NTU in Singapore.

References

1. R. Albert, H. Jeong, A.-L. Barabasi, Nature **406**, 378 (2000)
2. A. Aleksiejuk, J.A. Holyst, D. Stauffer, Physica A **310**, 260 (2002)
3. S. Aral, D. Walker, Science, **337**, 337 (2012)
4. R.P. Bagozzi, U.M. Dholakia, J. Interactive Marketing **16**, 2 (2002)

5. A.-L. Barabasi, R. Albert, Emergence of scaling in random networks [eprint arXiv:cond-mat/9910332], October (1999)
6. F. Bass, Management Sci. **15**, 215 (1969)
7. C.H. Bennett, Dissipation, information, computational complexity and the definition of organization, edited by David Pines, Emerging Syntheses in Science, (Reading, MA, 1987), p. 215
8. C.H. Bennett, Stud. History Philosophy Mod. Phys. **34**, 501 (2003)
9. C. Castellano, S. Fortunato, V. Loreto, Rev. Mod. Phys. **81**, 591 (2009)
10. D. Centola, M. Macy, Amer. J. Sociol. **113**, 702 (2007)
11. J.S. Coleman, E. Katz, H. Menzel, Medical innovation: A Diffusion Study, Vol. 46 (Bobbs-Merrill, 1966)
12. T.M. Cover, J.A. Thomas, *Elements of Information Theory*, Vol. 6 (Wiley-Interscience, 1991)
13. J.P. Crutchfield, C.J. Ellison, J.R. Mahoney, Phys. Rev. Lett. **103**, 094101 (2009)
14. J.P. Crutchfield, D.P. Feldman, Phys. Rev. E **55**, R1239 (1997)
15. J.P. Crutchfield, D.P. Feldman, Chaos **13**, 25 (2003)
16. U.M. Dholakia, R.P. Bagozzi, L.K. Pearo, Int. J. Res. Marketing **21**, 241 (2004)
17. T.L. Duncan, J.S. Semura, Found. Phys. **37**, 1767 (2007)
18. W. Ebeling, G. Nicolis, Chaos, Solitons Fractals **2**, 635 (1992)
19. P. Erdös, A. Rényi, Publicationes Mathematicae (Debrecen) **6**, 290 (1959)
20. Z. Fagyal, Social Networks **15** (2009)
21. S. Galam, S. Moscovici, Eur. J. Social Psychol. **21**, 49 (1991)
22. F. Geier, J. Timmer, C. Fleck, BMC Systems Biol. **1**, 11 (2007)
23. J.-D.J. Han, N. Bertin, T. Hao, D.S. Goldberg, G.F. Berriz, L.V. Zhang, D. Dupuy, A.J.M. Walhout, M.E. Cusick, F.P. Roth, M. Vidal, Nature **430** (2004)
24. E.T. Jaynes, The Phys. Rev. **106**, 620 (1957)
25. R. Landauer, IBM J. Res. Develop. **5**, 183 (1961)
26. R. Landauer, Phys. Lett. A **217**, 188 (1996)
27. M. Leone, A. Vázquez, A. Vespignani, R. Zecchina, Eur. Phys. J. B **28**, 191 (2002)
28. J. Leskovec, L.A. Adamic, B.A. Huberman, ACM Trans. Web **1** (2007)
29. Y.-Y. Liu, J.-J. Slotine, A.-L. Barabasi, Nature **473**, 167 (2011)
30. S. Lloyd, Nature **406**, 1047 (2000)
31. S. Lloyd, Programming the Universe: A Quantum Computer Scientist Takes On the Cosmos. Knopf, 2006
32. B. Mashhoon, Found. Phys. **15**, 497 (1985)
33. B. Mashhoon, D. Theiss, Relativistic effects in the motion of the moon, in *Gyros, Clocks, Interferometers: Testing Relativistic Gravity in Space*, edited by Claus Lämmerzahl, C. Everitt, and Friedrich Hehl, Lecture Notes in Physics, Vol. 562 (Springer Berlin / Heidelberg, 2001), p. 310
34. N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, E. Teller, J. Chem. Phys. **21**, 1087 (1953)
35. M.E.J. Newman, SIAM Rev. **45**, 167 (2003)
36. S.-C. Niu, Math. Probl. Eng. **8**, 249 (2002)
37. E.M. Rogers, Diff. Innovations (Free Press, 2003)
38. B. Ryan, N.C Gross, Rural Sociology **8**, 15 (1943)
39. V. Sood, T. Antal, S. Redner, Phys. Rev. E **77**, 1 (2008)
40. T.W. Valente, R.L. Davis, The Annals of the American Academy (1999)
41. K. Wiesner, Chaos **20**, 037114 (2010)